

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2023 年 8 月 31 日 (31.08.2023)



(10) 国际公布号
WO 2023/159986 A1

(51) 国际专利分类号:
H04W 28/14 (2009.01) *H04L 67/568* (2022.01)
H04W 4/44 (2018.01) *H04L 67/5683* (2022.01)
H04W 4/46 (2018.01) *H04W 40/32* (2009.01)
H04W 4/06 (2009.01) *H04L 45/00* (2022.01)

(21) 国际申请号: PCT/CN2022/126959

(22) 国际申请日: 2022 年 10 月 24 日 (24.10.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210185518.1 2022年2月28日 (28.02.2022) CN

(71) 申请人: 南京邮电大学(NANJING UNIVERSITY OF
POSTS AND TELECOMMUNICATIONS) [CN/CN];

中国江苏省南京市鼓楼区新模范马路
66号, Jiangsu 210003 (CN)。

(72) 发明人: 赵海涛(ZHAO, Haitao); 中国江苏省南京市鼓楼区新模范马路66号, Jiangsu 210003 (CN)。 赵越(ZHAO, Yue); 中国江苏省南京市鼓楼区新模范马路66号, Jiangsu 210003 (CN)。 夏文超(XIA, Wenchao); 中国江苏省南京市鼓楼区新模范马路66号, Jiangsu 210003 (CN)。 倪艺洋(NI, Yiyang); 中国江苏省南京市鼓楼区新模范马路66号, Jiangsu 210003 (CN)。 王琴(WANG, Qin); 中国江苏省南京市鼓楼区新模范马路66号, Jiangsu 210003 (CN)。 朱洪波(ZHU, Hongbo); 中国江苏省南京市鼓楼区新模范马路66号, Jiangsu 210003 (CN)。

(54) Title: COLLABORATIVE CACHING METHOD IN HIERARCHICAL NETWORK ARCHITECTURE

(54) 发明名称: 一种分层网络体系结构中的协作缓存方法

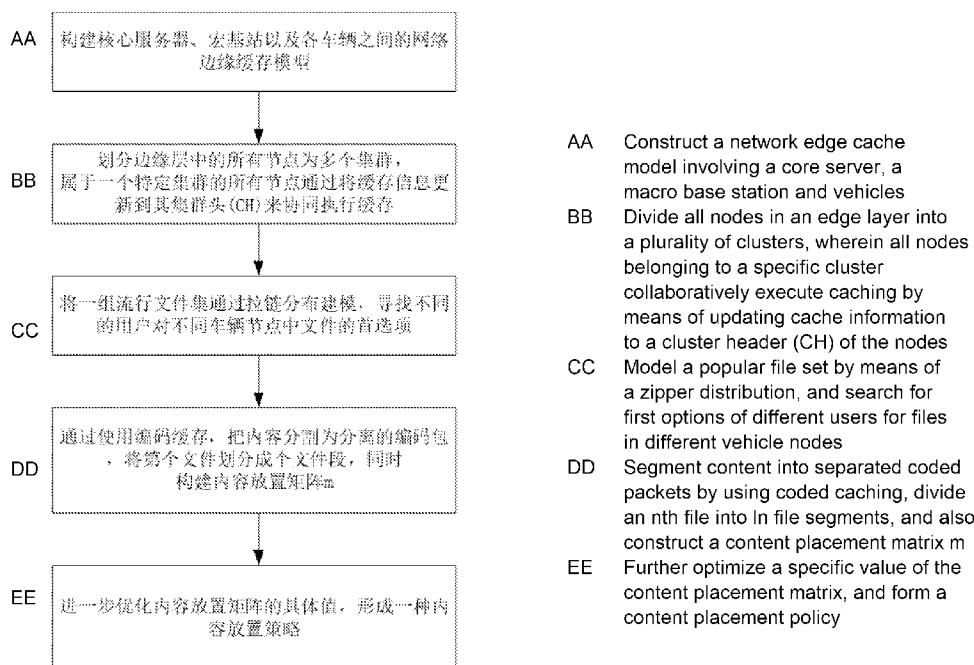


图 2

(57) Abstract: A collaborative caching method in a hierarchical network architecture. The method comprises: constructing a network edge cache model involving a core server, a macro base station and vehicles; dividing nodes in an edge layer into a plurality of clusters by area, wherein the number of nodes in each cluster does not exceed a threshold value β ; modeling a popular file set by means of a zipper distribution, and searching for the most popular files in different vehicle nodes; segmenting content into separated coded packets by using coded caching, and constructing a content placement matrix m; and further optimizing a specific value of the content placement

