



(12)发明专利

(10)授权公告号 CN 104239552 B

(45)授权公告日 2017.09.01

(21)申请号 201410494326.4

(22)申请日 2014.09.24

(65)同一申请的已公布的文献号

申请公布号 CN 104239552 A

(43)申请公布日 2014.12.24

(73)专利权人 北京百度网讯科技有限公司

地址 100085 北京市海淀区上地十街10号

百度大厦2层

(72)发明人 门聪

(74)专利代理机构 北京金律言科知识产权代理

事务所(普通合伙) 11461

代理人 罗延红 逯博

(51)Int.Cl.

G06F 17/30(2006.01)

(56)对比文件

CN 103258025 A, 2013.08.21,

CN 102110126 A, 2011.06.29,

CN 103927354 A, 2014.07.16,

CN 103258025 A, 2013.08.21,

US 2008/0071740 A1, 2008.03.20,

审查员 李艳霞

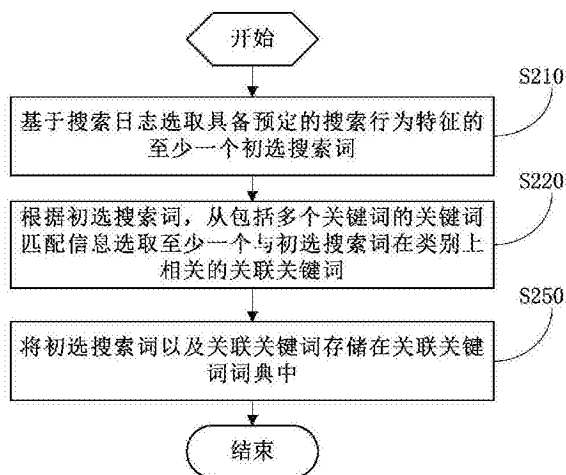
权利要求书4页 说明书7页 附图3页

(54)发明名称

生成关联关键词、提供关联关键词的方法及系统

(57)摘要

本发明提供一种生成关联关键词、提供关联关键词的方法及系统。所述生成关联关键词的方法包括：基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词；根据所述初选搜索词，从包括多个关键词的关键词匹配信息选取至少一个与初选搜索词在类别上相关的关联关键词；将所述初选搜索词以及所述关联关键词存储在关联关键词词典中。所述生成关联关键词、提供关联关键词的方法及系统为指定类目下的搜索词选取在类别上相关的关联关键词，从而当接收到包括所述搜索词的搜索请求时，可将所述关联关键词提供给用户，为用户提供上下级类别的推荐条目。



1. 一种生成关联关键词的方法,所述方法包括:

基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词,包括:对搜索日志中的任一搜索词,计算其以下第一搜索行为参数:所述搜索词是否包含预定的主类目关键词,所述搜索词是否包含预定的子类目关键词,所述搜索词的搜索结果中包含所述主类目关键词的搜索结果被查看的第一频率,和所述搜索词的搜索结果中包含所述子类目关键词的搜索结果被查看的第二频率;

根据所述初选搜索词,从包括多个关键词的关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词;

将所述初选搜索词以及所述关联关键词存储在关联关键词词典中。

2. 根据权利要求1所述的方法,其特征在于,所述第一搜索行为参数还包括:

在使用所述搜索词搜索后使用包含所述主类目关键词的其他搜索词进行搜索的第三频率,和

在使用所述搜索词搜索后使用包含所述子类目关键词的其他搜索词进行搜索的第四频率。

3. 根据权利要求1或2所述的方法,其特征在于,所述基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词的处理还包括:

对所述任一搜索词的所述第一搜索行为参数进行加权求和,计算所述搜索词的第一需求强度值,

如果所述搜索词的第一需求强度值超过第一预定需求强度阈值,则将所述搜索词选取为初选搜索词。

4. 根据权利要求3所述的方法,其特征在于,所述关键词匹配信息还包括所述多个关键词的价目。

5. 根据权利要求4所述的方法,其特征在于,所述根据所述初选搜索词,从所述关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词的处理包括:

对任一初选搜索词,在关键词匹配信息中查找与所述初选搜索词匹配的关键词及其价目,如果查找到的价目超过预定的价目阈值,则将匹配到的关键词匹配信息中的关键词选取为关联关键词。

6. 根据权利要求5所述的方法,其特征在于,所述方法还包括:

根据所述搜索日志和所述关联关键词,为所述初选搜索词生成包括所述关联关键词及其第二需求强度值的关键词需求列表;

根据所述关键词匹配信息以及第二需求强度值,对为所述初选搜索词生成的关键词需求列表中的关联关键词排序,并且

所述将所述初选搜索词以及所述关联关键词存储在关联关键词词典中的处理包括:将所述初选搜索词及其关键词需求列表存储在关联关键词词典中。

7. 根据权利要求6所述的方法,其特征在于,所述根据所述搜索日志和所述关联关键词,为所述初选搜索词生成包括所述关联关键词及其第二需求强度值的关键词需求列表的处理包括:

为任一所述初选搜索词,基于每个关联关键词的以下第二搜索行为参数计算所述关联关键词的第二需求强度值:

所述初选搜索词是否包含所述关联关键词，

所述初选搜索词的搜索结果中包含所述关联关键词的搜索结果被查看的第五频率，和
在使用所述初选搜索词搜索后使用包含所述关联关键词的其他搜索词进行搜索的第六频率，

将所述关联关键词及其第二需求强度值插入所述初选搜索词的关键词需求列表。

8. 根据权利要求7所述的方法，其特征在于，所述为任一所述初选搜索词，基于每个关联关键词的以下第二搜索行为参数计算所述关联关键词的第二需求强度值包括：

对计算的所述第二搜索行为参数进行加权求和，为所述初选搜索词求得所述关联关键词的第二需求强度值。

9. 根据权利要求7所述的方法，其特征在于，所述根据所述关键词匹配信息以及第二需求强度值，对为所述初选搜索词生成的关键词需求列表中的关联关键词排序的处理包括：

对为所述初选搜索词生成的关键词需求列表中的关联关键词，分别从所述关键词匹配信息提取其价目，对所述关联关键词的价目以及第二需求强度值进行加权求和，求得所述关联关键词的综合评估值，

根据所述综合评估值，对为所述初选搜索词生成的关键词需求列表中的关联关键词进行排序。

10. 根据权利要求9所述的方法，其特征在于，所述第一搜索行为参数还包括：

搜索词排除参数：所述任一搜索词是否包含预定的排除关键词，并且

在所述对所述任一搜索词的所述第一搜索行为参数进行加权求和的处理中，为所述搜索词排除参数赋予负的权值。

11. 一种生成关联关键词的系统，所述系统包括：

搜索词初选模块，用于基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词，所述搜索词初选模块用于对搜索日志中的任一搜索词，计算其以下第一搜索行为参数：所述搜索词是否包含预定的主类目关键词，所述搜索词是否包含预定的子类目关键词，所述搜索词的搜索结果中包含所述主类目关键词的搜索结果被查看的第一频率，和所述搜索词的搜索结果中包含所述子类目关键词的搜索结果被查看的第二频率；

关联关键词选取模块，用于根据搜索词初选模块选取的初选搜索词，从包括多个关键词的关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词；

词典管理装置，用于将搜索词初选模块选取的初选搜索词以及关联关键词选取模块选取的关联关键词存储在关联关键词词典中。

12. 根据权利要求11所述的系统，其特征在于，所述第一搜索行为参数还包括：

在使用所述搜索词搜索后使用包含所述主类目关键词的其他搜索词进行搜索的第三频率，和

在使用所述搜索词搜索后使用包含所述子类目关键词的其他搜索词进行搜索的第四频率。

13. 根据权利要求11或12所述的系统，其特征在于，所述搜索词初选模块还用于对所述任一搜索词的所述第一搜索行为参数进行加权求和，计算所述搜索词的第一需求强度值，如果所述搜索词的第一需求强度值超过第一预定需求强度阈值，则将所述搜索词选取为初选搜索词。

14. 根据权利要求13所述的系统,其特征在于,所述关键词匹配信息还包括所述多个关键词的价目。

15. 根据权利要求14所述的系统,其特征在于,所述关联关键词选取模块用于对任一初选搜索词,在关键词匹配信息中查找与所述初选搜索词匹配的关键词及其价目,如果查找到的价目超过预定的价目阈值,则将匹配到的关键词匹配信息中的关键词选取为关联关键词。

16. 根据权利要求15所述的系统,其特征在于,所述系统还包括:

关键词需求列表生成模块,用于根据所述搜索日志和所述关联关键词,为所述初选搜索词生成包括所述关联关键词及其第二需求强度值的关键词需求列表;

关键词需求列表排序模块,用于根据所述关键词匹配信息以及第二需求强度值,对为所述初选搜索词生成的关键词需求列表中的关联关键词排序,并且

所述词典管理装置用于将所述初选搜索词及其关键词需求列表存储在关联关键词词典中。

17. 根据权利要求16所述的系统,其特征在于,关键词需求列表生成模块用于为任一所述初选搜索词,基于每个关联关键词的以下第二搜索行为参数计算所述关联关键词的第二需求强度值,并且将所述关联关键词及其第二需求强度值插入所述初选搜索词的关键词需求列表:

所述初选搜索词是否包含所述关联关键词,

所述初选搜索词的搜索结果中包含所述关联关键词的搜索结果被查看的第五频率,和

在使用所述初选搜索词搜索后使用包含所述关联关键词的其他搜索词进行搜索的第六频率。

18. 根据权利要求17所述的系统,其特征在于,关键词需求列表生成模块在所述为任一所述初选搜索词,基于每个关联关键词的以下第二搜索行为参数计算所述关联关键词的第二需求强度值的处理中,对计算的第二搜索行为参数进行加权求和,为所述初选搜索词求得所述关联关键词的第二需求强度值。

19. 根据权利要求17所述的系统,其特征在于,所述关键词需求列表排序模块,用于对为所述初选搜索词生成的关键词需求列表中的关联关键词,分别从所述关键词匹配信息提取其价目,对所述关联关键词的价目以及第二需求强度值进行加权求和,求得所述关联关键词的综合评估值,并且根据所述综合评估值,对为所述初选搜索词生成的关键词需求列表中的关联关键词进行排序。

