



(12)发明专利

(10)授权公告号 CN 109767253 B

(45)授权公告日 2020.09.04

(21)申请号 201811416514.X

H04M 3/42(2006.01)

(22)申请日 2018.11.26

审查员 曹姝妹

(65)同一申请的已公布的文献号

申请公布号 CN 109767253 A

(43)申请公布日 2019.05.17

(73)专利权人 北京盘古影艺文化传播有限公司

地址 100000 北京市丰台区大成里秀园13
号楼东侧楼三层(卢沟桥企业集中办
公区)

(72)发明人 翟润天

(74)专利代理机构 北京冠和权律师事务所

11399

代理人 朱健 张国香

(51)Int.Cl.

G06Q 30/02(2012.01)

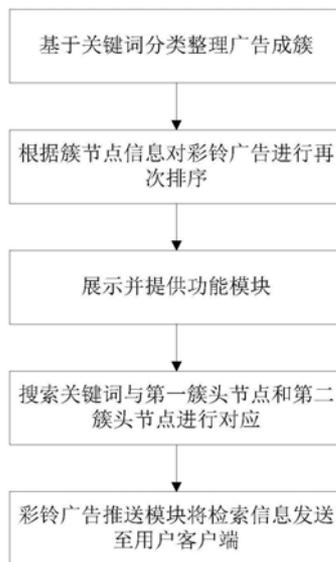
权利要求书7页 说明书15页 附图1页

(54)发明名称

一种彩铃广告推送方法及其推送系统

(57)摘要

本发明属于广告信息推送技术领域,具体涉及一种精确性更高、推送更准确的彩铃广告推送方法及其推送系统。彩铃广告推送系统包括彩铃广告收集模块、彩铃广告排序模块、彩铃广告WEB展示与服务模块、彩铃广告在线和离线检索模块、彩铃广告推送模块。该发明的框架创造性的通过关键词成簇设计将彩铃广告中的关键词与检索词进行对应,提高了彩铃广告推送的精准度。通过语义检索的方式,本发明提供了一种能够提供检索词和关键词相对应的方法,本发明能够进一步准确的将用户想要的结果检索出来并推送至用户面前。



1. 一种彩铃广告推送方法,彩铃广告推送系统包括彩铃广告收集模块、彩铃广告排序模块、彩铃广告WEB展示与服务模块、彩铃广告在线和离线检索模块、彩铃广告推送模块和服务器,所述服务器用于推送彩铃广告,其特征在于,包括如下步骤:

(1)所述彩铃广告收集模块基于第一簇头节点和第二簇头节点将彩铃广告整理成簇发送至所述彩铃广告排序模块;

(2)所述彩铃广告排序模块根据所述第一簇头节点和所述第二簇头节点将接收的所述彩铃广告再次排序并发送至所述彩铃广告WEB展示与服务模块;

(3)所述彩铃广告WEB展示与服务模块用于展示所述彩铃广告;

(4)所述彩铃广告在线和离线检索模块搜索关键词与第一簇头节点和第二簇头节点进行匹配,并将匹配成功的彩铃广告发送至所述彩铃广告推送模块;

(5)所述彩铃广告推送模块将所述匹配成功的彩铃广告发送至用户客户端供用户使用;

其中,所述彩铃广告收集模块基于第一簇头节点和第二簇头节点将彩铃广告整理成簇发送至所述彩铃广告排序模块,包括:

(1.1)选取初始特征簇头;

所述彩铃广告收集模块与所述服务器连接,且能选取所述服务器中n个彩铃广告作为n个节点,每个彩铃广告的第一个关键词作为第一簇头节点,提取所述服务器中n个彩铃广告中的每个广告的第二关键词作为第二簇头节点;

(1.2)成簇并建立特征簇头地图;

将第一和第二簇头节点按照频率由高到低的顺序设置为 $i(x_i, y_i)$ 、 $j(x_j, y_j)$, x_i 为第一簇头节点频率值, y_i 为第一簇头节点顺序号, x_j 为第二簇头节点频率值, y_j 为第二簇头节点顺序号,按照坐标形式建立成特征簇头地图; $i = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, n$;

(1.3)重新选取簇头;

计算第一簇头节点到第二簇头节点之间的距离:

$$D_{ij}(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2};$$

每个节点根据自己的第二关键词被检索的次数进行加权和 W_s 计算,选择加权和最小的节点作为第二簇头节点,

$$W_s = \delta D_s + \lambda \frac{1}{B};$$

$$D_s = \frac{\sum_{j=1}^n D_{ij}(i, j)}{\sum_{i=1}^n \sum_{j=1}^n D_{ij}(i, j)};$$

$$B = \frac{b_i}{\sum_{i=1}^n b_i + \sum_{j=1}^n b_j};$$

δ 和 λ 为加权系数, $\delta + \lambda = 1$, D_s 为第一簇头节点 (x_i, y_i) 与所有第二簇头节点距离

总和占有所有簇头节点间距离综合的比重; B 表示第一簇头节点 (x_i, y_i) 被检索次数 b_i 占有所有簇头节点被检索次数的比重;

若新的第二簇头节点与步骤(1.1)中的簇头节点不一致,即簇头节点发生改变,则重新执行步骤(1.1);若簇头节点没有改变,则继续执行步骤(1.4);

(1.4)新一轮簇头节点选取成功后,具有相同第二簇头节点的彩铃广告为一簇,所述彩铃广告收集模块按照簇的序号发送至所述彩铃广告WEB展示与服务模块;

所述彩铃广告排序模块根据所述第一簇头节点和所述第二簇头节点将接收的所述彩铃广告再次排序并发送至所述彩铃广告WEB展示与服务模块,包括:

(2.1)分别计算第一簇头节点与第二簇头节点在同一彩铃广告内出现的次数,设第一簇头节点在彩铃广告内出现 p 次,第二簇头节点出现了 q 次,判断 p 与 q 两个值的大小,取两者比较的小值作为在同一彩铃广告内的共现次数,同时将值赋给 r_i ,第一簇头节点与第二簇头节点在所述服务器内所有彩铃广告中出现次数 r 为:

$$r = r_1 + r_2 + r_3 + \dots + r_n = \sum_{i=1}^n r_i;$$

n 为所述服务器内所有彩铃广告的总数,当第一簇头节点与第二簇头节点未在同一彩铃广告内出现时, $r_i = 0$;

(2.2)计算第一簇头节点到第二簇头节点之间的预计距离:

$$D = \frac{D_{\max} + D_{\min}}{2};$$

D_{\max} 为 D 的最大值, D_{\min} 为 D 的最小值,

第一簇头节点与第二簇头节点在彩铃广告的距离取所有彩铃广告内距离的平均值;

$$\bar{D} = \frac{\bar{D} * r_s + D}{r_s + 1};$$

r_s 为第一簇头节点与第二簇头节点共现在彩铃广告的次数;

