



(12)发明专利申请

(10)申请公布号 CN 106649787 A

(43)申请公布日 2017. 05. 10

(21)申请号 201611235701.9

(22)申请日 2016.12.28

(71)申请人 北京奇虎科技有限公司

地址 100088 北京市西城区新街口外大街
28号D座112室(德胜园区)

(72)发明人 彭睿棋 彭露

(74)专利代理机构 北京市立方律师事务所

11330

代理人 王增鑫

(51) Int. Cl.

G06F 17/30(2006.01)

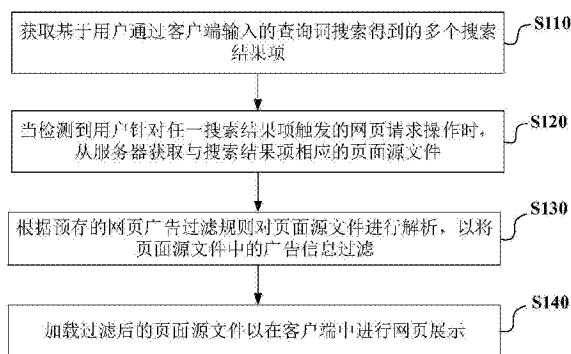
权利要求书2页 说明书8页 附图2页

(54)发明名称

移动终端客户端中广告过滤方法及装置

(57)摘要

本发明提供了移动终端客户端中广告过滤方法及装置,包括:获取基于用户通过客户端输入的查询词搜索得到的多个搜索结果项;当检测到用户针对任一搜索结果项触发的网页请求操作时,从服务器获取与搜索结果项相应的页面源文件;根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤;加载过滤后的页面源文件以在客户端中进行网页展示。通过本发明实施例提供的移动终端客户端中广告过滤方法,能够减少冗余信息的干扰,降低用户浏览网页的成本,从而提高用户体验。



1. 一种移动终端客户端中广告过滤方法,其特征在于,包括:
获取基于用户通过客户端输入的查询词搜索得到的多个搜索结果项;
当检测到所述用户针对任一搜索结果项触发的网页请求操作时,从服务器获取与所述搜索结果项相应的页面源文件;
根据预存的网页广告过滤规则对所述页面源文件进行解析,以将页面源文件中的广告信息过滤;
加载过滤后的页面源文件以在所述客户端中进行网页展示。
2. 根据权利要求1所述的方法,其特征在于,根据预存的网页广告过滤规则对所述页面源文件进行解析,包括:
根据预存的网页广告过滤规则判断所述页面源文件是否属于预置的网页地址白名单;
若是,则禁止对属于所述网页地址白名单的页面源文件进行过滤。
3. 根据权利要求1或2所述的方法,其特征在于,根据预存的网页广告过滤规则对所述源文件进行解析,以将页面源文件中的广告信息过滤,包括:
当判断所述页面源文件不属于所述网页地址白名单时,根据预存的网页广告过滤规则判断所述页面源文件中的任一统一资源定位符指向的页面元素是否包含广告信息;
若是,将指向的页面元素包含广告信息的统一资源定位符过滤。
4. 根据权利要求1-3中任一项所述的方法,其特征在于,所述方法还包括:
当判断所述页面源文件不属于所述网页地址白名单时,根据预存的网页广告过滤规则判断所述页面源文件中包括的任一页面元素是否属于预定类型的页面元素;
若是,则加载所述页面元素;
当检测到加载过滤后的页面源文件完成后,删除所述页面元素相应的加载结果。
5. 根据权利要求1-4中任一项所述的方法,其特征在于,所述页面元素包括脚本语言、层叠样式表、多媒体信息中的至少一项。
6. 根据权利要求1-5中任一项所述的方法,其特征在于,从服务器获取所述预存的网页广告过滤规则。
7. 一种移动终端客户端中广告过滤装置,其特征在于,包括:
第一获取模块,用于获取基于用户通过客户端输入的查询词搜索得到的多个搜索结果项;
第二获取模块,用于当检测到所述用户针对任一搜索结果项触发的网页请求操作时,从服务器获取与所述搜索结果相应的页面源文件;
解析模块,用于根据预存的网页广告过滤规则对所述页面源文件进行解析,以将页面源文件中的广告信息过滤;
加载模块,用于加载过滤后的页面源文件以在所述客户端中进行网页展示。
8. 根据权利要求7所述的装置,其特征在于,所述解析模块包括:
白名单判断单元,用于根据预存的网页广告过滤规则判断所述页面源文件是否属于预置的网页地址白名单;
禁止过滤单元,用于当所述页面源文件属于预置的网页地址白名单时,禁止对属于所述网页地址白名单的页面源文件进行过滤。
9. 根据权利要求7或8所述的装置,其特征在于,所述解析模块包括:

元素判断单元,用于当判断所述页面源文件不属于所述网页地址白名单时,根据预存的网页广告过滤规则判断所述页面源文件中的任一统一资源定位符指向的页面元素是否包含广告信息;

信息过滤单元,用于当所述页面源文件中的任一统一资源定位符指向的页面元素包含广告信息时,将指向的页面元素包含广告信息的统一资源定位符过滤。

10. 根据权利要求7-9中任一项所述的装置,其特征在于,所述装置还包括:

判断模块,用于当判断所述页面源文件不属于所述网页地址白名单时,根据预存的网页广告过滤规则判断所述页面源文件中包括的任一页面元素是否属于预定类型的页面元素;

加载模块,用于当所述页面源文件中包括的任一页面元素属于预定类型的页面元素时,加载所述页面元素;

删除模块,用于当检测到加载过滤后的页面源文件完成后,删除所述页面元素相应的加载结果。

11. 根据权利要求7-10中任一项所述的装置,其特征在于,所述页面元素包括脚本语言、层叠样式表、多媒体信息中的至少一项。

12. 根据权利要求7-11中任一项所述的装置,其特征在于,所述装置还包括:

第三获取模块,用于从服务器获取所述预存的网页广告过滤规则。

移动终端客户端中广告过滤方法及装置

技术领域

[0001] 本发明涉及终端设备技术领域,具体而言,本发明涉及一种移动终端客户端中广告过滤方法以及一种移动终端客户端中广告过滤装置。

背景技术

[0002] 随着智能终端设备功能的丰富以及移动互联网的快速发展,智能终端设备早已从单纯的通话设备变成了包含强大功能的移动设备。人们在工作之余可以通过智能终端设备播放音乐和视频、拍摄影像或上网浏览新闻。智能终端设备已然取代了多种常用电子设备,并且改变着人们的生活方式。

[0003] 智能终端设备上大多数的网页页面都包含广告信息。然而,由于智能终端设备屏幕较小,因此通过智能终端设备进行网页浏览的成本较高。此外,通过手动的方式关闭网页页面中的广告的操作相对繁琐,给用户带来了使用负担。

[0004] 例如,当用户输入查询词进行信息搜索时,浏览器除下载搜索结果页面的源文件之外,还会在用户浏览该搜索结果页面的过程中,在该搜索结果页面中嵌入第三方广告或其他不相干网页,并且自动触发浏览器获取第三方广告或其他不相干网页的源文件。此时虽然用户没有触发打开第三方广告或其他不相干网页网站地址的操作,但是也会自动展示第三方广告或其他不相干网页。

发明内容

[0005] 为克服上述技术问题或者至少部分地解决上述技术问题,特提出以下技术方案:

[0006] 本发明的一个实施例提出了一种移动终端客户端中广告过滤方法,包括:

[0007] 获取基于用户通过客户端输入的查询词搜索得到的多个搜索结果项;

[0008] 当检测到用户针对任一搜索结果项触发的网页请求操作时,从服务器获取与搜索结果项相应的页面源文件;

[0009] 根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤;

[0010] 加载过滤后的页面源文件以在客户端中进行网页展示。

[0011] 优选地,根据预存的网页广告过滤规则对页面源文件进行解析,包括:

[0012] 根据预存的网页广告过滤规则判断页面源文件是否属于预置的网页地址白名单;

[0013] 若是,则禁止对属于网页地址白名单的页面源文件进行过滤。

[0014] 优选地,根据预存的网页广告过滤规则对源文件进行解析,以将页面源文件中的广告信息过滤,包括:

[0015] 当判断页面源文件不属于网页地址白名单时,根据预存的网页广告过滤规则判断页面源文件中的任一统一资源定位符指向的页面元素是否包含广告信息;

[0016] 若是,将指向的页面元素包含广告信息的统一资源定位符过滤。

[0017] 可选地,该方法还包括:

- [0018] 当判断页面源文件不属于网页地址白名单时,根据预存的网页广告过滤规则判断页面源文件中包括的任一页面元素是否属于预定类型的页面元素;
- [0019] 若是,则加载页面元素;
- [0020] 当检测到加载过滤后的页面源文件完成后,删除页面元素相应的加载结果。
- [0021] 优选地,页面元素包括脚本语言、层叠样式表、多媒体信息中的至少一项。
- [0022] 优选地,从服务器获取预存的网页广告过滤规则。
- [0023] 本发明的另一实施例提出了一种移动终端客户端中广告过滤装置,包括:
- [0024] 第一获取模块,用于获取基于用户通过客户端输入的查询词搜索得到的多个搜索结果项;
- [0025] 第二获取模块,用于当检测到用户针对任一搜索结果项触发的网页请求操作时,从服务器获取与搜索结果相应的页面源文件;
- [0026] 解析模块,用于根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤;
- [0027] 加载模块,用于加载过滤后的页面源文件以在客户端中进行网页展示。
- [0028] 优选地,解析模块包括:
- [0029] 白名单判断单元,用于根据预存的网页广告过滤规则判断页面源文件是否属于预置的网页地址白名单;
- [0030] 禁止过滤单元,用于当页面源文件属于预置的网页地址白名单时,禁止对属于网页地址白名单的页面源文件进行过滤。
- [0031] 优选地,解析模块包括:
- [0032] 元素判断单元,用于当判断页面源文件不属于网页地址白名单时,根据预存的网页广告过滤规则判断页面源文件中的任一统一资源定位符指向的页面元素是否包含广告信息;
- [0033] 信息过滤单元,用于当页面源文件中的任一统一资源定位符指向的页面元素包含广告信息时,将指向的页面元素包含广告信息的统一资源定位符过滤。
- [0034] 可选地,该装置还包括:
- [0035] 判断模块,用于当判断页面源文件不属于网页地址白名单时,根据预存的网页广告过滤规则判断页面源文件中包括的任一页面元素是否属于预定类型的页面元素;
- [0036] 加载模块,用于当页面源文件中包括的任一页面元素属于预定类型的页面元素时,加载页面元素;
- [0037] 删除模块,用于当检测到加载过滤后的页面源文件完成后,删除页面元素相应的加载结果。
- [0038] 优选地,页面元素包括脚本语言、层叠样式表、多媒体信息中的至少一项。
- [0039] 可选地,该装置还包括:
- [0040] 第三获取模块,用于从服务器获取预存的网页广告过滤规则。
- [0041] 通过本发明实施例提供的移动终端客户端中广告过滤方法,根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤,通过加载过滤后的页面源文件可以将包含少量广告信息的甚至是不包含广告信息的与网页请求相应的网页页面进行展示,以减少冗余信息的干扰,降低用户浏览网页的成本,从而提高用户体验。

[0042] 本发明附加的方面和优点将在下面的描述中部分给出,这些将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0043] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0044] 图1为本发明一个实施例的移动终端客户端中广告过滤方法的流程示意图;

[0045] 图2为本发明一个优选实施例的移动终端客户端中广告过滤方法的流程示意图;

[0046] 图3为本发明另一优选实施例的移动终端客户端中广告过滤方法的流程示意图;

[0047] 图4为本发明一个实施例的移动终端客户端中广告过滤装置的结构示意图。

具体实施方式

[0048] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,仅用于解释本发明,而不能解释为对本发明的限制。

[0049] 本技术领域技术人员可以理解,除非特意声明,这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是,本发明的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件,但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。应该理解,当我们称元件被“连接”或“耦接”到另一元件时,它可以直接连接或耦接到其他元件,或者也可以存在中间元件。此外,这里使用的“连接”或“耦接”可以包括无线连接或无线耦接。这里使用的措辞“和/或”包括一个或多个相关联的列出项的全部或任一单元和全部组合。

[0050] 本技术领域技术人员可以理解,除非另外定义,这里使用的所有术语(包括技术术语和科学术语),具有与本发明所属领域中的普通技术人员的一般理解相同的意义。还应该理解的是,诸如通用字典中定义的那些术语,应该被理解为具有与现有技术的上下文中的意义一致的意义,并且除非像这里一样被特定定义,否则不会用理想化或过于正式的含义来解释。

[0051] 本技术领域技术人员可以理解,这里所使用的“终端”、“终端设备”既包括无线信号接收器的设备,其仅具备无发射能力的无线信号接收器的设备,又包括接收和发射硬件的设备,其具有能够在双向通信链路上,进行双向通信的接收和发射硬件的设备。这种设备可以包括:蜂窝或其他通信设备,其具有单线路显示器或多线路显示器或没有多线路显示器的蜂窝或其他通信设备;PCS(Personal Communications Service,个人通信系统),其可以组合语音、数据处理、传真和/或数据通信能力;PDA(Personal Digital Assistant,个人数字助理),其可以包括射频接收器、寻呼机、互联网/内联网访问、网络浏览器、记事本、日历和/或GPS(Global Positioning System,全球定位系统)接收器;常规膝上型和/或掌上型计算机或其他设备,其具有和/或包括射频接收器的常规膝上型和/或掌上型计算机或其他设备。这里所使用的“终端”、“终端设备”可以是便携式、可运输、安装在交通工具(航空、海运和/或陆地)中的,或者适合于和/或配置为在本地运行,和/或以分布形式,运行在地球和/或空间的任何其他位置运行。这里所使用的“终端”、“终端设备”还可以是通信终端、上