20. 根据权利要求19所述的系统,其特征在于,所述第一搜索行为参数还包括:

搜索词排除参数:所述任一搜索词是否包含预定的排除关键词,并且

所述搜索词初选模块在所述对所述任一搜索词的所述第一搜索行为参数进行加权求和的处理中,为所述搜索词排除参数赋予负的权值。

21. 一种提供关联关键词的方法,所述方法包括:

接收第一搜索词;

根据所述第一搜索词从关联关键词词典提取至少一个关联关键词,所述关联关键词词典包括第二搜索词及至少一个关联关键词,所述关联关键词词典中的至少一个关联关键词的生成处理包括:对搜索日志中的任一搜索词,计算其以下第一搜索行为参数:所述搜索词

是否包含预定的主类目关键词,所述搜索词是否包含预定的子类目关键词,所述搜索词的搜索结果中包含所述主类目关键词的搜索结果被查看的第一频率,和所述搜索词的搜索结果中包含所述子类目关键词的搜索结果被查看的第二频率;

发送提取的关联关键词。

22. 根据权利要求21所述的方法,其特征在于,所述至少一个关联关键词按照其需求强度排列。

23. 一种提供关联关键词的系统,所述系统包括:

搜索请求接收装置,用于接收第一搜索词;

搜索请求处理装置,用于根据搜索请求接收装置接收的第一搜索词从关联关键词词典提取至少一个关联关键词,所述关联关键词词典包括第二搜索词及至少一个关联关键词,所述关联关键词词典中的至少一个关联关键词的生成处理包括:对搜索日志中的任一搜索词,计算其以下第一搜索行为参数:所述搜索词是否包含预定的主类目关键词,所述搜索词是否包含预定的子类目关键词,所述搜索词的搜索结果中包含所述主类目关键词的搜索结果被查看的第一频率,和所述搜索词的搜索结果中包含所述子类目关键词的搜索结果被查看的第二频率;

推荐关键词发送装置,用于发送搜索请求处理装置提取的关联关键词。

24. 根据权利要求23所述的系统,其特征在于,所述至少一个关联关键词按照其需求强度排列。

生成关联关键词、提供关联关键词的方法及系统

技术领域

[0001] 本发明涉及数据处理技术,尤其涉及一种生成关联关键词、提供关联关键词的技术。

背景技术

[0002] 在搜索引擎技术中,除了提供搜索结果以外,通常搜索服务商还会在搜索结果网页的指定位置(如页面右侧)还提供一些与用户输入的搜索词相关的其他推荐信息。目前搜索服务商的推荐策略倾向于推荐同位的概念,对相关的泛需求推荐较少,并且推荐缺乏层次。

[0003] 图1示出在现有技术中搜索结果网页的示例。参照图1,用户输入“迅雷游戏”进行搜索后,在搜索结果网页的左侧显示如“迅雷游戏官网首页”、“迅雷游戏盒子”等结果条目。此外,在所述搜索结果网页的右侧还提供了如“相关游戏”类和“相关软件”类的同位概念条目。

[0004] 此外,搜索例如“封神无双”游戏的用户的主需求是该网页游戏,目前在搜索结果页面右侧推荐的策略是推荐同位的概念(如其他网页游戏),缺乏对其相关的上位概念(如网页游戏)的推荐,推荐维度不够丰富。

发明内容

[0005] 本发明的目的在于提供一种生成关联关键词的方法、提供关联关键词的方法和使用所述方法的系统,通过对搜索日志进行挖掘,提取与指定类目下的搜索词在类别上相关的更多关键词,以将这些关键词推荐给用户,增强用户体验。

[0006] 根据本发明的一方面,提供一种生成关联关键词的方法,所述方法包括:基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词;根据所述初选搜索词,从包括多个关键词的关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词;将所述初选搜索词以及所述关联关键词存储在关联关键词词典中。

[0007] 根据本发明的另一方面,提供一种生成关联关键词的系统,所述系统包括:搜索词初选模块,用于基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词;关联关键词选取模块,用于根据搜索词初选模块选取的初选搜索词,从包括多个关键词的关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词;词典管理装置,用于将搜索词初选模块选取的初选搜索词以及关联关键词选取模块选取的关联关键词存储在关联关键词词典中。

[0008] 根据本发明的另一方面,提供一种提供关联关键词的方法,所述方法包括:接收第一搜索词;根据所述第一搜索词从关联关键词词典提取至少一个关联关键词,所述关联关键词词典包括第二搜索词及至少一个关联关键词;发送提取的关联关键词。

[0009] 根据本发明的另一方面,提供一种提供关联关键词的系统,所述系统包括:搜索请求接收装置,用于接收第一搜索词;搜索请求处理装置,用于根据搜索请求接收装置接收的

第一搜索词从关联关键词词典提取至少一个关联关键词,所述关联关键词词典包括第二搜索词及至少一个关联关键词;推荐关键词发送装置,用于发送搜索请求处理装置提取的关联关键词。

[0010] 根据本发明的示例性实施例的生成关联关键词的方法、提供关联关键词的方法和使用所述方法的系统可以为指定类目下的搜索词选取在类别上相关的关联关键词,从而当接收到包括所述搜索词的搜索请求时,可将所述关联关键词提供给用户,为用户提供上下级类别的推荐条目。