(74) 代理人: 南京经纬专利商标代理有限公司 (NANJING JINGWEI PATENT & TRADEMARK AGENCY CO., LTD.); 中国江苏省南京市鼓楼区中山路179号12楼B座, Jiangsu 210005 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, 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)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

matrix m by taking the minimization of the total energy consumption as an optimization objective, and formulating a content placement policy. The present method is aimed at selecting a content cache node and performing collaborative placement on content when the energy consumption is minimized, and thus achieves higher content transmission efficiency.

(57) 摘要: 一种分层网络体系结构中的协作缓存方法, 包括: 构建核心服务器、宏基站以及各车辆之间的网络边缘缓存模型; 将边缘层中的节点按区域划分为多个集群, 每个集群节点数不超过阈值 β ; 将一组流行文件集通过拉链分布建模, 寻找不同车辆节点中最流行的文件; 通过使用编码缓存, 把内容分割为分离的编码包, 构建内容放置矩阵 m ; 以最小化总能耗为优化目标, 进一步优化内容放置矩阵 m 的具体值, 制定内容放置策略。本方法旨在最小化能耗的情况下, 选择内容缓存节点, 并对内容进行协作放置, 具有更高的内容传递效率。

一种分层网络体系结构中的协作缓存方法

技术领域

本发明涉及车辆交通技术领域，具体而言涉及一种车联网环境下分层网络体系结构中的协作缓存方法。

背景技术

智能交通系统在当今社会发挥着至关重要的作用，大多数自动驾驶汽车用户的兴趣更加关注娱乐服务。用户在互联网上的请求出现了前所未有的增长，通常包括在不同的平台上观看和分享视频、音频和图像。为了获得更好的服务质量，无缝的内容交付显得至关重要。然而，在高度动态的环境下，很难实现长期的车对车（V2V）和车对基础设施（V2I）的连接。此外，路侧单元(RSU)和基站(BS)等基础设施组件的短传输范围将为 V2I 通信提供低效的解决方案。由于范围有限和高移动性，频繁的连接丢失和连接重建将带来大量的网络流量开销和内容检索延迟。

在物联网中，设备将连接到网络，并将收集到的数据共享到云服务器。在云服务器上进行数据分析和处理后，用户经常会请求有用的信息。不同的用户可以在同一时间段内请求相同的内容，从而导致相同的数据内容的重复传输。这促使运营商重新考虑他们当前的网络架构，并寻求更先进的技术，以低延迟和节能的方式使内容更接近最终用户。移动边缘计算/缓存在宏基站(MBS)覆盖范围内的设备，通过分层异构网络架构，使内容缓存降低能耗。

然而，由于边缘设备内存有限，以及车辆的高移动速度，在异构车辆网络中多层边缘缓存服务器(例如 MBS 和 RSU)的交错覆盖范围内，不同服务器之间的跨层协作可以提供无缝连接，以方便内容下载。

专利号为 CN112839082A 的发明中提及一种异构边缘缓存分配方法、系统、介质及应用，在考虑内容缓存分配的同时也将功率分配问题考虑在内，为不同的内容分配不同的传输功率，以达到用户体验的最优化。但该专利不涉及编码缓存的问题。

专利号为 CN113543071A 的发明中提及一种车联网环境下异构节点协作边缘缓存方法，旨在最小化车辆获取总内容所需延迟、成本以及高命中率的情况下，对内容进行协作放置。但该发明采用的是喷泉码。