网终端、音乐/视频播放终端,例如可以是PDA、MID (Mobile Internet Device,移动互联网设备)和/或具有音乐/视频播放功能的移动电话,也可以是智能电视、机顶盒等设备。

[0052] 通过移动终端浏览的网页页面的源文件包括HTML (HyperText Markup Language, 超文本标记语言) 文件。HTML是用来描述网页的一种语言。在请求网页时,移动终端可以根据与网页请求操作相应的URL (Uniform Resource Locator, 统一资源定位符) 来下载与上述网页请求操作相应的HTML文件并解析HTML文件。需要说明的是,URL可以包括模式(或协议)部分、网站域名(或IP地址)部分、路径部分和文件名部分等,简单的URL可以由字母组成,如“www.XXX.com”,或由数字组成,如192.46.20.40。在具体应用场景中,移动终端中的浏览器通过URL从服务器请求网页。

[0053] 进一步地,为了增强网页页面的动态性和交互性,可以在部分页面源文件中增加JS (JavaScript, 脚本语言) 元素。另外,还可以在页面源文件中增加用于渲染网页页面样式的CSS (Cascading Style Sheets, 层叠样式表) 元素、多媒体信息。其中,多媒体信息可以包括图片、视频信息、音频信息等。故而,如果页面源文件包括JS元素、CSS元素或多媒体信息等页面元素,在移动终端对页面源文件进行解析的过程中,当解析到需要加载的页面元素时,可以从服务器下载上述页面元素。

[0054] 需要说明的是,每一个页面源文件或页面元素的下载都对应着一个URL。在一个网页页面的展示过程中,移动终端需要从服务器下载数十个甚至上百个页面源文件和页面元素。因此,一个网页页面的展示过程可以涉及数十个甚至上百个URL。

[0055] 图1为本发明一个实施例的移动终端客户端中广告过滤方法的流程示意图。

[0056] 步骤S110:获取基于用户通过客户端输入的查询词搜索得到的多个搜索结果项。

[0057] 例如,用户在移动终端如智能手机中安装了“XX搜索”客户端,当用户需要对某一词汇进行搜索时,可以在该客户端中输入该词汇并进行搜索。此时,智能手机可以获取上述词汇并基于上述词汇进行搜索以得到与该词汇相应的多个搜索结果项。

[0058] 步骤S120:当检测到用户针对任一搜索结果项触发的网页请求操作时,从服务器获取与网页请求相应的页面源文件。

[0059] 需要说明的是,用户针对任一搜索结果项触发的网页请求操作可以包括用户对搜索结果项进行点击的操作等。

[0060] 例如,当检测到用户在移动终端客户端中输入查询词“中国美食”时,获取基于“中国美食”搜索得到的多个搜索结果项并展示多个搜索结果项;接着,用户对多个搜索结果项中的一个产生兴趣时,点击感兴趣的搜索结果项,其中,与该结果项相应的网站域名为“www.XXX.com”,移动终端可以生成与包含网站域名“www.XXX.com”的网页请求信息;接着,可以将网页请求信息发送至与网站域名“www.XXX.com”相应的服务器;随后,该服务器接收到来自移动终端的网页请求信息,并返回与网页请求信息相应的HTML文件。

[0061] 步骤S130:根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤。

[0062] 可选地,可以从服务器获取预存的网页广告过滤规则。需要说明的是,首先,服务器可以对海量网页页面进行分析并且通过分析结果确定出哪些页面源文件中的URL包含广告信息;随后,可以将广告信息进行分类;接着,通过不同的标识信息将与不同类型的广告信息相应的URL进行标记,对于不同类型的广告信息相应的URL存在不同的处理方式;最后,

基于URL、与不同类型的广告信息相应的URL的标识信息以及与标识信息相应的URL处理方式生成网页广告过滤规则并下发至移动终端。

[0063] 优选地,如图2所示,根据预存的网页广告过滤规则对页面源文件进行解析的步骤包括步骤S210和步骤S220:步骤S210:根据预存的网页广告过滤规则判断页面源文件是否属于预置的网页地址白名单;步骤S220:若是,则禁止对属于网页地址白名单的页面源文件进行过滤。

[0064] 需要说明的是,移动终端不需要对展示在某些特殊网站的网页广告进行过滤。例如,在网站“淘宝网”的网页页面的侧边栏展示的产品广告信息不作为广告过滤,即不需要对该网站的网页页面进行过滤。因此,服务器在生成网页广告过滤规则时,可以将“淘宝网”的网页地址加入网页地址白名单。当服务器将包含“淘宝网”的网页地址的网页地址白名单的网页广告过滤规则下发至移动终端时,移动终端接收并存储该网页广告过滤规则。故而,当移动终端打开网站“淘宝网”的网页页面并下载“淘宝网”的页面源文件时,根据上述网页广告过滤规则可以判断“淘宝网”的页面源文件属于预置的网页地址白名单。进而,禁止对属于网页地址白名单的“淘宝网”的页面源文件进行过滤,以避免对特殊网站的网页广告的错误。

