



LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

(84) **指定国** (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT,

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

本发明公开了一种基于云-端协同的移动浏览器资源加载优化方法。本方法为: 1)在用户的手机端建立一代理服务器, 在该用户的个人云上搭建一云服务器; 2)手机端代理服务器将接收并解析浏览器的 URL 请求, 如果手机端代理服务器已缓存该 URL 请求对应的资源, 或者通过 URL 匹配算法计算出存在对应的资源, 则直接将该资源返回给浏览器; 否则将其发送给云服务器; 3)云服务器检查是否已缓存该 URL 请求对应的资源, 如果已缓存则将该资源返回给代理服务器; 否则向该 URL 请求对应的目标网站进行资源请求并将结果返回给手机端代理服务器; 4)该手机端代理服务器将收到的返回结果依照 HTTP 响应的结构发回浏览器。本发明大大提高了加载速度。

一种基于云-端协同的移动浏览器资源加载优化方法

技术领域

本发明是一种基于云-端协同的移动浏览器资源加载优化方法，属于软件技术领域。

背景技术

根据中国互联网络信息中心于 2015 年 1 月发布的《第 35 次中国互联网络发展状况统计报告》，截至 2014 年 12 月，我国手机网民规模达 5.57 亿，网民中使用手机上网的人群占比已达到 85.8%。连续几年来，手机网民规模保持稳定增长，手机也早已成为第一大上网终端。

越来越多的人们正在利用手机或者平板电脑来通过 Wi-Fi 或 3G 网络上网，随之而来的，Web 浏览器成为了移动设备上最常用的应用之一。据研究表明，浏览器占据了移动设备 63% 的显示时间。但是，现有移动浏览器的用户体验却远不尽如人意。资源加载是影响浏览器性能重要因素之一，有实验表明，网页加载 65% 的时间都消耗在资源加载上。由于移动设备的网络状况通常并不理想，尤其是可能出现网络拥挤或是用户正在移动的状况，这个问题就变得更为严重了。此外，下载资源的过程本身会消耗许多电量，而电量恰恰是移动设备上最为稀缺的资源。

为了加速 Web 资源加载，缓存机制被广泛运用。通过缓存的方法，可以为用户减少资源下载量，从而提高 Web 浏览性能。这种缓存机制给移动设备带来的好处更大，因为在移动设备上，该机制不仅可以减少网页加载时间，还可以减少网络流量和电量消耗。然而在实际运用中，缓存却常常并未被很好地配置与实现。有研究表明，缓存机制的理论性能与实际性能之间有很大差距因此当前有这样一个需求，在不改变现有浏览器与服务器实现的前提下，对现有的缓存机制进行优化，提高缓存的利用效果。希望可以以此进一步提高移动设备用户在浏览网页时的体验，减少用户的流量与电量消耗。

发明内容

针对现有技术中存在的问题，本发明的目的在于提供一种基于云-端协同的移动浏览器资源加载优化方法。本发明是在不改变现有用户浏览习惯与 Web 应用缓存原有配置的前提下，最大程度减少 Web 资源的传输流量并提高加载速度。其核心思想是：在原有的 Web 应用架构的基础上，在手机端搭建代理服务器，并借助云服务器预取 Web 资源，即将原有的从浏览器到服务器的两层架构改变为四层架构。

该系统在手机端搭建 HTTP 代理服务器，并将手机端浏览器的 HTTP 代理配置为该手机端代理服务器；在个人云上为每个手机用户搭建云服务器。手机端代理服务器接收手机端浏览器的请求，并向手机端浏览器返回资源响应。手机端代理服务器并不直接向互联网请求资源，而向在个人云上搭建的云服务器请求资源，由云服务器从互联网上下载资源，并有选择地向手机端进行推送。在此过程中，通过 URL 匹配算法、资源校验码比对、数据压缩与网站预取等手段达到最大程度减少传输流量并提高加载速度的目的。

本发明的技术方案为：

一种云-端协同的移动浏览器缓存优化方法，具体步骤为：

- 1) 在手机端建立代理服务器，该代理服务器可以接收浏览器的请求，并解析请求，对于代理服务器已缓存的资源，或者通过 URL 匹配算法计算出存在对应的资源，可以直接将资源返回给浏览器；对于本地未存储的资源，向云端服务器进行资源的查询。
- 2) 在云端为每个手机用户建立一个云服务器，该云服务器可以进行 Internet 连接。云服务器可以预取网站的资源，并接收手机端代理服务器的资源请求。如果资源已经被缓存在云服务器上，则云服务器直接向手机端代理服务器返回该资源，如果没有，则向目标网站进行请求并将结果返回。
- 3) 用户在手机端将浏览器的代理配置为手机端的代理服务器。