发明内容

本发明针对现有技术中的不足，提供一种车联网环境下分层网络体系结构中的协作缓存方法，旨在最小化能耗的情况下，选择内容缓存节点，并对内容进行协作放置，具有更高的

内容传递效率。

为实现上述目的，本发明采用以下技术方案：

本发明实施例提出了一种分层网络体系结构中的协作缓存方法，所述协作缓存方法包括以下步骤：

S1，构建核心服务器、宏基站以及各车辆之间的网络边缘缓存模型，该网络边缘缓存模型设置在一个宏基站覆盖范围内，包含核心层和边缘层的两层典型的分层网络体系结构；其中，将该宏基站和其所包含的每个车辆均视为其中一个缓存节点，边缘层中的节点协作执行内容缓存；包括宏基站在内的核心层支持大带宽和快速的数据包传输的同时，缓存边缘层未缓存部分；

S2，将边缘层中的节点按区域划分为多个集群，每个集群节点数不超过阈值 β ；属于一个特定集群的所有节点通过将缓存信息更新到其集群头来协同执行缓存，节点阈值 β 与集群节点内的通信开销相关；

S3，将一组流行文件集通过拉链分布建模，寻找不同的用户对不同车辆节点中最流行的文件；

S4，通过使用编码缓存，把内容分割为分离的编码包，将第 n 个文件划分成 l_n 个文件段，且每个文件大小相等；同时将流行文件集中的文件按照受欢迎程度由高到低的顺序依次进行缓存，构建内容放置矩阵 m ；当用户 u_i 请求第 n 个文件时，其中 m_{ki} 个文件段从本地关联的簇中第 k 个车辆节点获得，剩余的 $l_n - m_{ki}$ 个文件段将通过宏基站的编码多播传输获得；

S5，以最小化总能耗为优化目标，进一步优化内容放置矩阵 m 的具体值，制定内容放置策略。

进一步地，步骤 S1 中，边缘层中的节点协作执行内容缓存是指，只允许选定的车辆缓存其他车辆缓存事件的信息。

进一步地，步骤 S2 中，将边缘层中的节点按区域划分为多个集群的过程包括以下步骤：

将整个边缘网络进行划分，同一片区域里的车辆节点分为一组，且每个集群节点数不超过阈值 β ，每组对应一个边缘簇，每个簇进一步分为各种节点类型：集群头节点、网关节点和集群成员；属于一个特定集群的所有节点通过将缓存信息更新到其集群头节点来协同执行缓存，集群成员之间不存在直接通信。

进一步地，集群头节点的选择过程包括以下步骤：

采用下述公式计算集群内每个节点的优先级评估值：

$$w_k = w_1 \frac{1/d_k}{\overline{M_d}} + w_2 \frac{T_k}{\overline{T}} + w_3 \frac{H_k}{\overline{H}} + w_4 \frac{p_k}{\overline{P}} + w_5 \frac{c_k}{\overline{C}}$$

式中， w 是不同参数对应的权重因子， $w_1 + w_2 + w_3 + w_4 + w_5 = 1$ ； d_k 是节点 k 范围内的节点数， $\overline{M_d}$ 是所有网络节点的节点度的平均逆； T_k 是节点 k 与其范围内的节点的传输时间， $\overline{T}$ 是所有网络节点的平均传输时间， H_k 是节点 k 与其范围内所有节点距离的加权和， $\overline{H}$ 是所有系统宽跳的平均加权和， p_k 是节点 k 的传输功率， $\overline{P}$ 是所有网络节点的平均功率利用率， c_k 是节点 k 的高速缓存利用率， $\overline{C}$ 是所有网络节点的平均缓存空间利用率。

进一步地，步骤 S3 中，将一组流行文件集通过拉链分布建模，寻找不同的用户对不同车辆节点中最流行的文件的过程包括以下步骤：

S31，根据下述公式计算一个用户请求第 k 个节点中的第 n 个文件的概率：

$$q_{kn} = \frac{(\gamma_{kn})^{-\alpha_k}}{\sum_{i=1}^N i^{-\alpha_k}}$$

式中， q_{kn} 表示第 k 个 SBS 的用户请求第 n 个文件的概率； α_k 是压缩指数，用于描述请求模式中的偏态， α_k 越大意味着内容请求更集中在高排名的内容文件上； γ_{kn} 是第 k 个节点中第 n 个文件的排名顺序；