第二簇头节点在彩铃广告中的信息增益为:

$$\alpha_j = -\sum_{j=1}^n \frac{\beta_{jt}}{\sum_{j=1}^n \beta_{jt}} \ln \frac{\beta_{jt}}{\sum_{j=1}^n \beta_{jt}} ; ;$$

α_j 表示第二簇头节点在彩铃广告中的信息增益;

β_{jt} 为第二簇头节点在彩铃广告 t 中出现的次数;

(2.3)计算第一簇头节点与第二簇头节点的相关度为:

$$\beta_{ij} = \frac{r \frac{a_j}{a_j + \bar{D}}}{a_j};$$

(2.4) 计算第一簇头节点在彩铃广告中的权重;

$$\gamma = \beta_{ij} \times \log \frac{N + 0.05}{r_s};$$

(2.5) 将第一簇头节点按照与第二簇头节点的权重值相同的顺序进行排序,根据第一簇头节点的排序对彩铃广告进行排序。

2. 根据权利要求1所述的一种彩铃广告推送方法,其特征在于,所述彩铃广告在线和离线检索模块根据关键词与第一簇头节点和第二簇头节点进行匹配,具体包括:

(3.1) 根据所述彩铃广告在线和离线检索模块的关键词提取同义词库:

$$W = \{w_0, w_1, w_2, \dots, w_v\};$$

w_1 为第一个关键词的同义词; w_2 为第二个关键词的同义词; w_v 为关键词的第v个同义词; w_0 为有歧义的关键词, w_0 的歧义包括:

$$\text{mean}(w_0) = \{w_{01}, w_{02}, w_{03}, \dots, w_{0m}\}; ;$$

m为歧义词的数量;

(3.2) 估计W和彩铃广告的关键词的语义 s_j 之间的相关度,

$$C_s = \arg \max_{1 \leq j \leq m} \text{Rele}(W, G(s_j)); j \in \{1, \dots, q\}$$

C_s 表示同义词库与关键词的语义之间的相关度;

$\text{Rele}(W, G(s_j))$ 表示同义词库与语义 s_j 的相关词集合 $G(s_j)$ 的相关度, $G(s_j)$ 表示通过推理机得到的语义 s_j 的相关词集合;

$$\text{Rele}(W, G(s_j)) = \sum_{j=1}^v \frac{\text{Wei}(w_j, s_j)}{D(w_j, w_0)};$$

$\text{Wei}(w_j, s_j)$ 表示 w_j 与 s_j 之间的相关度, $D(w_j, w_0)$ 表示 w_j 与 w_0 之间的距离;

(3.3) 通过相关度阈值 Δ 确定同义词库与关键词之间的匹配关系是否成立,如果 $C_s \geq \Delta$,则同义词库与关键词之间相关;如果 $C_s < \Delta$,则依次删除 $\text{mean}(w_0)$ 中的歧义

词,重新执行步骤(3.1),直到 $C_s \geq \Delta$;

(3.4)将匹配结果发送至所述彩铃广告推送模块。

3.根据权利要求1所述的一种彩铃广告推送方法,其特征在于,所述彩铃广告推送模块将所述匹配成功的彩铃广告发送至用户客户端供用户使用包括:

所述彩铃广告推送模块获取匹配结果信息,根据所述匹配结果信息获取彩铃广告并发送给用户;若所述用户进行了下载或确认操作则确定对应结果,将检测结果储存于所述服务器或者云端以备下次使用。

4.根据权利要求3所述的一种彩铃广告推送方法,其特征在于,所述推送信息包括彩铃广告名称、音频、出版信息和格式。

5.一种彩铃广告推送系统,包括彩铃广告收集模块、彩铃广告排序模块、彩铃广告WEB展示与服务模块、彩铃广告在线和离线检索模块、彩铃广告推送模块和服务器,所述服务器用于推送彩铃广告,其特征在于:所述彩铃广告收集模块基于第一簇头节点和第二簇头节点将彩铃广告分类整理成簇发送至所述彩铃广告排序模块;所述彩铃广告收集模块与所述服务器连接,选取所述服务器中n个彩铃广告作为n个节点,每个彩铃广告的第一个关键词作为第一簇头节点,提取所述服务器中n个彩铃广告中每个广告的第二关键词作为第二簇头节点;将第一和第二簇头节点按照频率由高到低的顺序设置为 $i(x_i, y_i)$ 、 $j(x_j, y_j)$, x_i 为第一簇头节点频率值, y_i 为第一簇头节点顺序号, x_j 为第二簇头节点频率值, y_j 为第二簇头节点顺序号,按照坐标形式建立成特征簇头地图; $i=1, 2, 3, \dots, n$; $j=1, 2, 3, \dots, n$;计算第一簇头节点到第二簇头节点之间的距离:

$$D_{ij}(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2};$$

每个节点根据自己的第二关键词被检索的次数进行加权和 W_s 计算,选择加权和最小的节点作为第二簇头节点,

$$W_s = \delta D_s + \lambda \frac{1}{B};$$

$$D_s = \frac{\sum_{j=1}^n D_{ij}(i, j)}{\sum_{i=1}^n \sum_{j=1}^n D_{ij}(i, j)};$$

$$B = \frac{b_i}{\sum_{i=1}^n b_i + \sum_{j=1}^n b_j};$$

δ 和 λ 为加权系数, $\delta + \lambda = 1$, D_s 为第一簇头节点 (x_i, y_i) 与所有第二簇头节点距离总和占所有簇头节点间距离综合的比重; B 表示第一簇头节点 (x_i, y_i) 被检索次数 b_i 占所述所有簇头节点被检索次数的比重;

若新的第二簇头节点与簇头节点不一致,即簇头节点发生改变,则重新选取初始特征簇头;若簇头节点没有改变,则新一轮簇头节点选取成功后,具有相同第一簇头节点的彩铃广告为一簇,所述彩铃广告收集模块按照簇的序号发送至所述彩铃广告WEB展示与服务模块;

所述彩铃广告排序模块根据所述第一簇头节点和所述第二簇头节点信息将接收的所述彩铃广告再次排序并发送至所述彩铃广告WEB展示与服务模块;分别计算第一簇头节点与第二簇头节点在同一彩铃广告内出现的次数,设第一簇头节点在彩铃广告内出现 p 次,第二簇头节点出现了 q 次,判断 p 与 q 两个值的大小,取两者比较的小值作为在同一彩铃广告内的共现次数,同时将值赋给 r_i ,第一簇头节点与第二簇头节点在所述服务器内所有彩铃广告中出现次数 r 为:

$$r = r_1 + r_2 + r_3 + \dots + r_n = \sum_{i=1}^n r_i ;$$

n 为服务器所有彩铃广告的总数,当第一簇头节点与第二簇头节点未在同一彩铃广告内出现时, $r_i = 0$;

计算第一簇头节点到第二簇头节点之间的预计距离:

$$D = \frac{D_{\max} + D_{\min}}{2} ;$$

D_{\max} 为 D 的最大值, D_{\min} 为 D 的最小值,

第一簇头节点与第二簇头节点在彩铃广告的距离取所有彩铃广告内距离的平均值;

$$\bar{D} = \frac{\bar{D} * r_s + D}{r_s + 1} ;$$

r_s 为第一簇头节点与第二簇头节点共现在彩铃广告的次数;