[0011] 在此基础上,还可以为指定类目下的搜索词选取在类别上相关并且商业价值高的关联关键词,并且通过综合评估关联关键词的第二需求强度值和价目指标将综合评估值高的关联关键词排在前面,将综合价值高的关键词优先推荐给用户。

附图说明

[0012] 图1示出在现有技术中搜索结果网页的示例;

[0013] 图2是示出根据本发明的示例性实施例的生成关联关键词的方法的流程图;

[0014] 图3是示出根据本发明的另一示例性实施例的生成关联关键词的方法的流程图;

[0015] 图4是示出根据本发明的示例性实施例的提供关联关键词的方法的流程图;

[0016] 图5示出本发明实施例的生成关联关键词的方法和提供关联关键词的方法提供的搜索结果网页的示例;

[0017] 图6是示出根据本发明的示例性实施例的生成关联关键词的系统的逻辑框图;

[0018] 图7是示出根据本发明的示例性实施例的提供关联关键词的系统的逻辑框图。

具体实施方式

[0019] 本发明的总体构思是,通过对搜索日志进行分析,提取与指定类目相关的至少一个初选搜索词,并且通过与关键词匹配信息匹配选取从类别上与所述初选搜索词相关的关联关键词,从而将所述初选搜索词与所述关联关键词关联起来;在接收到初选搜索词的搜索请求时,提供所述关联关键词。

[0020] 通过为指定类目的搜索词提供关联关键词,能够为搜索用户提供更多上位的推荐信息,提高用户体验。

[0021] 下面结合附图对根据本发明的示例性实施例的生成关联关键词的方法、提供关联关键词的方法以及实现所述方法的系统进行详细描述。

[0022] 图2是示出根据本发明的示例性实施例的生成关联关键词的方法的流程图。

[0023] 参照图2,在步骤S210,基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词。

[0024] 根据本发明的示例性实施例,步骤S210可包括:对搜索日志中的任一搜索词,计算其多个第一搜索行为参数;对所述任一搜索词的所述第一搜索行为参数进行加权求和,计算所述搜索词的第一需求强度值;如果所述搜索词的第一需求强度值超过第一预定需求强度阈值,则将所述搜索词选取为初选搜索词。

[0025] 具体地,搜索词的所述第一搜索行为参数包括:所述搜索词是否包含预定的主类目关键词(参数1)、所述搜索词是否包含预定的子类目关键词(参数2)、所述搜索词的搜索

结果中包含所述主类目关键词的搜索结果被查看的第一频率(参数3)、所述搜索词的搜索结果中包含所述子类目关键词的搜索结果被查看的第二频率(参数4)。可选地,搜索词的所述第一搜索行为参数可还包括:在使用所述搜索词搜索后使用包含所述主类目关键词的其他搜索词进行搜索的第三频率(参数5),和在使用所述搜索词搜索后使用包含所述子类目关键词的其他搜索词进行搜索的第四频率(参数6)。

[0026] 这里,所述第一频率可以是,但不限于,例如,查看使用所述搜索词得到的搜索结果中标题包含所述主类目关键词的用户数量在使用所述搜索词进行搜索的全部用户数量中所占的百分比;所述第二频率可以是,但不限于,例如,查看使用所述搜索词得到的搜索结果中标题包含所述子类目关键词的用户数量在使用所述搜索词进行搜索的全部用户数量中所占的百分比;所述第三频率可以是,但不限于,例如,在使用所述搜索词搜索后使用包含所述主类目关键词的其他搜索词进行搜索的用户数量在使用所述搜索词进行搜索的全部用户数量中所占的百分比;所述第四频率可以是,但不限于,例如,在使用所述搜索词搜索后使用包含所述子类目关键词的其他搜索词进行搜索的用户数量在使用所述搜索词进行搜索的全部用户数量中所占的百分比。

[0027] 此外,对于所述搜索词是否包含预定的主类目关键词的第一搜索行为参数,可使用1或0来进行赋值;对于所述搜索词是否包含预定的子类目关键词的第一搜索行为参数,也使用1或0来进行赋值。

[0028] 以游戏类目为例,可预先设定主类目关键词为“游戏”,并且设定搜索频率高的“网页游戏”为子类目关键词。假设有100个用户使用搜索词“迅雷游戏”进行搜索,点击搜索结果中包含“游戏”的有30个用户,点击搜索结果中包含“网页游戏”的有20个用户,此后,又使用搜索词“网页游戏”进行搜索的有5个用户,则参数1的值为1,参数2的值为0,参数3的值为0.3,参数4的值为0.2,参数6的值为0.05。

[0029] 在所述加权求和的处理中,各第一搜索行为参数的权值可以是经验值;并且,可根据筛选要求或选取的初选搜索词的数量设置所述第一预定需求强度阈值。

[0030] 此外,搜索词中的一些关键词不利于初选搜索词的选取。例如,在包含“游戏”的搜索词中,如果还包含“电影”、“歌曲”等类别关键词,则无法确定所述搜索词是否实际与“游戏”相关。因此,根据本发明的优选实施例,所述第一搜索行为参数可还包括搜索词排除参数,即搜索词是否包含预定的排除关键词(可使用1或0进行赋值)。在此情况下,在所述对搜索词的所述第一搜索行为参数进行加权求和的处理中,为所述搜索词排除参数赋予负的权值,以降低包含预定的排除关键词的搜索词的第一需求强度值。