S32，在每个传输周期 t 中，第 k 个车辆节点将平均接收来自其覆盖区域内所有相关用户的 U_k 请求；在第 k 个节点的覆盖范围内，第 n 个文件的平均请求数定义为 ω_{kn} ，用 $\omega_{kn} = U_k q_{kn}$ 表示；

S33，在多播周期 t 内，在第 k 个节点覆盖范围内的所有用户请求文件共 j 次的概率为：

$$P_{kn}^j = \frac{(\omega_{kn})^j e^{-\omega_{kn}}}{j!}。$$

进一步地，步骤 S4 中，所述内容放置矩阵 m 为：

$$m = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1N} \\ m_{21} & m_{22} & \cdots & m_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1} & m_{K2} & \cdots & m_{KN} \end{bmatrix}$$

式中, m_{kn} 表示每个簇内第 k 个车辆节点缓存第 n 个文件的文件段的数量; m 的值满足以下约束条件:

$$\sum_{n=1}^N m_{kn} s_0 \leq M_k, \forall k, \sum_{k=1}^K M_k \leq M$$

式中, M_k 表示每个车辆节点可以缓存的最大容量, M 表示总文件的大小。

进一步地, 步骤 S5 中, 优化内容放置矩阵的建模为:

$$\begin{aligned} \min & E_C + E_t^{MBS} + E_t^V \\ \text{s.t. C1: } & \sum_{n=1}^N m_{kn} s_0 \leq M / K, \forall k \\ & \text{C2: } 0 \leq m_{kn} \leq l_n, \forall k, n \end{aligned}$$

式中, 限制条件 C1 和 C2 是对缓存容量的限制, s_0 是每个文件段的大小, M 是总文件的

大小; E_C 是内容缓存的能量消耗, $E_C = e_{ca} \sum_{k=1}^K \sum_{n=1}^N m_{kn} s_0$; E_t^V 是各车辆节点传输的能耗,

$E_t^V = e_t^v \sum_{k=1}^K \sum_{n=1}^N (1 - P_{kn}^0) (m_{kn} s_0 + (l_n - m_{kn}) s_0)$; P_{kn}^0 表示用户在第 k 个节点范围内请求文件 n 次数为

零的概率; E_t^{MBS} 是 MBS 的传输能耗, $E_t^{MBS} = e_T^{MBS} B$, e_T^{MBS} 表示 MBS 传输的每比特数据的能量消耗, B 表示 MBS 通过编码组播传输的总比特数; N 为流行文件集的文件总数。

本发明的有益效果是:

本发明提出的分层网络体系结构中的协作缓存方法, 对于编码缓存, 在放置阶段, 内容未编码并存储在边缘服务器中, 但在传输阶段, 可以通过编码的多播包同时满足不同的用户请求。本发明可以实现最佳的内存率权衡。

附图说明

图 1 是本发明实施例的网络边缘缓存模型的结构示意图。

图 2 为本发明实施例的分层网络体系结构中的协作缓存方法流程图。

具体实施方式

现在结合附图对本发明作进一步详细的说明。

需要注意的是, 发明中所引用的如“上”、“下”、“左”、“右”、“前”、“后”等的用语, 亦仅为便于叙述的明了, 而非用以限定本发明可实施的范围, 其相对关系的改变或调整, 在

无实质变更技术内容下，当亦视为本发明可实施的范畴。

图 2 为本发明实施例的分层网络体系结构中的协作缓存方法流程图。参见图 2，该缓存方法包括以下步骤：

S1，构建核心服务器、宏基站以及各车辆之间的网络边缘缓存模型，该模型设置在一个宏基站覆盖范围内。

图 1 为具有缓存能力的、两层分层的车辆网络架构。基于图 1，本实施例的模型构建原理包括：设计由核心层和边缘层组成的两层典型的分层网络体系结构，核心层包含宏基站以及远程服务器，边缘层由车辆用户组成；将该宏基站和其所包含的每个车辆均视为其中一个缓存节点，边缘层中的节点协作执行内容缓存，核心层支持大带宽和快速的数据包传输，同时缓存边缘层未缓存部分。