第二簇头节点在彩铃广告中的信息增益为:

$$\alpha_j = -\sum_{j=1}^n \frac{\beta_{jt}}{\sum_{j=1}^n \beta_{jt}} \ln \frac{\beta_{jt}}{\sum_{j=1}^n \beta_{jt}} ;$$

α_j 表示第二簇头节点在彩铃广告中的信息增益;

β_{jt} 为第二簇头节点在彩铃广告 t 中出现的次数;

计算第一簇头节点与第二簇头节点的相关度为:

$$\beta_{ij} = \frac{r \frac{a_j}{a_j + \bar{D}}}{a_j};$$

计算第一簇头节点在彩铃广告中的权重;

$$\gamma = \beta_{ij} \times \log \frac{N + 0.05}{r_s};$$

将第一簇头节点按照与第二簇头节点的权重值的预设顺序进行排序,根据第一簇头节点的排序对彩铃广告进行排序;

所述彩铃广告在线和离线检索模块根据关键词与第一簇头节点和第二簇头节点进行匹配,并将匹配结果发送至所述彩铃广告推送模块;根据所述彩铃广告在线和离线检索模块获取关键词:

$$W = \{w_0, w_1, w_2, \dots, w_v\};$$

w_1 为第一个关键词的同义词; w_2 为第二个关键词的同义词; w_v 为关键词的第v个同义词; w_0 为有歧义的关键词, w_0 的歧义包括:

$$\text{mean}(w_0) = \{w_{01}, w_{02}, w_{03}, \dots, w_{0m}\};$$

m为歧义词的数量;

估计W和彩铃广告的关键词的语义 s_j 之间的相关度,

$$C_s = \arg \max_{1 \leq j \leq m} \text{Rele}(W, G(s_j)); j \in \{1, \dots, q\}$$

C_s 表示同义词库与关键词的语义之间的相关度;

$\text{Rele}(W, G(s_j))$ 表示同义词库与语义 s_j 的相关词集合 $G(s_j)$ 的相关度, $G(s_j)$ 表示通过推理机得到的语义 s_j 的相关词集合;

$$\text{Rele}(W, G(s_j)) = \sum_{j=1}^v \frac{\text{Wei}(w_j, s_j)}{D(w_j, w_0)};$$

$\text{Wei}(w_j, s_j)$ 表示 w_j 与 s_j 之间的相关度, $D(w_j, w_0)$ 表示 w_j 与 w_0 之间的距离;

通过相关度阈值 Δ 确定同义词库与关键词之间的匹配关系是否成立,如果 $C_s \geq \Delta$,则同义词库与关键词之间相关;如果 $C_s < \Delta$,则依次删除 $\text{mean}(w_0)$ 中的歧义词,直到

$C_S \geq \Delta$;

将匹配结果发送至所述彩铃广告推送模块;

所述彩铃广告推送模块将匹配成功的彩铃广告发送至用户客户端供用户使用;所述彩铃广告推送模块获取匹配结果信息,根据匹配结果信息获取彩铃广告的推送信息并发送给用户;若用户进行了下载或确认操作则确定对应结果,将检测结果储存于所述服务器或者云端以备下次使用;所述推送信息包括:彩铃广告名称、音频、出版信息、格式。

一种彩铃广告推送方法及其推送系统

技术领域

[0001] 本发明属于广告信息推送技术领域,具体涉及一种精确性更高、推送更准确的彩铃广告推送方法及其推送系统。

背景技术

[0002] 韩国SK电信与2002年3月率先在韩国推出彩铃业务,短时间内就风靡全韩国。2003年,彩铃业务被中国移动率先引入中国市场,并取得了巨大的商业成功,而后中国联通首先在它高端用户占主流的CDMA网中开展了炫铃业务,商用后进一步扩大了炫铃的用户容量,使联通的专用号段的网络用户也可以与CDMA用户一样享受炫铃的优质服务。2005年中国移动增值业务彩铃的收益远远超过了全国传统唱片市场的收入,彩铃从此成为移动运营商的重要收入来源,被认为是在继短信业务以后,移动增值业务中最可能再现行业“神话”的业务。作为一项全新的语音业务,彩铃业务不像短信、呼叫转移等业务那样对话音业务有替代作用,对运营商而言,它所带来的完全是全新的收益,所以一开始就收到移动运营商的青睐;对大户用户而言,它所提供是一种全新视听效果与优质的服务,所以一开始就得到广大用户的认可。

[0003] 个性化回铃音业务也称为彩铃业务即,客户可以根据自己的喜好和要求定制自己电话的回铃音。回铃音既可以是音乐、广告,也可以是客户录制的语音。彩铃业务适合于所有的移动用户,包括移动、联通及电信用户。彩铃业务是一项由被叫或主叫用户定制,为主叫用户提供一段悦耳的音乐或一句问候语来替代普通回铃音的业务。用户申请开通彩铃业务之后,可以自行设定个性化回铃音,在其作为被叫时,为主叫用户播放个性化定制的音乐或录音,来代替普通的回铃音。当被叫用户处于忙、不在服务区、关机等非空闲状态情况下时,仍播放原网络系统提供的语音通知。

[0004] 移动增值业务,指凭借移动网络资源及相关技术设备而开发的附加通信业务,是对移动语音业务的延伸与发展,其实现的价值使原有网络的经济价值大大增加,目前通常所讲的移动增值业务是指除传统基础移动语音通信之外的各项其它业务。近年来,随着通信、计算机、微电子、软件等技术的进步,移动通信领域的发展日新月异。第四代移动通信系统在全球的推出使移动通信网络的承载力得到了极大的提升,各种新的移动增值业务应运而生、层出不穷,移动终端的样式与种类也极大丰富,手机成为第五媒体的说法已被广泛接受。随着移动增值业务的极大丰富,未来移动终端的功能将会更多、更强,与人们工作、生活、娱乐、安全保障等日常生活息息相关的功能会最大程度汇集到移动终端上。与此同时,行业间的融合也在移动终端上得以体现。从近几年移动增值业务发展来看,我们可以发现,新技术和新业务层出不穷,从最开始的短信业务到后来的语音杂志、彩信,再到高速发展且很受大众喜欢的彩铃以及手机上网业务等。每一项新的增值业务都不同程度的体现了移动增值业务发展速度快、渗透力强、满足个性需求的特点。移动通信的日益普及使用户得以随时、随地、随心地进行沟通交流。而通信技术的飞速发展,使移动用户在满足通话需求的基础上。在信息沟通、娱乐方面的需求也提出了更高的要求。

[0005] 其中彩铃广告就是极具发展潜力的移动增值业务之一,通过极具特点的彩铃对手机或者WEB终端的用户进行点对点推送,进而达到精准推销的移动业务,极大节约了广告成本并有效提高了推销效率。