进一步的，每一个资源都会在云服务器端计算出该文件的校验值，作为标识资源内容的唯一依据，可以依据该值对缓存资源进行检索。在服务器端将资源传回手机端时，会先检索该校验值的文件是否已存在于手机端的缓存中，如果已存在，则仅返回该资源的校验值，手机端代理服务器会根据该校验值提取资源。

进一步的，云端服务器会定时预取用户可能访问的网站，提前建立网站的缓存以供用户访问。

进一步的，所有服务器到手机端的传输均经过压缩，以节省未经压缩资源的传输流量。

使用本发明的框架，资源访问主要包含三个步骤：

1. 预取目标资源：云服务器端会定期对用户可能访问的资源进行预取。
2. 接收浏览器请求并检查手机端缓存：手机端代理服务器接收浏览器请求，如果手机端没有缓存资源，则向云端服务器请求资源；如果该缓存存在或可以根据 URL 匹配算法找到相应资源，则直接向浏览器返回资源。
3. 检查云服务器端缓存：如果云服务器端没有缓存资源，则向原网站请求，然后将资源传回手机端代理服务器。手机端代理服务器再将结果依照 HTTP 响应的结构发回浏览器。

上述步骤的具体技术方案如下：

1. 预取目标资源：服务器端会对用户可能访问的资源进行提前预取，具体的资源列表会根据用户的访问习惯产生。服务器会将预取好的资源存储，同时计算好校验值，并将条目整理到数据库中。

2. 代理服务器接收浏览器请求并检查手机端缓存：本发明可以与浏览器的原有缓存同时使用，浏览器会将原本需要向远程服务器请求的资源发到手机端代理服务器上。手机端的代理服务器会接收请求并根据 HTTP 协议对请求进行解析。之后在本地检测资源是否已被手机端代理服务器缓存。如果缓存不存在，手机会将解析到的资源 URL 通过 URL 匹配算法在缓存列表中进行查询，如果查询到相关资源，则返回资源，否则将资源请求通过 HTTP2 或 SPDY 协议发送给服务器。

URL 匹配算法包含两部分，URL 模式识别与 URL 匹配查询，其中模式识别在云服务器端进行，匹配查询在手机端进行。

云服务器会对同一网站定时进行的抓取与解析，由此可以得到 URL 和资源的对应列表。同一资源对应的 URL 可能会有多个。本算法的目的在于根据 URL 检索到潜在对应的资源，因此需要建立每个资源对应的 URL 模式。算法首先对 URL 根据常见 URL 分隔符进行切分，对于同一资源，计算得出 URL 每一部分发生变化时，该资源保持不变的可能性，这个值我们称为段落关键系数，计算方法为：

$$\text{段落关键系数} = \frac{\text{段落出现的次数}}{\text{相应资源对应 URL 总数} + 1}$$

对于所有资源，该值永远介于 0 至 1 之间。

当服务器将资源与其 URL 发回手机端时，该 URL 中所有段落的关键系数都将被传回。

当手机端代理服务接收到一个 URL 的请求时，会首先进行匹配查询，将该 URL 与相应网站的所有 URL 记录进行相似度的对比（先根据该 URL 找到相应网站，然后将该网站的所有 URL 记录进行对比），相似度的计算方法如下：

$$\text{相似度} = \prod_{i=\text{URL 内的每个段落}} (1 - \text{段落 } i \text{ 的关键系数} * \alpha_i)$$

其中 α_i 表示检索中的待请求 URL 的段落是否与目标 URL 的段落匹配，如果不匹配则为 1，匹配则为 0。特别地，当待请求 URL 与目标 URL 完全相同时，该相似度结果为 1。当相似度结果大于 0.5 时，我们认为目标 URL 的资源就是待请求 URL 的资源，并取最大相似度的结果作为待请求 URL 的资源，如果所有资源的相似度均不足 0.5，则认为匹配失败。缓存中