将所有的网络用户用集合 $U = \{u_1, u_2, \dots, u_M\}$ 表示，该网络的所有节点都用 K 表示，其中， $k \in K$ 指定任何网络节点，为了有效地利用资源，只允许选定的车辆缓存其他车辆缓存事件的信息；将流行文件集用 $N = \{1, 2, \dots, N\}$ 表示；对应的文件大小用 $s = \{s_1, s_2, \dots, s_N\}$ 表示。

由于边缘层节点缓存空间有限，网络车辆速度快，用户只能接收所请求的多媒体文件的一部分，通过编码传输将剩余部分缓存在核心层。

S2，划分边缘层中的所有节点为多个集群。

每个集群进一步分为各种节点类型，集群头节点(CH)、网关节点(GN)和集群成员(CM)。属于一个特定集群的所有节点通过将缓存信息更新到其集群头(CH)来协同执行缓存，其中，与传统的协作缓存不同，单个集群的成员之间不可能存在直接通信。

CH 主要负责缓存其所有 CM 的更新后的缓存事件信息，并且通过 GN 提供自己集群中的节点和邻近集群中的节点的通信链接。GN 的主要职责是集群间的数据包传输，并且参与集群头的选择过程。剩余节点都被视为集群成员，CM 是根据内容的流行程度而实际进行缓存的节点，所有的 CM 都参与了集群头的选择过程。

聚类头 CH 选择方法：传统设计的 WCA 算法考虑四个因素来计算特定节点的节点度与节点阈值 β 之间的权重函数。对于 WCA，一个权重函数是使用所有因素一起计算的，这些因素反过来产生一个值来对每个网络节点进行评分。选择所有节点中得分最低的节点作为集群头。本专利也遵循 WCA 方法进行聚类，同时考虑了下述公式所示的不同因素：

$$w_k = w_1 \frac{1/d_k}{M_d} + w_2 \frac{T_k}{T} + w_3 \frac{H_k}{H} + w_4 \frac{p_k}{P} + w_5 \frac{c_k}{C}$$

式中, w_1 、 w_2 、 w_3 、 w_4 、 w_5 是不同参数对应的权重因子, 表现为:

$$w_1 + w_2 + w_3 + w_4 + w_5 = 1。$$

第一个因素对应于节点 k 的 d_k 度(即节点 k 的传输范围内的节点, 也称为节点 k 的邻居), 并计算表示为:

$$d_k = Neighbor(k) = \sum_{q \in K, q \neq k} dist(k, q) \quad dist(k, q) \in (0, 1)$$

式中 $dist(k, q)=1$ 对应于 q 在 k 的传输范围内, 所有网络节点 $\overline{M_d}$ 的节点度的平均逆可以用 $\overline{M_d} = \frac{1}{K} \sum_{k=1}^K \frac{1}{d_k}$ 表示。

第二个因素是从节点 k 到 k 范围内所有节点的平均传输时间 T_k 。这样, T_k 被计算表示为:

$$T_k = \frac{1}{S_k} \sum_{q \in S_k} t_{q,k}$$

式中, S_k 为 k 和 $t_{q,k}$ 范围内的节点集合, k 表示从节点 k 到任意节点 q 的传输时间, 所有网络节点的平均传输时间用 $\overline{T} = \frac{1}{K} \sum_{k=1}^K T_k$ 表示。

第三个因素是节点 k 与邻居跳点距离的加权和 H_k 。为了减少发送和接收缓存信息的网络负载, 从 CH 到 CM 的跳点计数应保持在最低限度。 H_k 的值是用 $H_k = \sum_{m=1}^{\infty} \frac{\gamma_m \times m}{count_m}$ 表示, 式中 $count_m$ 为节点 k 的第 m 跳个邻居的计数, γ_m 为权值因子。同样, 所有系统宽跳的平均加权和的计算用 $\overline{H} = \frac{1}{K} \sum_{i=1}^K H_i$ 表示。