发明内容

[0006] 本发明的目的在于提供一种精确推送内容、提高推销效率的彩铃广告推送系统。

[0007] 本发明的目的还在于提供一种彩铃广告推送方法。

[0008] 本发明的目的是这样实现的:

[0009] 一种彩铃广告推送方法,彩铃广告推送系统包括彩铃广告收集模块、彩铃广告排序模块、彩铃广告WEB展示与服务模块、彩铃广告在线和离线检索模块、彩铃广告推送模块和服务服务器,所述服务器用于推送彩铃广告,包括如下步骤:

[0010] (1)所述彩铃广告收集模块基于关键词将彩铃广告分类整理成簇发送至所述彩铃广告排序模块;

[0011] (2)所述彩铃广告排序模块根据簇节点信息对接收的所述彩铃广告进行再次排序并发送至所述彩铃广告WEB展示与服务模块;

[0012] (3)所述彩铃广告WEB展示与服务模块用于展示所述彩铃广告;

[0013] (4)所述彩铃广告在线和离线检索模块根据用户的搜索关键词与第一簇头节点和第二簇头节点进行匹配,并将匹配成功的彩铃广告发送至所述彩铃广告推送模块;

[0014] (5)所述彩铃广告推送模块将所述匹配成功的彩铃广告发送至用户客户端供用户使用。

[0015] 在一个实施例中,所述彩铃广告收集模块基于关键词将彩铃广告分类整理成簇发送至所述彩铃广告排序模块具体步骤包括:

[0016] (1.1)选取初始特征簇头;

[0017] 所述彩铃广告收集模块与所述服务器连接,且能选取所述服务器中n个彩铃广告作为n个节点,每个彩铃广告的第一个关键词作为第一簇头节点,提取所述服务器中n个彩铃广告中第二个关键词作为第二簇头节点;

[0018] (1.2)成簇并建立特征簇头地图;

[0019] 将第一和第二簇头节点按照频率由高到低的顺序设置为 $i(x_i, y_i)$ 、 $j(x_j, y_j)$, x_i 为第一簇头节点频率值(关键词出现的次数除以关键词出现的总次数), y_i 为第一簇头节点顺序号,为第二簇头节点频率值, y_j 为第二簇头节点顺序号,按照坐标形式建立成特征簇头地图; $i=1, 2, 3 \dots n$; $j=1, 2, 3 \dots n$;

[0020] (1.3)重新选取簇头;

[0021] 计算第一簇头到第二簇头之间的距离:

$$[0022] D_{ij}(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2};$$

[0023] 每个节点根据自己的第二关键词被检索的次数进行加权和 W_s 计算,选择加权和最小的节点作为第二簇头节点,

$$[0024] \quad W_s = \delta D_s + \lambda \frac{1}{B};$$

$$[0025] \quad D_s = \frac{\sum_{j=1}^n D_{ij}(i,j)}{\sum_{i=1}^n \sum_{j=1}^n D_{ij}(i,j)};$$

$$[0026] \quad B = \frac{b_i}{\sum_{i=1}^n b_i + \sum_{j=1}^n b_j};$$

[0027] δ 和 λ 为加权系数, $\delta + \lambda = 1$, D_s 为第一簇头节点 (x_i, y_i) 与所有第二簇头节点距离总和占所有簇头节点间距离综合的比重; B 表示第一簇头节点 (x_i, y_i) 被检索次数 b_i 占相关簇头节点被检索次数的比重;

[0028] 若新的第二簇头节点与步骤(1.1)中的簇头簇头节点不一致, 即簇头节点发生改变, 则重新执行步骤(1.1); 若簇头节点没有改变, 则继续执行步骤(1.4);

[0029] (1.4) 新一轮簇头节点选取成功后, 具有相同第二簇头节点的彩铃广告为一簇, 所述彩铃广告收集模块按照簇的序号发送至所述彩铃广告WEB展示与服务模块。

[0030] 在一个实施例中, 所述彩铃广告排序模块根据簇节点信息对所述彩铃广告进行再次排序, 具体包括:

[0031] (2.1) 分别计算第一簇头节点与第二簇头节点在同一彩铃广告内出现的次数, 设第一簇头节点在彩铃广告内出现 p 次, 第二簇头节点出现了 q 次, 判断 p 与 q 两个值的大小, 取两者比较的小值作为在同一彩铃广告内的共现次数, 同时将值赋给 r_i , 第一簇头节点与第二簇头节点在所述服务器内所有彩铃广告中出现次数 r 为:

$$[0032] \quad r = r_1 + r_2 + r_3 + \dots + r_N = \sum_{i=1}^N r_i;$$

[0033] N 为所述服务器内所有彩铃广告的总数, 当第一簇头节点与第二簇头节点未在同一彩铃广告内出现时, $r_i = 0$;

[0034] (2.2) 计算第一簇头节点到第二簇头节点之间的预计距离:

$$[0035] \quad D = \frac{D_{\max} + D_{\min}}{2};$$

[0036] D_{\max} 为 D 的最大值, D_{\min} 为 D 的最小值,

[0037] 第一簇头节点与第二簇头节点在彩铃广告的距离取所有彩铃广告内距离的平均值;

$$[0038] \quad \bar{D} = \frac{\bar{D} * r_s + D}{r_s + 1};$$

[0039] r_s 为第一簇头节点与第二簇头节点共现在彩铃广告的次数;

[0040] 第二簇头节点在彩铃广告中的信息增益为:

$$[0041] \quad \alpha_j = - \sum_{j=1}^N \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}} \ln \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}};$$

[0042] β_{jt} 为第二簇头节点在文档 t 中出现的次数;

[0043] (2.3) 计算第一簇头节点与第二簇头节点的相关度为:

$$[0044] \quad \beta_{ij} = \frac{r \frac{\alpha_j}{\alpha_j + \bar{D}}}{\alpha_j};$$

[0045] (2.4) 计算第一簇头节点在彩铃广告中的权重;

$$[0046] \quad \gamma = \beta_{ij} \times \log \frac{N + 0.05}{r_s};$$

[0047] (2.5) 将第一簇头节点按照与第二簇头节点的权重值进行排序, 根据第一簇头节点的排序对彩铃广告进行排序。

[0048] 在一个实施例中, 所述彩铃广告在线和离线检索模块根据用户的搜索关键词与第一簇头节点和第二簇头节点进行匹配, 具体包括:

[0049] (3.1) 根据所述彩铃广告在线和离线检索模块的获得检索词提取同义词库:

$$[0050] \quad W = \{W_0, W_1, W_2, \dots, W_v\};$$

[0051] w_v 为检索词的同义词; w_0 为有歧义的检索词, w_0 的歧义包括:

$$[0052] \quad \text{mean}(w_0) = \{w_{01}, w_{02}, w_{03}, \dots, w_{0m}\};$$

[0053] m 为歧义词的数量;