[0065] 然而,对于一般的需要对网页广告进行过滤的网站,可以通过下述方法进行过滤:

[0066] 优选地,如图3所示,根据预存的网页广告过滤规则对源文件进行解析,以将页面源文件中的广告信息过滤的步骤包括步骤S310和步骤S320:步骤S310:当判断页面源文件不属于网页地址白名单时,根据预存的网页广告过滤规则判断页面源文件中的任一统一资源定位符指向的页面元素是否包含广告信息;步骤S320:若是,将指向的页面元素包含广告信息的统一资源定位符过滤。

[0067] 优选地,页面元素包括但不限于脚本语言、层叠样式表、多媒体信息中的至少一项。

[0068] 需要说明的是,当判断页面源文件中的任一统一资源定位符指向的页面元素包含广告信息时,在解析页面源文件的过程中,可以跳过上述统一资源定位符不对其进行加载。

[0069] 例如,在请求网页A时,移动终端从服务器获取网页A的源文件index.HTML,其中,index.HTML中包括URL“http://image.baidu.com/search/detail?ct=503316480”,该URL指向的页面元素为图片信息B;接着,根据预存的网页广告过滤规则判断网页A的网页地址是否属于预置的网页地址白名单;当判断网页A的网页地址不属于预置的网页地址白名单时,对网页A进行过滤并解析index.HTML;在解析index.HTML的过程中,当解析到上述URL时,判断该URL指向的图片信息B是否包含广告信息;当判断图片信息B包含广告信息时,可以跳过上述URL不对其进行加载,以实现不展示包含广告信息的图片信息B。

[0070] 可选地,本发明实施例提供的方法还包括步骤S150、步骤S160和步骤S170:步骤S150:当判断页面源文件不属于网页地址白名单时,根据预存的网页广告过滤规则判断页面源文件中包括的任一页面元素是否属于预定类型的页面元素;步骤S160:若是,则加载页面元素;步骤S170:当检测到加载过滤后的页面源文件完成后,删除页面元素相应的加载结果。

[0071] 需要说明的是,对于部分特殊的包含广告信息的页面元素,例如必须加载一个特定的包含广告信息的JS元素才能正常加载整个网页页面,首先可以通过JS注入的方式对上

述JS元素进行加载,随后再删除与该JS元素相应的加载结果,以达到先正常加载整个网页页面再删除广告部分的目的。

[0072] 在服务器生成网页广告过滤规则时,可以将上述部分特殊的包含广告信息的页面元素作为预定类型的页面元素放入网页广告过滤规则中。因此,当服务器将网页广告过滤规则下发至移动终端时,移动终端接收网页广告过滤规则并可以依据该网页广告过滤规则判断页面源文件中包括的任一页面元素是否属于预定类型的页面元素。

[0073] 步骤S140:加载过滤后的页面源文件以在客户端中进行网页展示。

[0074] 通过本发明实施例提供的移动终端客户端中广告过滤方法,根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤,通过加载过滤后的页面源文件可以将包含少量广告信息的甚至是不包含广告信息的与网页请求相应的网页页面进行展示,以减少冗余信息的干扰,降低用户浏览网页的成本,从而提高用户体验。

[0075] 图4为本发明一个实施例的移动终端客户端中广告过滤装置的结构示意图。该装置可以包括:第一获取模块410、第二获取模块420、解析模块430和加载模块440。以下针对各个模块的具体作用做进一步的说明:

[0076] 第一获取模块410获取基于用户通过客户端输入的查询词搜索得到的多个搜索结果项。

[0077] 例如,用户在移动终端如智能手机中安装了“XX搜索”客户端,当用户需要对某一词汇进行搜索时,可以在该客户端中输入该词汇并进行搜索。此时,智能手机可以获取上述词汇并基于上述词汇进行搜索以得到与该词汇相应的多个搜索结果项。