第四个因素是指定节点 k 作为 CH 期间的累积功率 P_k , 所有网络节点的平均功率利用率可以计算为 $\overline{P} = \frac{1}{K} \sum_{k=1}^K P_k$ 。

第五个因素是节点 n 的高速缓存利用率 C_k , 主要是在 CH 中, 所有网络节点的平均缓存空间利用率可以计算为 $\overline{C} = \frac{1}{K} \sum_{k=1}^K C_k$ 。

集群成员 CM 选择方法：为了减少集群节点内的通信开销，为每个集群设置一个节点阈值 β ，即，一个集群内除 CH 外可容纳的最大节点数的上限。从而限制了集群可以管理的最大节点数量。

S3，将一组流行文件集通过拉链分布建模，寻找不同的用户对不同车辆节点中文件的首选选项。

不同的用户对不同车辆节点中的文件有不同的首选项，因此一个车辆节点中的流行文件在其他节点中可能不流行。假设文件的流行度分布遵循拉链式分布，根据下述公式计算一个用户请求第 k 个节点中的第 n 个文件的概率：

$$q_{kn} = \frac{(\gamma_{kn})^{-\alpha_k}}{\sum_{i=1}^N i^{-\alpha_k}}$$

式中， q_{kn} 表示第 k 个 SBS 的用户请求第 n 个文件的概率， α_k 是压缩指数，用于描述请求模式中的偏态， α_k 越大意味着内容请求更集中在高排名的内容文件上； γ_{kn} 是第 k 个节点中第 n 个文件的排名顺序。

在每个传输周期 t 中，第 k 个车辆节点将平均接收来自其覆盖区域内所有相关用户的 U_k 请求。在第 k 个节点的覆盖范围内，第 n 个文件的平均请求数定义为 ω_{kn} ，用 $\omega_{kn} = U_k q_{kn}$ 表示。

用户在第 k 个车辆节点覆盖范围中对第 n 个文件的请求数遵循泊松过程。因此，在多播周期 t 内，在第 k 个节点覆盖范围内的所有用户请求文件共 j 次的概率可以写为：

$$P_{kn}^j = \frac{(\omega_{kn})^j e^{-\omega_{kn}}}{j!}$$

S4，通过使用编码缓存，把内容分割为分离的编码包。

由于每个集群节点缓存大小的限制和文件的异构性，边缘层中每个节点只能缓存部分文件。通过使用编码缓存，把内容分割为分离的编码包，将第 n 个文件划分成 l_n 个文件段，且每个文件段大小相等为 s_0 ， m_{kn} 表示每个簇内第 k 个边缘设备缓存第 n 个文件的文件段的数量，因此，内容放置矩阵 m 可以表示为：

$$m = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1N} \\ m_{21} & m_{22} & \cdots & m_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1} & m_{K2} & \cdots & m_{KN} \end{bmatrix}$$

用户 u_i 请求第 n 个文件，则需要得到它的 l_n 个文件段， m_{kn} 个文件段从本地关联的簇中第 k 个车辆节点获得，剩余的 $l_n - m_{kn}$ 个文件段将通过宏基站的编码多播传输获得。通过优化可以得到内容放置矩阵 m 的具体值，形成一种内容放置策略。由于车辆节点的高速缓存大小， m 的值应该满足以下约束条件：

$$\sum_{n=1}^N m_{kn} s_0 \leq M_k, \forall k, \sum_{k=1}^K M_k \leq M$$

式中， M_k 表示每个车辆节点可以缓存的最大容量， M 表示总文件的大小。

S5，以最小化用户获取内容能量消耗为优化目标，优化网络边缘缓存模型，优化内容放置矩阵 m 的具体值，对车辆传输内容进行放置；其中，将流行文件集中的流行文件按照人气由高到低的顺序依次进行缓存。