[0054] (3.2) 估计 W 和语义 s_q 之间的相关度,

$$[0055] \quad C_s = \arg \max_{1 \leq j \leq m} \text{Rele}(W, G(s_j)); j \in \{1, \dots, q\}$$

[0056] $\text{Rele}(W, G(s_j))$ 表示同义词库与语义 s_j 的相关词集合 $G(s_j)$ 的相关度, $G(s_j)$

表示通过推理机得到的语义 S_j 的相关词集合；

$$[0057] \quad Rele(W, G(s_j)) = \sum_{j=1}^v \frac{Wei(w_j, s_j)}{D(w_j, w_0)}$$

[0058] Wei表示 w_j 与 S_j 之间的相关度, $D(w_j, w_0)$ 表示 w_j 与 w_0 之间的距离；

[0059] (3.3)通过相关度阈值 Δ 确定同义词库与检索词之间的匹配关系是否成立,如果 $C_S \geq \Delta$,则同义词库与检索词之间相关;如果 $C_S < \Delta$,则依次删除 $mean(w_0)$ 中的歧义

词,重新执行步骤(3.1),直到 $C_S \geq \Delta$;

[0060] (3.4)将匹配结果发送至所述彩铃广告推送模块。

[0061] 在一个实施例中,所述彩铃广告推送模块将所述匹配成功的彩铃广告发送至用户客户端供用户使用包括:

[0062] 所述彩铃广告推送模块获取匹配结果信息,根据所述匹配结果信息获取彩铃广告并发送给用户;若所述用户进行了下载或确认操作则确定对应结果,将检测结果储存于所述服务器或者云端以备下次使用。

[0063] 在一个实施例中,所述推送信息包括彩铃广告名称、音频、出版信息和格式。

[0064] 一种彩铃广告推送系统,包括彩铃广告收集模块、彩铃广告排序模块、彩铃广告WEB展示与服务模块、彩铃广告在线和离线检索模块、彩铃广告推送模块和服务器,所述服务器用于推送彩铃广告,所述彩铃广告收集模块基于关键词将彩铃广告分类整理成簇发送至所述彩铃广告排序模块;所述彩铃广告收集模块与所述服务器连接,选取所述服务器中n个彩铃广告作为n个节点,每个彩铃广告的第一个关键词作为第一簇头节点,提取所述服务器中n个彩铃广告中第二个关键词作为第二簇头节点;将第一和第二簇头节点按照频率由高到低的顺序设置为 $i(x_i, y_i)$ 、 $j(x_j, y_j)$, x_i 为第一簇头节点频率值(关键词出现的次数除以关键词出现的总次数), y_i 为第一簇头节点顺序号, x_j 为第二簇头节点频率值, y_j 为第二簇头节点顺序号,按照坐标形式建立成特征簇头地图; $i=1,2,3\dots n$; $j=1,2,3\dots n$;计算第一簇头到第二簇头之间的距离:

$$[0065] \quad D_{ij}(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2};$$

[0066] 每个节点根据自己的第二关键词被检索的次数进行加权和 W_s 计算,选择加权和最小的节点作为第二簇头节点,

$$[0067] \quad W_s = \delta D_s + \lambda \frac{1}{B};$$

$$[0068] \quad D_s = \frac{\sum_{j=1}^n D_{ij}(i,j)}{\sum_{i=1}^n \sum_{j=1}^n D_{ij}(i,j)};$$

$$[0069] \quad B = \frac{b_i}{\sum_{i=1}^n b_i + \sum_{j=1}^n b_j};$$

[0070] δ 和 λ 为加权系数, $\delta + \lambda = 1$, D_s 为第一簇头节点 (x_i, y_i) 与所有第二簇头节点距离总和占所有簇头节点间距离综合的比重; 表示第一簇头节点 (x_i, y_i) 被检索次数 b_i 占相关簇头节点被检索次数的比重;

[0071] 若新的第二簇头节点与簇头节点不一致, 即簇头节点发生改变, 则重新选取初始特征簇头; 若簇头节点没有改变, 则新一轮簇头节点选取成功后, 具有相同第二簇头节点的彩铃广告为一簇, 所述彩铃广告收集模块按照簇的序号发送至所述彩铃广告WEB展示与服务模块;

[0072] 所述彩铃广告排序模块根据簇节点信息对彩铃广告进行再次排序并发送至所述彩铃广告WEB展示与服务模块; 分别计算第一簇头节点与第二簇头节点在同一彩铃广告内出现的次数, 设第一簇头节点在彩铃广告内出现 p 次, 第二簇头节点出现了 q 次, 判断 p 与 q 两个值的大小, 取两者比较的小值作为在同一彩铃广告内的共现次数, 同时将值赋给 r_i , 第一簇头节点与第二簇头节点在所述服务器内所有彩铃广告中出现次数 r 为:

$$[0073] \quad r = r_1 + r_2 + r_3 + \dots + r_N = \sum_{i=1}^N r_i;$$

[0074] N 为服务器所有彩铃广告的总数, 当第一簇头节点与第二簇头节点未在同一彩铃广告内出现时, $r_i = 0$;

[0075] 计算第一簇头节点到第二簇头节点之间的预计距离:

$$[0076] \quad D = \frac{D_{\max} + D_{\min}}{2};$$

[0077] D_{\max} 为 D 的最大值, D_{\min} 为 D 的最小值,

[0078] 第一簇头节点与第二簇头节点在彩铃广告的距离取所有彩铃广告内距离的平均值;

$$[0079] \quad \bar{D} = \frac{\bar{D} * r_s + D}{r_s + 1};$$

[0080] r_s 为第一簇头节点与第二簇头节点共现在彩铃广告的次数；

[0081] 第二簇头节点在彩铃广告中的信息增益为：

$$[0082] \quad \alpha_j = - \sum_{j=1}^N \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}} \ln \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}};$$

[0083] β_{jt} 为第二簇头节点在文档t中出现的次数；

[0084] 计算第一簇头节点与第二簇头节点的相关度为：

$$[0085] \quad \beta_{ij} = \frac{r \frac{\alpha_j}{\alpha_j + \bar{D}}}{\alpha_j};$$

[0086] 计算第一簇头节点在彩铃广告中的权重；

$$[0087] \quad \gamma = \beta_{ij} \times \log \frac{N + 0.05}{r_s};$$

[0088] 将第一簇头节点按照与第二簇头节点的权重值进行排序,根据第一簇头节点的排序对彩铃广告进行排序。

[0089] 所述彩铃广告在线和离线检索模块根据用户的搜索关键词与第一簇头节点和第二簇头节点进行匹配,并将匹配结果发送至所述彩铃广告推送模块;根据所述彩铃广告在线和离线检索模块获取检索词提取同义词库:

$$[0090] \quad W = \{W_0, W_1, W_2, \dots, W_v\};$$

[0091] w_v 为检索词的同义词; w_0 为有歧义的检索词, w_0 的歧义包括:

$$[0092] \quad \text{mean}(w_0) = \{w_{01}, w_{02}, w_{03}, \dots, w_{0m}\};$$

[0093] m为歧义词的数量;