[0031] 通过将用户使用与所述主类目关键词和子类目关键词相关的搜索词进行搜索以及查看相关搜索结果的行为作为初选搜索词的筛选条件,可将搜索频率高的包含与所述主类目关键词和子类目关键词相关的搜索词筛选出来。例如,当通过主类目关键词“游戏”和搜索频率高的子类目关键词“网页游戏”对搜索日志进行分析时,可将在游戏类目下,同样经常被搜索的包括关键词“小游戏”、“桌面游戏”等的搜索词可被选取为初选搜索词。

[0032] 在执行步骤S210后,根据所述初选搜索词,从包括多个关键词的关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词(步骤S220)。

[0033] 具体地,将在所述关键词匹配信息当中包含在所述初选搜索词中的关键词选取为关联关键词,并且/或者将与所述初选搜索词存在类别上下级从属关系的关键词选取为关

联关键词。

[0034] 在步骤S250,将所述初选搜索词以及所述关联关键词存储在关联关键词词典中。

[0035] 通过步骤S210、S220和S250的处理,可以为指定类目下的搜索词选取在类别上相关的关联关键词,从而当接收到包括所述搜索词的搜索请求时,可将所述关联关键词提供给用户,为用户提供上下级类别的推荐条目。

[0036] 搜索服务商通常希望在所述指定位置提供具有潜在商业价值的信息。然而,高商业价值的搜索词往往较长且搜索频率不高,而用户倾向于使用一些简短的搜索词或需求表达不明确的搜索词。这就造成了很难通过搜索使用的搜索词准确地定位潜在商业需求很大的用户。可以通过在不损伤用户体验的情况下对用户搜索进行适当引导解决这个问题。

[0037] 图3是示出根据本发明的另一示例性实施例的生成关联关键词的方法的流程图。在图3中步骤S210的处理与图2中相应步骤的处理类似,在此不再赘述。

[0038] 根据所述示例性实施例,所述关键词匹配信息还包括所述多个关键词的价目。将关键词的价目数据作为选取所述关联关键词的要素之一。

[0039] 具体地,参照图3,在步骤S220,所述根据所述初选搜索词,从包括多个关键词的关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词的处理包括:对任一初选搜索词,在关键词匹配信息中查找与所述初选搜索词匹配的关键词及其价目,如果查找到的价目超过预定的价目阈值,则将匹配到的关键词匹配信息中的关键词选取为关联关键词。同理,所述预定的价目阈值可以是经验值。

[0040] 此后,在步骤S230,根据所述搜索日志和所述关联关键词,为所述初选搜索词生成包括所述关联关键词及其第二需求强度值的关键词需求列表。

[0041] 具体地,为任一所述初选搜索词,基于每个关联关键词的第二搜索行为参数计算所述关联关键词的第二需求强度值,并且将所述关联关键词及其第二需求强度值插入所述初始搜索词的关键词需求列表。例如,所述第二搜索行为参数可包括,但不限于,所述初选搜索词是否包含所述关联关键词(以1或0赋值)、所述初选搜索词的搜索结果中包含所述关联关键词的搜索结果被查看的第五频率、以及在使用所述初选搜索词搜索后使用包含所述关联关键词的其他搜索词进行搜索的第六频率。

[0042] 其中,所述初选搜索词的搜索结果中包含所述关联关键词的搜索结果被查看第五频率可以是,但不限于,例如,查看使用所述初选搜索词得到的搜索结果中标题包含所述关联关键词的用户数量在使用所述初选搜索词进行搜索的全部用户数量中所占的百分比;所述在使用所述初选搜索词搜索后使用包含所述关联关键词的其他搜索词进行搜索的第六频率可以是,但不限于,例如,在使用所述初选搜索词搜索后使用包含所述关联关键词的其他搜索词进行搜索的用户数量在使用所述初选搜索词进行搜索的全部用户数量中所占的百分比。

[0043] 根据本发明的可选实施例,在所述为任一所述初选搜索词,基于每个关联关键词的第二搜索行为参数计算所述关联关键词的第二需求强度值的处理中,对计算的所述第二搜索行为参数进行加权求和,为所述初选搜索词求得所述关联关键词的第二需求强度值。

[0044] 在为所述初选搜索词生成包括所述关联关键词及其第二需求强度值的关键词需求列表后,在步骤S240,根据所述关键词匹配信息以及第二需求强度值,对为所述初选搜索词生成的关键词需求列表中的关联关键词排序。

[0045] 具体地,步骤S240可包括:对为所述初选搜索词生成的关键词需求列表中的关联关键词,分别从所述关键词匹配信息提取其价目,对所述关联关键词的价目以及第二需求强度值进行加权求和,求得所述关联关键词的综合评估值;根据所述综合评估值,对为所述初选搜索词生成的关键词需求列表中的关联关键词进行排序。这样,通过综合评估关联关键词的第二需求强度值和价目指标,可将综合评估值高的关联关键词排在前面;在向用户推荐关联关键词时,可将综合价值高的关键词优先推荐给用户。