首先分析内容缓存的能耗，定义为 E_c ，在放置阶段，每个文件都被部分缓存在车辆节点中，从而导致能源消耗。一般来说，缓存能耗与每个节点上存储的文件位数成正比。然后，只要已知内容放置矩阵 m ，就可以通过计算每个节点缓存的数据位数来获得缓存能耗，计算表示为：

$$E_C = e_{ca} \sum_{k=1}^K \sum_{n=1}^N m_{kn} s_0$$

式中， e_{ca} 为焦耳/比特的车辆节点缓存的能量系数， s_0 是每个文件段的大小。

其次分析车辆节点传输的能耗，即 E_t^V 。在传输阶段，当用户发起一个请求时，关联的节点将首先向用户传输其高速缓存的文件段。而其余的文件段由 MBS 传输给相关的车辆节点，然后关联的节点将其传输给用户。因此，从车辆节点传输文件的能耗包括两部分：一是向用户传输自己缓存文件段的能耗，另一部分是从 MBS 获得的未缓存文件段的能耗，总能耗计算为：

$$\begin{aligned} E_t^V &= e_t^v \sum_{k=1}^K \sum_{n=1}^N (1 - P_{kn}^0) (m_{kn} s_0 + (l_n - m_{kn}) s_0) \\ &= e_t^v \sum_{k=1}^K \sum_{n=1}^N (1 - e^{-U_k q_n t}) s_n \end{aligned}$$

式中， e_t^v 为车辆节点传输的每比特数据的能量消耗， P_{kn}^0 表示用户在第 k 个节点范围内请

求文件 n 次数为零的概率。

由于车辆节点的缓存大小的限制，不能从集群中获得部分文件段，这些文件段将通过回程链路从 MBS 传输到车辆节点。因此，在 MBS 上传输文件的能耗主要与在 MBS 上传输的文件位数成正比，MBS 的传输能耗可以表示为：

$$E_T^{MBS} = e_T^{MBS} B$$

式中， e_T^{MBS} 表示 MBS 传输的每比特数据的能量消耗， B 表示 MBS 通过编码组播传输的总比特数，表示为：

$$B = \sum_{n=1}^N s_n \sum_{j=1}^K \sum_{K_j \in K} \prod_{k \in K_j} (1 - P_{kn}^0) \prod_{k' \in K/K_j} P_{k'n}^0 \left(1 - \prod_{k \in K_j} \frac{m_{kn}}{l_n} \right)$$

式中， K_j 表示一个由 j , $1 \leq j \leq K$ 组成的车辆节点集， $K_j \subseteq K$ 。

假设每个车辆节点的缓存大小已知，将第 k 个车辆节点内容放置优化问题建模为：

$$\begin{aligned} \min_{m_k} & E_C + E_t^{MBS} + E_t^V \\ \text{s.t. C1:} & \sum_{n=1}^N m_{kn} s_0 \leq M / K, \forall k \\ \text{C2:} & 0 \leq m_{kn} \leq l_n, \forall k, n \end{aligned}$$

式中，限制条件 C1 和 C2 是对缓存容量的限制。

通过求出第 k 个车辆节点的内容放置，以此类推每个簇群中所有车辆节点的内容放置，得到内容放置矩阵 m 的具体值。

以上仅是本发明的优选实施方式，本发明的保护范围并不局限于上述实施例，凡属于本发明思路下的技术方案均属于本发明的保护范围。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前提下的若干改进和润饰，应视为本发明的保护范围。

权 利 要 求 书

1、一种分层网络体系结构中的协作缓存方法，其特征在于，所述协作缓存方法包括以下步骤：

S1，构建核心服务器、宏基站以及各车辆之间的网络边缘缓存模型，该网络边缘缓存模型设置在一个宏基站覆盖范围内，包含核心层和边缘层的两层典型的分层网络体系结构；其中，将该宏基站和其所包含的每个车辆均视为其中一个缓存节点，边缘层中的节点协作执行内容缓存；包括宏基站在内的核心层支持大带宽和快速的数据包传输的同时，缓存边缘层未缓存部分；

S2，将边缘层中的节点按区域划分为多个集群，每个集群节点数不超过阈值 β ；属于一个特定集群的所有节点通过将缓存信息更新到其集群头来协同执行缓存，节点阈值 β 与集群节点内的通信开销相关；