[0094] 估计W和语义 s_q 之间的相关度,

$$[0095] \quad C_s = \arg \max_{1 \leq j \leq m} \text{Rele}(W, G(s_j)); \quad j \in \{1, \dots, q\}$$

[0096] $\text{Rele}(W, G(s_j))$ 表示同义词库与语义 s_j 的相关词集合 $G(s_j)$ 的相关度, $G(s_j)$

表示通过推理机得到的语义 s_j 的相关词集合;

$$[0097] \quad Rele(W, G(s_j)) = \sum_{j=1}^v \frac{Wei(w_j, s_j)}{D(w_j, w_0)}$$

[0098] Wei表示 w_j 与 s_j 之间的相关度, $D(w_j, w_0)$ 表示 w_j 与 w_0 之间的距离;

[0099] 通过相关度阈值 Δ 确定同义词库与检索词之间的匹配关系是否成立,如果 $C_S \geq \Delta$, 则同义词库与检索词之间相关;如果 $C_S < \Delta$, 则依次删除 $mean(w_0)$ 中的歧义词,直到 $C_S \geq \Delta$;

[0100] 将匹配结果发送至所述彩铃广告推送模块;

[0101] 所述彩铃广告推送模块将匹配成功的彩铃广告发送至用户客户端供用户使用;所述彩铃广告推送模块获取匹配结果信息,根据匹配结果信息获取彩铃广告的推送信息并发送给用户;若用户进行了下载或确认操作则确定对应结果,将检测结果储存于所述服务器或者云端以备下次使用;所述推送信息包括:彩铃广告名称、音频、出版信息、格式。

[0102] 本发明的有益效果在于:

[0103] 该发明的框架创造性的通过关键词成簇设计将彩铃广告中的关键词与检索词进行对应,提高了彩铃广告推送的精准度。通过语义检索的方式,本发明提供了一种能够提供检索词和关键词相对应的方法,本发明能够进一步准确的将用户想要的结果检索出来并推送至用户面前。

附图说明

[0104] 图1为一种彩铃广告推送方法的流程图。

具体实施方式

[0105] 下面结合附图对本发明做进一步描述。

[0106] 信息技术的快速发展和互联网业务对于人民生活的不断渗透,深化了互联网广告产业链的日益成熟,业界对于广告投放技术的竞争也日趋激烈。彩铃广告,尤其是电子商务网站中的在线广告,是互联网生活中的潜在的广告展现形式之一。随着媒体和电商网站积累了越来越多的用户行为,浏览网页时各种彩铃的点击下载,有效的进行用户类别的挖掘和进行精准的投放,成为了广告增加收入的潜在手段。点击率预估技术是预估用户在特定场景下对特定广告点击概率的技术,是计算广告中进行精准投放的主要手段之一,但是面对彩铃业务网站海量的广告和流式的用户日志,广告排序和点击率模型的更新方式成了当前在线彩铃广告的技术难点。同时,随着日志数据的爆发式增长,电子商务网站对海量数据存储和并行计算框架的需求也日益迫切,本发明正是针对上述需求,设计并实现一个面向电子商务网站的在线彩铃广告推送系统,核心是在海量数据的基础上实现较为通用的电商网站精准彩铃广告框架。在线彩铃广告系统在帮助电商网站降低处理海量日志的技术门槛同时,提升有限流量场景下的彩铃广告投放转化效率。

[0107] 具体来说本发明的一种彩铃广告推送方法,彩铃广告推送系统包括彩铃广告收集

模块、彩铃广告排序模块、彩铃广告WEB展示与服务模块、彩铃广告在线和离线检索模块、彩铃广告推送模块和服务器,服务器用于推送彩铃广告,包括如下步骤:

[0108] (1)彩铃广告收集模块基于关键词将彩铃广告分类整理成簇发送至彩铃广告排序模块;

[0109] (2)彩铃广告排序模块根据簇节点信息对接收的彩铃广告进行再次排序并发送至彩铃广告WEB展示与服务模块;

[0110] (3)彩铃广告WEB展示与服务模块用于展示所述彩铃广告;

[0111] (4)彩铃广告在线和离线检索模块根据用户的搜索关键词与第一簇头节点和第二簇头节点进行匹配,并将匹配成功的彩铃广告发送至彩铃广告推送模块;

[0112] (5)彩铃广告推送模块将匹配成功的彩铃广告发送至用户客户端供用户使用。

[0113] 该发明的框架创造性的通过关键词成簇设计将彩铃广告中的关键词与检索词进行对应,提高了彩铃广告推送的精准度。

[0114] 进一步的,彩铃广告收集模块基于关键词将彩铃广告分类整理成簇发送至彩铃广告排序模块具体步骤包括:

[0115] (1.1)选取初始特征簇头;

[0116] 彩铃广告收集模块与服务器连接,且能选取服务器中n个彩铃广告作为n个节点,每个彩铃广告的第一个关键词作为第一簇头节点,提取服务器中n个彩铃广告中第二个关键词作为第二簇头节点;

[0117] (1.2)成簇并建立特征簇头地图;

[0118] 将第一和第二簇头节点按照频率由高到低的顺序设置为 $i(x_i, y_i)$ 、 $j(x_j, y_j)$, x_i 为第一簇头节点频率值(关键词出现的次数除以关键词出现的总次数), y_i 为第一簇头节点顺

序号, x_j 为第二簇头节点频率值, y_j 为第二簇头节点序号,按照坐标形式建立成特征簇头地图; $i=1, 2, 3 \dots n$; $j=1, 2, 3 \dots n$;

[0119] (1.3)重新选取簇头;

[0120] 计算第一簇头到第二簇头之间的距离:

$$[0121] \quad D_{ij}(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2};$$

[0122] 每个节点根据自己的第二关键词被检索的次数进行加权和 W_s 计算,选择加权和最小的节点作为第二簇头节点,

$$[0123] \quad W_s = \delta D_s + \lambda \frac{1}{B};$$

$$[0124] \quad D_s = \frac{\sum_{j=1}^n D_{ij}(i, j)}{\sum_{i=1}^n \sum_{j=1}^n D_{ij}(i, j)};$$

$$[0125] \quad B = \frac{b_i}{\sum_{i=1}^n b_i + \sum_{j=1}^n b_j};$$

[0126] δ 和 λ 为加权系数, $\delta + \lambda = 1$, D_s 为第一簇头节点 (x_i, y_i) 与所有第二簇头节点距离总和占所有簇头节点间距离综合的比重; B 表示第一簇头节点 (x_i, y_i) 被检索次数 b_i 占相关簇头节点被检索次数的比重;

[0127] 若新的第二簇头节点与步骤(1.1)中的簇头簇头节点不一致, 即簇头节点发生改变, 则重新执行步骤(1.1); 若簇头节点没有改变, 则继续执行步骤(1.4);