[0046] 在步骤S250,将所述初选搜索词及其关键词需求列表存储在关联关键词词典中。

[0047] 通过图3中步骤S210~S250的处理,可以为指定类目下的搜索词选取在类别上相关并且商业价值高的关联关键词,并且通过综合评估关联关键词的第二需求强度值和价目指标将综合评估值高的关联关键词排在前面,将综合价值高的关键词优先推荐给用户。

[0048] 图4是示出根据本发明的示例性实施例的提供关联关键词的方法的流程图。

[0049] 参照图4,在步骤S410,接收第一搜索词。在本发明中,所述搜索词包括类目关键词或从属于所述类目关键词的关键词,如“迅雷游戏”、“网页游戏”等。

[0050] 在步骤S420,根据所述第一搜索词从关联关键词词典提取至少一个关联关键词,所述关联关键词词典包括第二搜索词及至少一个关联关键词。根据本发明的优选实施例,所述至少一个关联关键词按照其需求强度排列。

[0051] 图5示出本发明实施例的生成关联关键词的方法和提供关联关键词的方法提供的搜索结果网页的示例。

[0052] 图5示出的是当用户使用“迅雷游戏”进行搜索时的搜索结果页面。在所述搜索结果页面的左侧,显示了通常的搜索结果条目;在搜索结果的页面的右侧,除了提供如“相关游戏”和“相关软件”等推荐项目以外,还提供如“网页游戏”、“小游戏”、“客户端游戏”和“三国杀网页游戏”等搜索频率高并且商业价值高的“游戏相关热搜词”。

[0053] 图6是示出根据本发明的示例性实施例的生成关联关键词的系统的逻辑框图。

[0054] 参照图6,根据本发明的示例性实施例的生成关联关键词的系统包括搜索词初选模块610、关联关键词选取模块620和词典管理装置650。

[0055] 搜索词初选模块610用于基于搜索日志选取具备预定的搜索行为特征的至少一个初选搜索词。

[0056] 具体地,搜索词初选模块610可用于对搜索日志中的任一搜索词,计算其以下第一搜索行为参数:所述搜索词是否包含预定的主类目关键词、所述搜索词是否包含预定的子类目关键词、所述搜索词的搜索结果中包含所述主类目关键词的搜索结果被查看的第一频率以及所述搜索词的搜索结果中包含所述子类目关键词的搜索结果被查看的第二频率。

[0057] 根据本发明的优选实施例,所述第一搜索行为参数还包括:在使用所述搜索词搜索后使用包含所述主类目关键词的其他搜索词进行搜索的第三频率,以及在使用所述搜索词搜索后使用包含所述子类目关键词的其他搜索词进行搜索的第四频率。

[0058] 根据本发明的另一优选实施例,所述第一搜索行为参数还包括搜索词排除参数,即所述任一搜索词是否包含预定的排除关键词。

[0059] 在此基础上,搜索词初选模块610还用于对所述任一搜索词的所述第一搜索行为参数进行加权求和,计算所述搜索词的第一需求强度值,如果所述搜索词的第一需求强度值超过第一预定需求强度阈值,则将所述搜索词选取为初选搜索词。

[0060] 关联关键词选取模块620用于根据搜索词初选模块610选取的初选搜索词,从包括多个关键词的关键词匹配信息选取至少一个与所述初选搜索词在类别上相关的关联关键词。

[0061] 例如,关联关键词选取模块620可将在所述关键词匹配信息当中包含在所述初选搜索词中的关键词选取为关联关键词,并且/或者将与所述初选搜索词存在类别上下级从属关系的关键词选取为关联关键词。

[0062] 词典管理装置650用于将搜索词初选模块610选取的初选搜索词以及关联关键词选取模块620选取的关联关键词存储在关联关键词词典中。

[0063] 根据本发明的优选实施例,所述关键词匹配信息还包括所述多个关键词的价目。关联关键词选取模块620用于对任一初选搜索词,在关键词匹配信息中查找与所述初选搜索词匹配的关键词及其价目,如果查找到的价目超过预定的价目阈值,则将匹配到的关键词匹配信息中的关键词选取为关联关键词。

[0064] 根据本发明的优选实施例,所述系统还包括关键词需求列表生成模块630和关键词需求列表排序模块640。

[0065] 关键词需求列表生成模块640用于根据所述搜索日志和所述关联关键词,为所述初选搜索词生成包括所述关联关键词及其第二需求强度值的关键词需求列表。

[0066] 具体地,关键词需求列表生成模块640用于为任一所述初选搜索词,基于每个关联关键词的以下第二搜索行为参数计算所述关联关键词的第二需求强度值,并且将所述关联关键词及其第二需求强度值插入所述初始搜索词的关键词需求列表:所述初选搜索词是否包含所述关联关键词、所述初选搜索词的搜索结果中包含所述关联关键词的搜索结果被查看的第五频率、以及在使用所述初选搜索词搜索后使用包含所述关联关键词的其他搜索词进行搜索的第六频率。此后,关键词需求列表生成模块630在所述为任一所述初选搜索词,基于每个关联关键词的以下第二搜索行为参数计算所述关联关键词的第二需求强度值的处理中,对计算的所述第二搜索行为参数进行加权求和,为所述初选搜索词求得所述关联关键词的第二需求强度值。