[0078] 第二获取模块420当检测到用户针对任一搜索结果项触发的网页请求操作时,从服务器获取与网页请求相应的页面源文件。

[0079] 需要说明的是,用户针对任一搜索结果项触发的网页请求操作可以包括用户对搜索结果项进行点击的操作等。

[0080] 例如,当检测到用户在移动终端客户端中输入查询词“中国美食”时,获取基于“中国美食”搜索得到的多个搜索结果项并展示多个搜索结果项;接着,用户对多个搜索结果项中的一个产生兴趣时,点击感兴趣的搜索结果项,其中,与该结果项相应的网站域名为“www.XXX.com”,移动终端可以生成与包含网站域名“www.XXX.com”的网页请求信息;接着,可以将网页请求信息发送至与网站域名“www.XXX.com”相应的服务器;随后,该服务器接收到来自移动终端的网页请求信息,并返回与网页请求信息相应的HTML文件。

[0081] 解析模块430根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤。

[0082] 可选地,本发明实施例提供的装置还包括第三获取模块:第三获取模块从服务器获取预存的网页广告过滤规则。需要说明的是,首先,服务器可以对海量网页页面进行分析并且通过分析结果确定出哪些页面源文件中的URL包含广告信息;随后,可以将广告信息进行分类;接着,通过不同的标识信息将与不同类型的广告信息相应的URL进行标记,对于不同类型的广告信息相应的URL存在不同的处理方式;最后,基于URL、与不同类型的广告信息相应的URL的标识信息以及与标识信息相应的URL处理方式生成网页广告过滤规则并下发至移动终端。

[0083] 优选地,解析模块420包括白名单判断单元和禁止过滤单元:白名单判断单元根据

预存的网页广告过滤规则判断页面源文件是否属于预置的网页地址白名单；禁止过滤单元当页面源文件属于预置的网页地址白名单时，禁止对属于网页地址白名单的页面源文件进行过滤。

[0084] 需要说明的是，移动终端不需要对展示在某些特殊网站的网页广告进行过滤。例如，在网站“淘宝网”的网页页面的侧边栏展示的产品广告信息不作为广告过滤，即不需要对该网站的网页页面进行过滤。因此，服务器在生成网页广告过滤规则时，可以将“淘宝网”的网页地址加入网页地址白名单。当服务器将包含“淘宝网”的网页地址的网页地址白名单的网页广告过滤规则下发至移动终端时，移动终端接收并存储该网页广告过滤规则。故而，当移动终端打开网站“淘宝网”的网页页面并下载“淘宝网”的页面源文件时，根据上述网页广告过滤规则可以判断“淘宝网”的页面源文件属于预置的网页地址白名单。进而，禁止对属于网页地址白名单的“淘宝网”的页面源文件进行过滤，以避免对特殊网站的网页广告的错误。

[0085] 然而，对于一般的需要对网页广告进行过滤的网站，可以通过下述装置进行过滤：

[0086] 优选地，解析模块420包括元素判断单元和信息过滤单元：元素判断单元当判断页面源文件不属于网页地址白名单时，根据预存的网页广告过滤规则判断页面源文件中的任一统一资源定位符指向的页面元素是否包含广告信息；信息过滤单元当页面源文件中的任一统一资源定位符指向的页面元素包含广告信息时，将指向的页面元素包含广告信息的统一资源定位符过滤。

[0087] 优选地，页面元素包括但不限于脚本语言、层叠样式表、多媒体信息中的至少一项。

[0088] 需要说明的是，当判断页面源文件中的任一统一资源定位符指向的页面元素包含广告信息时，在解析页面源文件的过程中，可以跳过上述统一资源定位符不对其进行加载。

[0089] 例如，在请求网页A时，移动终端从服务器获取网页A的源文件index.HTML，其中，index.HTML中包括URL“http://image.baidu.com/search/detail?ct=503316480”，该URL指向的页面元素为图片信息B；接着，根据预存的网页广告过滤规则判断网页A的网页地址是否属于预置的网页地址白名单；当判断网页A的网页地址不属于预置的网页地址白名单时，对网页A进行过滤并解析index.HTML；在解析index.HTML的过程中，当解析到上述URL时，判断该URL指向的图片信息B是否包含广告信息；当判断图片信息B包含广告信息时，可以跳过上述URL不对其进行加载，以实现不展示包含广告信息的图片信息B。

[0090] 可选地，本发明实施例提供的装置还包括判断模块、加载模块和删除模块：判断模块当判断页面源文件不属于网页地址白名单时，根据预存的网页广告过滤规则判断页面源文件中包括的任一页面元素是否属于预定类型的页面元素；加载模块当页面源文件中包括的任一页面元素属于预定类型的页面元素时，加载页面元素；删除模块当检测到加载过滤后的页面源文件完成后，删除页面元素相应的加载结果。

[0091] 需要说明的是，对于部分特殊的包含广告信息的页面元素，例如必须加载一个特定的包含广告信息的JS元素才能正常加载整个网页页面，首先可以通过JS注入的方式对上述JS元素进行加载，随后再删除与该JS元素相应的加载结果，以达到先正常加载整个网页页面再删除广告部分的目的。