[0128] (1.4)新一轮簇头节点选取成功后, 具有相同第二簇头节点的彩铃广告为一簇, 彩铃广告收集模块按照簇的序号发送至彩铃广告WEB展示与服务模块。

[0129] 该步骤详细的解释了对彩铃广告的成簇过程, 简单梳理了服务器中广告彩铃的分类, 有助于彩铃广告的有序性排列。

[0130] 针对彩铃广告排序模块根据簇节点信息对彩铃广告进行再次排序, 具体包括:

[0131] (2.1)分别计算第一簇头节点与第二簇头节点在同一彩铃广告内出现的次数, 设第一簇头节点在彩铃广告内出现 p 次, 第二簇头节点出现了 q 次, 判断 p 与 q 两个值的大小, 取两者比较的小值作为在同一彩铃广告内的共现次数, 同时将值赋给 r_i , 第一簇头节点与第二簇头节点在服务器内所有彩铃广告中出现次数 r 为:

$$[0132] \quad r = r_1 + r_2 + r_3 + \dots + r_N = \sum_{i=1}^N r_i;$$

[0133] N 为服务器内所有彩铃广告的总数, 当第一簇头节点与第二簇头节点未在同一彩铃广告内出现时, $r_i = 0$;

[0134] (2.2)计算第一簇头节点到第二簇头节点之间的预计距离:

$$[0135] \quad D = \frac{D_{\max} + D_{\min}}{2};$$

[0136] D_{\max} 为 D 的最大值, 为 D 的最小值,

[0137] 第一簇头节点与第二簇头节点在彩铃广告的距离取所有彩铃广告内距离的平均值;

$$[0138] \quad \bar{D} = \frac{\bar{D} * r_s + D}{r_s + 1};$$

[0139] r_s 为第一簇头节点与第二簇头节点共现在彩铃广告的次数;

[0140] 第二簇头节点在彩铃广告中的信息增益为：

$$[0141] \quad \alpha_j = - \sum_{j=1}^N \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}} \ln \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}};$$

[0142] β_{jt} 为第二簇头节点在文档t中出现的次数；

[0143] (2.3) 计算第一簇头节点与第二簇头节点的相关度为：

$$[0144] \quad \beta_{ij} = \frac{r \frac{\alpha_j}{\alpha_j + \bar{D}}}{\alpha_j};$$

[0145] (2.4) 计算第一簇头节点在彩铃广告中的权重；

$$[0146] \quad \gamma = \beta_{ij} \times \log \frac{N + 0.05}{r_s};$$

[0147] (2.5) 将第一簇头节点按照与第二簇头节点的权重值进行排序，根据第一簇头节点的排序对彩铃广告进行排序。

[0148] 本发明根据相关度和权重在简单梳理了彩铃广告排序后进一步进行了精准排序，因此本发明分两次进行有效彩铃广告的排序能够进一步提高广告的检索精度。

[0149] 彩铃广告在线和离线检索模块根据用户的搜索关键词与第一簇头节点和第二簇头节点进行匹配，具体包括：

[0150] (3.1) 根据彩铃广告在线和离线检索模块的获得检索词提取同义词库：

$$[0151] \quad W = \{W_0, W_1, W_2, \dots, W_V\};$$

[0152] w_v 为检索词的同义词； w_0 为有歧义的检索词， w_0 的歧义包括：

$$[0153] \quad \text{mean}(w_0) = \{w_{01}, w_{02}, w_{03}, \dots, w_{0m}\};$$

[0154] m为歧义词的数量；

[0155] (3.2) 估计W和语义 S_q 之间的相关度，

$$[0156] \quad C_S = \arg \max_{1 \leq j \leq m} \text{Rele}(W, G(s_j)); \quad j \in \{1, \dots, q\}$$

[0157] $\text{Rele}(W, G(s_j))$ 表示同义词库与语义 S_j 的相关词集合 $G(s_j)$ 的相关度， $G(s_j)$

表示通过推理机得到的语义 S_j 的相关词集合；

$$[0158] \quad Rele(W, G(s_j)) = \sum_{j=1}^v \frac{Wei(w_j, s_j)}{D(w_j, w_0)}$$

[0159] Wei表示 w_j 与 s_j 之间的相关度, $D(w_j, w_0)$ 表示 w_j 与 w_0 之间的距离;

[0160] (3.3)通过相关度阈值 Δ 确定同义词库与检索词之间的匹配关系是否成立,如果 $C_s \geq \Delta$,则同义词库与检索词之间相关;如果 $C_s < \Delta$,则依次删除 $mean(w_0)$ 中的歧义词,重新执行步骤(3.1),直到 $C_s \geq \Delta$;

[0161] (3.4)将匹配结果发送至彩铃广告推送模块。

[0162] 通过语义检索的方式,本发明提供了一种能够提供检索词和关键词相对应的方法,本发明能够进一步准确的将用户想要的结果检索出来并推送至用户面前。

[0163] 彩铃广告推送模块将匹配成功的彩铃广告发送至用户客户端供用户使用包括:

[0164] 彩铃广告推送模块获取匹配结果信息,根据匹配结果信息获取彩铃广告并发送给用户;若用户进行了下载或确认操作则确定对应结果,将检测结果储存于服务器或者云端以备下次使用。

[0165] 推送信息包括彩铃广告名称、音频、出版信息和格式。

[0166] 一种彩铃广告推送系统,包括彩铃广告收集模块、彩铃广告排序模块、彩铃广告WEB展示与服务模块、彩铃广告在线和离线检索模块、彩铃广告推送模块和服务器,服务器用于推送彩铃广告,彩铃广告收集模块基于关键词将彩铃广告分类整理成簇发送至彩铃广告排序模块;彩铃广告收集模块与服务器连接,选取服务器中 n 个彩铃广告作为 n 个节点,每个彩铃广告的第一个关键词作为第一簇头节点,提取服务器中 n 个彩铃广告中第二个关键词作为第二簇头节点;将第一和第二簇头节点按照频率由高到低的顺序设置为 $i(x_i, y_i)$ 、, x_i 为第一簇头节点频率值(关键词出现的次数除以关键词出现的总次数), y_i 为第一簇头节点顺序号, x_j 为第二簇头节点频率值, y_j 为第二簇头节点顺序号,按照坐标形式建立成特征簇头地图; $i=1,2,3\dots n$; $j=1,2,3\dots n$;计算第一簇头到第二簇头之间的距离:

$$[0167] \quad D_{ij}(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2};$$

[0168] 每个节点根据自己的第二关键词被检索的次数进行加权和 W_s 计算,选择加权和最小的节点作为第二簇头节点,

$$[0169] \quad W_s = \delta D_s + \lambda \frac{1}{B};$$

$$[0170] \quad D_s = \frac{\sum_{j=1}^n D_{ij}(i,j)}{\sum_{i=1}^n \sum_{j=1}^n D_{ij}(i,j)};$$

$$[0171] \quad B = \frac{b_i}{\sum_{i=1}^n b_i + \sum_{j=1}^n b_j};$$

[0172] δ 和 λ 为加权系数, $\delta + \lambda = 1$, D_s 为第一簇头节点 (x_i, y_i) 与所有第二簇头节点距离总和占有所有簇头节点间距离综合的比重; B 表示第一簇头节点 (x_i, y_i) 被检索次数 b_i 占相关簇头节点被检索次数的比重;