不存在该资源，需要发往服务器进一步确认。

3. 检查服务器端缓存：服务器端接收到手机端的资源请求后，会根据相关算法计算出该 URL 对应的数据库条目，或者 URL 并不在数据库中，即资源并未缓存。如果资源没有被缓存，服务器会向原网站请求该资源，得到资源后，会将资源的校验值计算出并将条目存储到数据库中。

4. 检查资源是否在手机端已存在并向浏览器发送响应：在服务器端准备好资源后，会依照资源的校验码在服务器端存储的手机资源列表中检索资源是否存在，如果资源存在，则不会返回数据体，只会返回资源对应的校验值，手机端会根据此校验值从本地列表中提取出相应资源。如果资源不存在，则服务器端会将资源压缩并通过 HTTP2 或 SPDY 协议将内容返回给手机端。手机端接收到数据后会进行解压缩，并按照 HTTP 响应的标准得到该资源的响应，最终将相应发回给浏览器。

与现有技术相比，本发明的积极效果为：

本发明通过四层架构进行 Web 浏览的优化，可以减少传输流量并提高加载速度。实验表明，使用该框架，在用户第一次访问网页时，平均可以节省流量 17%，提高加载速度 29%；再次访问同一网页，平均可以为第二次访问节省流量 66%，提高加载速度 9%。

附图说明

图 1 为本发明的系统架构图。

具体实施方式

本节以某新闻网站为例，给出用户利用本发明提供的框架先后两次访问该新闻网站的后台操作过程。

云服务器端会定期预取该网站的该资源，当用户第一次访问该网站时，浏览器会先把 HTML 的请求发送给手机端代理服务器，由于 HTML 属于最经常更新的资源，对于 HTML 的请求，手机端会默认此前的缓存不可用，而向服务器发起询问。代理服务解析出 URL 后，向云端的服务器发起询问，由于云端的服务器此前已经预取过该资源，服务器可以根据 URL 检索到该资源。此后服务器会在手机端资源列表中检索该资源，由于此前并未访问过此网站，该资源不存在手机端上，因此服务器会将资源进行压缩，并传回手机端代理服务器。手机端代理服务器会解析服务器传回的内容，并将其整理为 HTTP 响应，发回浏览器。浏览器将解析 HTML 的内容，并向代理服务发送其余资源的请求。对于剩余的资源，由于此前并未访问

过该网站，缓存中不存在这些资源，通过 URL 匹配算法也无法计算得出与该资源相似的已缓存资源，因此手机端代理服务也会向服务器请求这些资源。

在第二次访问该网站时，浏览器会再次把 HTML 的请求发送给手机端代理，手机端会将该资源的请求发给云服务器，云服务器在本地检索到该资源后，再根据校验值在手机缓存列表中检索，如果 HTML 本身的内容并未改变，该资源的校验值会在手机缓存列表中存在，该资源本身将不会被返回，否则需要返回整个资源，并在手机端更新该缓存。

对于 HTML 解析出的某些资源，虽然在手机端不存在该 URL 的缓存数据，但这种 URL 的改变可能是由于网页的动态解析造成的，资源本身并未改变，通过 URL 匹配算法，可以检测到这类资源 URL 的改变从而返回相应的已缓存资源。如果算法判定找不到相关资源，手机端仍需将该请求发送给服务器进行询问。

当手机端将该资源的请求发送给云服务器端后，若云服务器检索或请求到该资源之后，发现该资源对应的校验值在手机端的资源列表中已存在，则不会返回该资源本身，而只会向手机端返回响应头和校验值。手机端会接收到该信息，然后从本地提取出资源本体，将两者（该响应头和资源本体）按照 HTTP 相应的结构进行重新组合，并将此内容发给浏览器。