[0092] 在服务器生成网页广告过滤规则时，可以将上述部分特殊的包含广告信息的页面

元素作为预定类型的页面元素放入网页广告过滤规则中。因此,当服务器将网页广告过滤规则下发至移动终端时,移动终端接收网页广告过滤规则并可以依据该网页广告过滤规则判断页面源文件中包括的任一页面元素是否属于预定类型的页面元素。

[0093] 加载模块440加载过滤后的页面源文件以在客户端中进行网页展示。

[0094] 通过本发明实施例提供的移动终端客户端中广告过滤装置,根据预存的网页广告过滤规则对页面源文件进行解析,以将页面源文件中的广告信息过滤,通过加载过滤后的页面源文件可以将包含少量广告信息的甚至是不包含广告信息的与网页请求相应的网页页面进行展示,以减少冗余信息的干扰,降低用户浏览网页的成本,从而提高用户体验。

[0095] 本技术领域技术人员可以理解,本发明包括涉及用于执行本申请中所述操作中的一项或多项的设备。这些设备可以为所需的目的而专门设计和制造,或者也可以包括通用计算机中的已知设备。这些设备具有存储在其内的计算机程序,这些计算机程序选择性地激活或重构。这样的计算机程序可以被存储在设备(例如,计算机)可读介质中或者存储在适于存储电子指令并分别耦联到总线的任何类型的介质中,所述计算机可读介质包括但不限于任何类型的盘(包括软盘、硬盘、光盘、CD-ROM、和磁光盘)、ROM(Read-Only Memory,只读存储器)、RAM(Random Access Memory,随机存储器)、EPROM(Erasable Programmable Read-Only Memory,可擦写可编程只读存储器)、EEPROM(Electrically Erasable Programmable Read-Only Memory,电可擦可编程只读存储器)、闪存、磁性卡片或光线卡片。也就是,可读介质包括由设备(例如,计算机)以能够读的形式存储或传输信息的任何介质。

[0096] 本技术领域技术人员可以理解,可以用计算机程序指令来实现这些结构图和/或框图和/或流图中的每个框以及这些结构图和/或框图和/或流图中的框的组合。本技术领域技术人员可以理解,可以将这些计算机程序指令提供给通用计算机、专业计算机或其他可编程数据处理方法的处理器来实现,从而通过计算机或其他可编程数据处理方法的处理器来执行本发明公开的结构图和/或框图和/或流图的框或多个框中指定的方案。

[0097] 本技术领域技术人员可以理解,本发明中已经讨论过的各种操作、方法、流程中的步骤、措施、方案可以被交替、更改、组合或删除。进一步地,具有本发明中已经讨论过的各种操作、方法、流程中的其他步骤、措施、方案也可以被交替、更改、重排、分解、组合或删除。进一步地,现有技术中的具有与本发明中公开的各种操作、方法、流程中的步骤、措施、方案也可以被交替、更改、重排、分解、组合或删除。