[0067] 关键词需求列表排序模块640用于根据所述关键词匹配信息以及第二需求强度值,对为所述初选搜索词生成的关键词需求列表中的关联关键词排序。

[0068] 具体地,关键词需求列表排序模块640用于对为所述初选搜索词生成的关键词需求列表中的关联关键词,分别从所述关键词匹配信息提取其价目,对所述关联关键词的价目以及第二需求强度值进行加权求和,求得所述关联关键词的综合评估值,并且根据所述综合评估值,对为所述初选搜索词生成的关键词需求列表中的关联关键词进行排序。

[0069] 根据本发明的优选实施例,词典管理装置650用于将所述初选搜索词及其关键词需求列表存储在关联关键词词典中。

[0070] 图7是示出根据本发明的示例性实施例的提供关联关键词的系统的逻辑框图。

[0071] 参照图7,根据本发明的示例性实施例的提供关联关键词的系统包括搜索请求接收装置710、搜索请求处理装置720和推荐关键词发送装置730。

[0072] 搜索请求接收装置710用于接收第一搜索词。

[0073] 搜索请求处理装置720用于根据搜索请求接收装置710接收的第一搜索词从关联关键词词典提取至少一个关联关键词,所述关联关键词词典包括第二搜索词及至少一个关

联关键词。

[0074] 推荐关键词发送装置730用于发送搜索请求处理装置720提取的关联关键词。

[0075] 优选地,所述至少一个关联关键词按照其需求强度排列。

[0076] 综上所述,根据本发明的示例性实施例的生成关联关键词的方法、提供关联关键词的方法和使用所述方法的系统可以为指定类目下的搜索词选取在类别上相关的关联关键词,从而当接收到包括所述搜索词的搜索请求时,可将所述关联关键词提供给用户,为用户提供上下级类别的推荐条目。

[0077] 在此基础上,还可以为指定类目下的搜索词选取在类别上相关并且商业价值高的关联关键词,并且通过综合评估关联关键词的第二需求强度值和价目指标将综合评估值高的关联关键词排在前面,将综合价值高的关键词优先推荐给用户。

[0078] 需要指出,根据实施的需要,可将本申请中描述的各个步骤拆分为更多步骤,也可将两个或多个步骤或者步骤的部分操作组合成新的步骤,以实现本发明的目的。

[0079] 上述根据本发明的方法可在硬件、固件中实现,或者被实现为可存储在记录介质(诸如CD ROM、RAM、软盘、硬盘或磁光盘)中的软件或计算机代码,或者被实现通过网络下载的原始存储在远程记录介质或非暂时机器可读介质中并将被存储在本地记录介质中的计算机代码,从而在此描述的方法可被存储在使用通用计算机、专用处理器或者可编程或专用硬件(诸如ASIC或FPGA)的记录介质上的这样的软件处理。可以理解,计算机、处理器、微处理器控制器或可编程硬件包括可存储或接收软件或计算机代码的存储组件(例如,RAM、ROM、闪存等),当所述软件或计算机代码被计算机、处理器或硬件访问且执行时,实现在此描述的处理方法。此外,当通用计算机访问用于实现在此示出的处理的代码时,代码的执行将通用计算机转换为用于执行在此示出的处理的专用计算机。