权利要求书

1. 一种基于云-端协同的移动浏览器资源加载优化方法，其步骤为：
 - 1) 在用户的手机端建立一代理服务器，在该用户的个人云上搭建一云服务器；
 - 2) 手机端代理服务器将接收并解析浏览器的 URL 请求，如果手机端代理服务器已缓存该 URL 请求对应的资源，或者通过 URL 匹配算法计算出存在对应的资源，则直接将该资源返回给浏览器；否则将该 URL 请求发送给云服务器；
 - 3) 云服务器检查是否已缓存该 URL 请求对应的资源，如果已缓存则将该资源返回给该手机端代理服务器；否则向该 URL 请求对应的目标网站进行资源请求并将结果返回给该手机端代理服务器；
 - 4) 该手机端代理服务器将收到的返回结果依照 HTTP 响应的结构发回浏览器。
2. 如权利要求 1 所述的方法，其特征在于，所述云服务器为一虚拟机。
3. 如权利要求 1 或 2 所述的方法，其特征在于，所述云服务器将资源与其 URL 返回手机端代理服务器时，将该 URL 相关的段落关键系数同时传回。
4. 如权利要求 3 所述的方法，其特征在于，所述段落关键系数为（段落的出现次数）/（相同资源对应 URL 总数+1）。
5. 如权利要求 1 或 2 所述的方法，其特征在于，当该代理服务接收到一个 URL 请求时，先进行匹配查询，将该 URL 与相应网站的所有 URL 记录进行相似度对比，将相似度最大且超过设定阈值的 URL 记录对应的缓存资源返回给浏览器。
6. 如权利要求 5 所述的方法，其特征在于，所述

$$\text{相似度} = \prod_{i=\text{URL 内的每个段落}} (1 - \text{段落 } i \text{ 的关键系数} * \alpha_i)$$
 ；其中， α_i 表示检索中的待请求 URL 的段落是否与目标 URL 的段落匹配，如果不匹配则为 1，匹配则为 0。
7. 如权利要求 1 所述的方法，其特征在于，所述云服务器计算获取的资源的校验值并将该资源及其校验值存储到数据库中。
8. 如权利要求 1 所述的方法，其特征在于，所述云服务器根据用户的访问习惯生成该用户的资源列表，然后根据该资源列表进行资源预取，并将预取好的资源及其校验值存储到数据库中。
9. 如权利要求 7 或 8 所述的方法，其特征在于，所述云服务器将资源返回手机端代理服务器之前，检查该资源对应的校验值在手机端的资源列表中是否已存在，如果存在则返回响应

头和该资源的校验值；该手机端代理服务器接收到该信息后从本地提取出资源本体，将该响应头和资源本体按照 HTTP 相应的结构进行重新组合，并将此内容发给浏览器；如果不存在，则将该资源及其校验值发送给手机端代理服务器。

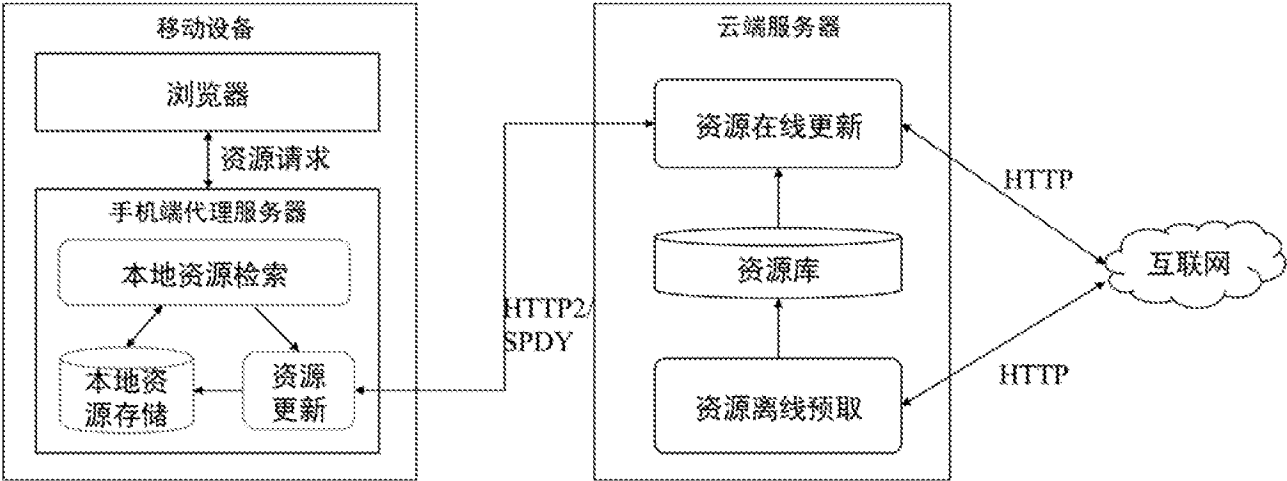


图 1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/098289

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04W; H04Q; H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: cloud, server, brows??, increas???, optimiz???, mobile, web, agency server, URL, cache, webpage, speed
up, match, paragraph