[0173] 若新的第二簇头节点与簇头节点不一致, 即簇头节点发生改变, 则重新选取初始特征簇头; 若簇头节点没有改变, 则新一轮簇头节点选取成功后, 具有相同第二簇头节点的彩铃广告为一簇, 彩铃广告收集模块按照簇的序号发送至彩铃广告WEB展示与服务模块;

[0174] 彩铃广告排序模块根据簇节点信息对彩铃广告进行再次排序并发送至彩铃广告WEB展示与服务模块; 分别计算第一簇头节点与第二簇头节点在同一彩铃广告内出现的次数, 设第一簇头节点在彩铃广告内出现 p 次, 第二簇头节点出现了 q 次, 判断 p 与 q 两个值的大小, 取两者比较的小值作为在同一彩铃广告内的共现次数, 同时将值赋给 r_i , 第一簇头节点与第二簇头节点在服务器内所有彩铃广告中出现次数 r 为:

$$[0175] \quad r = r_1 + r_2 + r_3 + \dots + r_N = \sum_{i=1}^N r_i;$$

[0176] N 为服务器所有彩铃广告的总数, 当第一簇头节点与第二簇头节点未在同一彩铃广告内出现时, $r_i = 0$;

[0177] 计算第一簇头节点到第二簇头节点之间的预计距离:

$$[0178] \quad D = \frac{D_{\max} + D_{\min}}{2};$$

[0179] D_{\max} 为 D 的最大值, D_{\min} 为 D 的最小值,

[0180] 第一簇头节点与第二簇头节点在彩铃广告的距离取所有彩铃广告内距离的平均值;

$$[0181] \quad \bar{D} = \frac{\bar{D} * r_s + D}{r_s + 1};$$

[0182] r_s 为第一簇头节点与第二簇头节点共现在彩铃广告的次数;

[0183] 第二簇头节点在彩铃广告中的信息增益为:

$$[0184] \quad \alpha_j = - \sum_{j=1}^N \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}} \ln \frac{\beta_{jt}}{\sum_{j=1}^N \beta_{jt}};$$

[0185] β_{jt} 为第二簇头节点在文档t中出现的次数;

[0186] 计算第一簇头节点与第二簇头节点的相关度为:

$$[0187] \quad \beta_{ij} = \frac{r \frac{\alpha_j}{\alpha_j + \bar{D}}}{\alpha_j};$$

[0188] 计算第一簇头节点在彩铃广告中的权重;

$$[0189] \quad \gamma = \beta_{ij} \times \log \frac{N + 0.05}{r_s};$$

[0190] 将第一簇头节点按照与第二簇头节点的权重值进行排序,根据第一簇头节点的排序对彩铃广告进行排序。

[0191] 彩铃广告在线和离线检索模块根据用户的搜索关键词与第一簇头节点和第二簇头节点进行匹配,并将匹配结果发送至彩铃广告推送模块;根据彩铃广告在线和离线检索模块获取检索词提取同义词库:

$$[0192] \quad W = \{w_0, w_1, w_2, \dots, w_v\};$$

[0193] w_v 为检索词的同义词; w_0 为有歧义的检索词, w_0 的歧义包括:

$$[0194] \quad \text{mean}(w_0) = \{w_{01}, w_{02}, w_{03}, \dots, w_{0m}\};$$

[0195] m为歧义词的数量;

[0196] 估计W和语义 s_q 之间的相关度,

$$[0197] \quad C_s = \arg \max_{1 \leq j \leq m} \text{Rele}(W, G(s_j)); \quad j \in \{1, \dots, q\}$$

[0198] $\text{Rele}(W, G(s_j))$ 表示同义词库与语义 s_j 的相关词集合 $G(s_j)$ 的相关度, $G(s_j)$

表示通过推理机得到的语义 s_j 的相关词集合;

$$[0199] \quad \text{Rele}(W, G(s_j)) = \sum_{j=1}^v \frac{\text{Wei}(w_j, s_j)}{D(w_j, w_0)};$$

[0200] Wei 表示 w_j 与 s_j 之间的相关度, $D(w_j, w_0)$ 表示 w_j 与 w_0 之间的距离;

[0201] 通过相关度阈值 Δ 确定同义词库与检索词之间的匹配关系是否成立, 如果 $C_S \geq \Delta$, 则同义词库与检索词之间相关; 如果 $C_S < \Delta$, 则依次删除 $mean(w_0)$ 中的歧义词, 直到 $C_S \geq \Delta$;

[0202] 将匹配结果发送至彩铃广告推送模块;

[0203] 彩铃广告推送模块将匹配成功的彩铃广告发送至用户客户端供用户使用; 彩铃广告推送模块获取匹配结果信息, 根据匹配结果信息获取彩铃广告的推送信息并发送给用户; 若用户进行了下载或确认操作则确定对应结果, 将检测结果储存于服务器或者云端以备下次使用; 推送信息包括: 彩铃广告名称、音频、出版信息、格式。

[0204] 需要指出的是, 本发明只是详细描述有关彩铃广告推送的有关技术, 而推送系统中的其他技术问题, 如信息的提取、发送、云端或者服务器的存储由于都是本领域的公知技术, 因此不在本发明中做进一步描述。

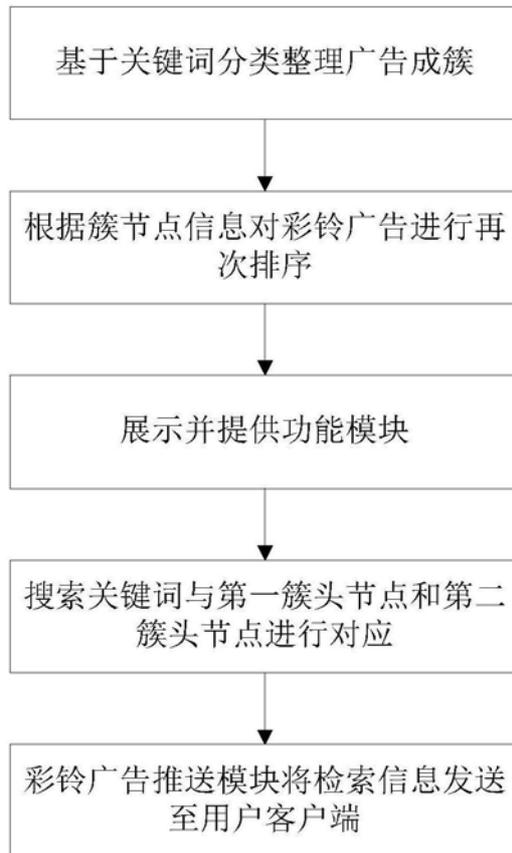


图1