[0080] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。



图1

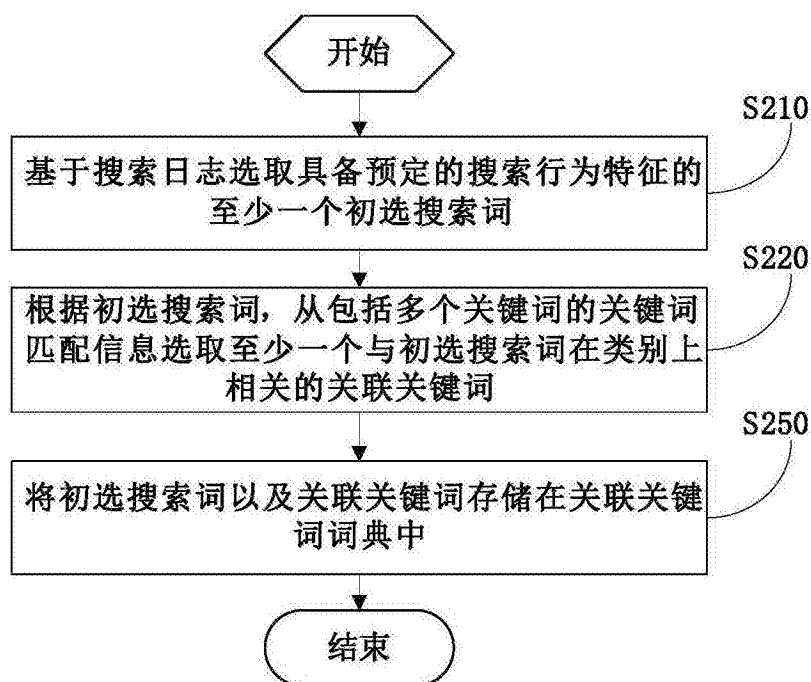


图2

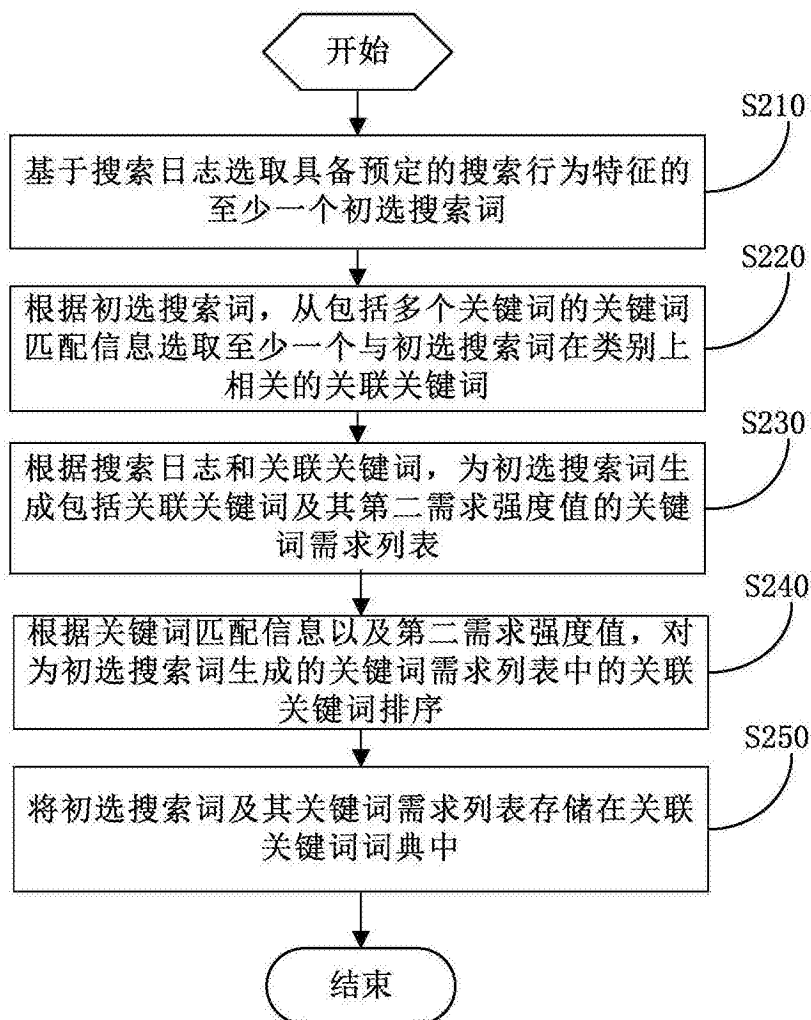


图3

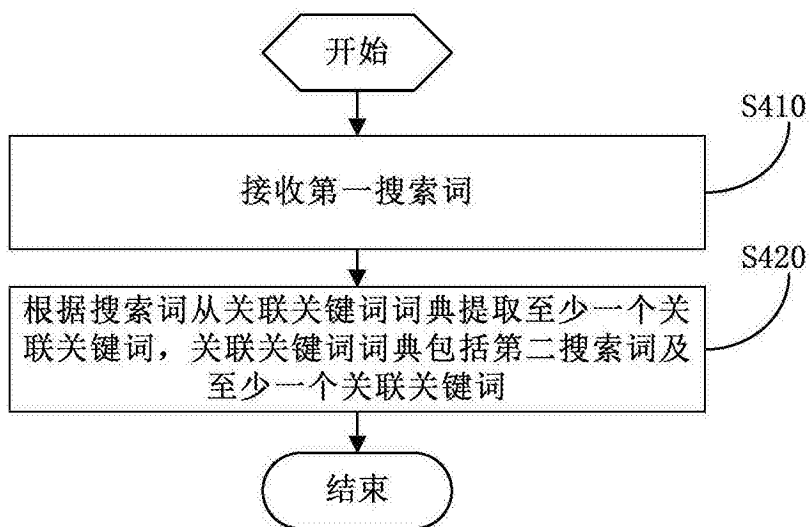


图4



图5

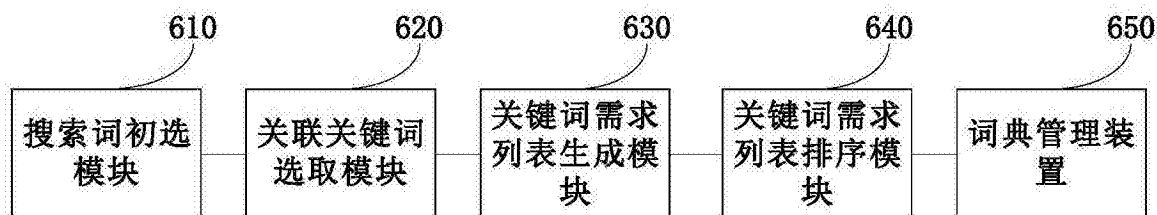


图6



图7