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105610909 A (PEKING UNIVERSITY) 25 May 2016 (25.05.2016) claims 1-9	1-9
X	CN 103795690 A (HUAWEI TECHNOLOGIES CO., LTD.) 14 May 2014 (14.05.2014) description, paragraphs [0002], [0016]-[0020], [0030], [0031] and [0044]	1, 2, 4, 6-9
Y	CN 103795690 A (HUAWEI TECHNOLOGIES CO., LTD.) 14 May 2014 (14.05.2014) description, paragraphs [0002], [0016]-[0020], [0030], [0031] and [0044]	3, 5
Y	CN 101814086 A (SHANDONG NORMAL UNIVERSITY) 25 August 2010 (25.08.2010) description, paragraphs [0011] and [0012]	3, 5

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search
07 November 2016

Date of mailing of the international search report
30 December 2016

Name and mailing address of the ISA
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No. (86-10) 62019451

Authorized officer

SUN, Wei

Telephone No. (86-10) 62413854

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/CN2016/098289

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101163156 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 16 April 2008 (16.04.2008) description, page 5, line 22 to page 6, line 16	1, 2, 4, 6-9
A	CN 104202403 A (ANYI HENG TONG BEIJING TECHNOLOGY CO., LTD.) 10 December 2014 (10.12.2014) the whole document	1-9
A	US 2009177514 A1 (MICROSOFT CORP.) 09 July 2009 (09.07.2009) the whole document	1-9

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/098289

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105610909 A	25 May 2016	None	
CN 103795690 A	14 May 2014	None	
CN 101814086 A	25 August 2010	None	
CN 104202403 A	10 December 2014	None	
US 2009177514 A1	09 July 2009	US 2009178109 A1	09 July 2009
		US 2009178131 A1	09 July 2009
		US 2009178132 A1	09 July 2009
		US 2009178108 A1	09 July 2009

国际检索报告

国际申请号

PCT/CN2016/098289

A. 主题的分类

H04L 29/08(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

H04W; H04Q; H04L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT, CNKI, WPI, EPODOC: 云, 代理服务器, 缓存, 网页, 浏览, 加速, 优化, 提速, 匹配, 段落, cloud, server, brows??, increas??, optimiz??, mobile, web, agency server, URL

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 105610909 A (北京大学) 2016年 5月 25日 (2016 - 05 - 25) 权利要求1-9	1-9
X	CN 103795690 A (华为技术有限公司) 2014年 5月 14日 (2014 - 05 - 14) 说明书第[0002]、[0016]-[0020]、[0030]-[0031]、[0044]段	1、2、4、6-9
Y	CN 103795690 A (华为技术有限公司) 2014年 5月 14日 (2014 - 05 - 14) 说明书第[0002]、[0016]-[0020]、[0030]-[0031]、[0044]段	3、5
Y	CN 101814086 A (山东师范大学) 2010年 8月 25日 (2010 - 08 - 25) 说明书第[0011]-[0012]段	3、5
X	CN 101163156 A (腾讯科技深圳有限公司) 2008年 4月 16日 (2008 - 04 - 16) 说明书第5页第22行-第6页第16行	1、2、4、6-9
A	CN 104202403 A (安一恒通北京科技有限公司) 2014年 12月 10日 (2014 - 12 - 10) 全文	1-9
A	US 2009177514 A1 (MICROSOFT CORP.) 2009年 7月 9日 (2009 - 07 - 09) 全文	1-9

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2016年 11月 7日

国际检索报告邮寄日期

2016年 11月 30日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

孙伟

电话号码 01062413854

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/098289

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	105610909	A	2016年 5月 25日	无			
CN	103795690	A	2014年 5月 14日	无			
CN	101814086	A	2010年 8月 25日	无			
CN	104202403	A	2014年 12月 10日	无			
US	2009177514	A1	2009年 7月 9日	US	2009178109	A1	2009年 7月 9日
				US	2009178131	A1	2009年 7月 9日
				US	2009178132	A1	2009年 7月 9日
				US	2009178108	A1	2009年 7月 9日

表 PCT/ISA/210 (同族专利附件) (2009年7月)