[0098] 以上所述仅是本发明的部分实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

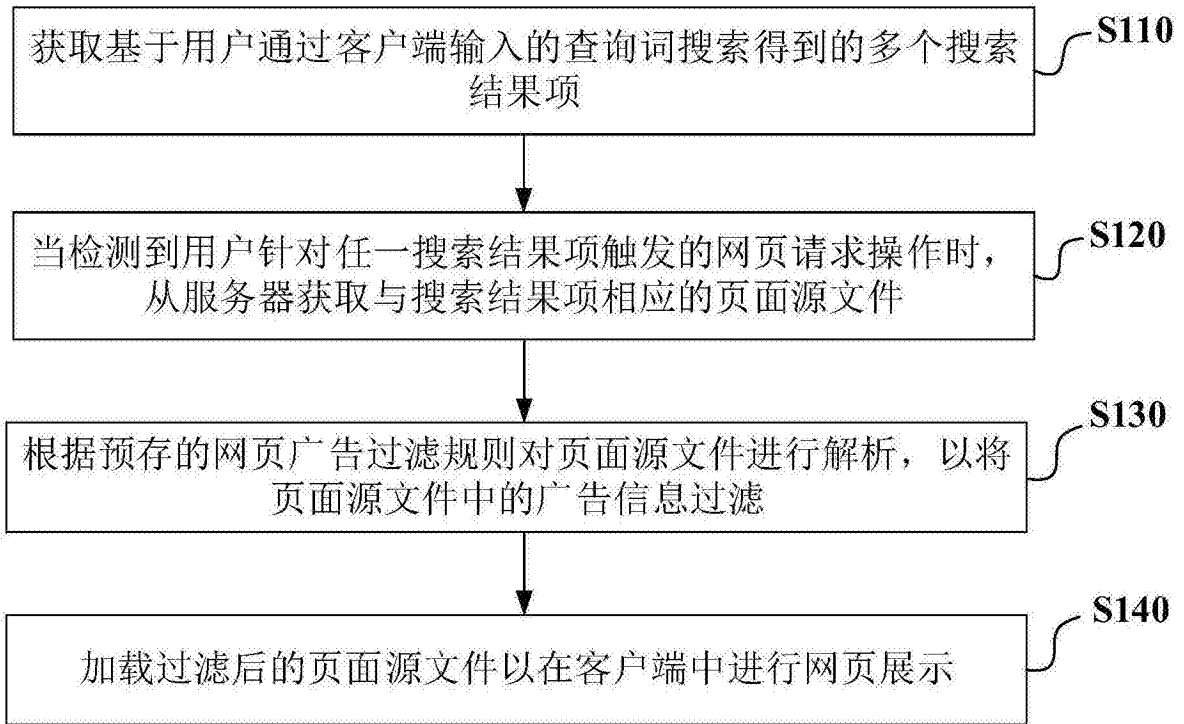


图1

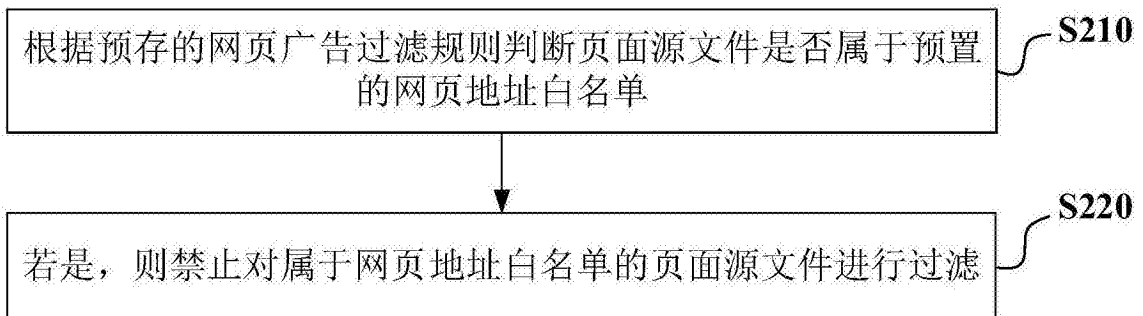


图2

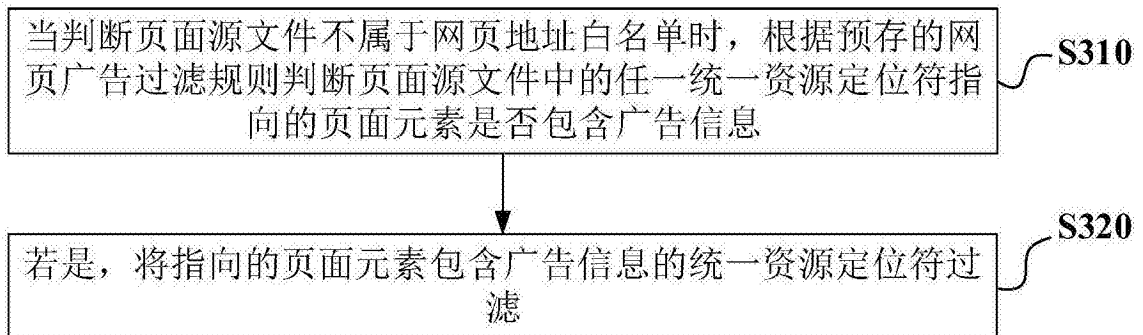


图3

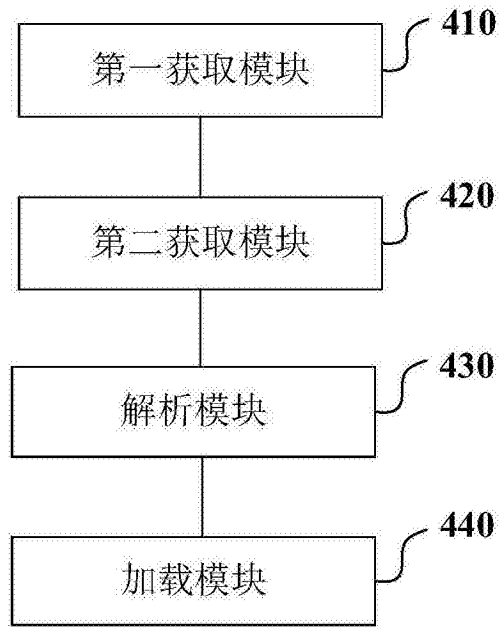


图4