S3，将一组流行文件集通过拉链分布建模，寻找不同的用户对不同车辆节点中最流行的文件；

S4，通过使用编码缓存，把内容分割为分离的编码包，将第 n 个文件划分成 l_n 个文件段，且每个文件大小相等；同时将流行文件集中的文件按照受欢迎程度由高到低的顺序依次进行缓存，构建内容放置矩阵 m ；当用户 u_l 请求第 n 个文件时，其中 m_{kn} 个文件段从本地关联的簇中第 k 个车辆节点获得，剩余的 $l_n - m_{kn}$ 个文件段将通过宏基站的编码多播传输获得；

S5，以最小化总能耗为优化目标，进一步优化内容放置矩阵 m 的具体值，制定内容放置策略。

2、根据权利要求1所述的分层网络体系结构中的协作缓存方法，其特征在于，步骤S1中，边缘层中的节点协作执行内容缓存是指，只允许选定的车辆缓存其他车辆缓存事件的信息。

3、根据权利要求1所述的分层网络体系结构中的协作缓存方法，其特征在于，步骤S2中，将边缘层中的节点按区域划分为多个集群的过程包括以下步骤：

将整个边缘网络进行划分，同一片区域里的车辆节点分为一组，且每个集群节点数不超过阈值 β ，每组对应一个边缘簇，每个簇进一步分为各种节点类型：集群头节点、网关节点和集群成员；属于一个特定集群的所有节点通过将缓存信息更新到其集群头节点来协同执行缓存，集群成员之间不存在直接通信。

4、根据权利要求3所述的分层网络体系结构中的协作缓存方法，其特征在于，集群头节

点的选择过程包括以下步骤:

采用下述公式计算集群内每个节点的优先级评估值:

$$w_k = w_1 \frac{1/d_k}{\overline{M_d}} + w_2 \frac{T_k}{\overline{T}} + w_3 \frac{H_k}{\overline{H}} + w_4 \frac{p_k}{\overline{P}} + w_5 \frac{c_k}{\overline{C}}$$

式中, w 是不同参数对应的权重因子, $w_1 + w_2 + w_3 + w_4 + w_5 = 1$; d_k 是节点 k 范围内的节点数, $\overline{M_d}$ 是所有网络节点的节点度的平均逆; T_k 是节点 k 与其范围内的节点的传输时间, $\overline{T}$ 是所有网络节点的平均传输时间, H_k 是节点 k 与其范围内所有节点距离的加权和, $\overline{H}$ 是所有系统宽跳的平均加权和, p_k 是节点 k 的传输功率, $\overline{P}$ 是所有网络节点的平均功率利用率, c_k 是节点 k 的高速缓存利用率, $\overline{C}$ 是所有网络节点的平均缓存空间利用率。

5、根据权利要求 1 所述的分层网络体系结构中的协作缓存方法, 其特征在于, 步骤 S3 中, 将一组流行文件集通过拉链分布建模, 寻找不同的用户对不同车辆节点中最流行的文件的过程包括以下步骤:

S31, 根据下述公式计算一个用户请求第 k 个节点中的第 n 个文件的概率:

$$q_{kn} = \frac{(\gamma_{kn})^{-\alpha_k}}{\sum_{i=1}^N i^{-\alpha_k}}$$

式中, q_{kn} 表示第 k 个 SBS 的用户请求第 n 个文件的概率; α_k 是压缩指数, 用于描述请求模式中的偏态, α_k 越大意味着内容请求更集中在高排名的内容文件上; γ_{kn} 是第 k 个节点中第 n 个文件的排名顺序;

S32, 在每个传输周期 t 中, 第 k 个车辆节点将平均接收来自其覆盖区域内所有相关用户的 U_k 请求; 在第 k 个节点的覆盖范围内, 第 n 个文件的平均请求数定义为 ω_{kn} , 用 $\omega_{kn} = U_k q_{kn}$ 表示;

S33, 在多播周期 t 内, 在第 k 个节点覆盖范围内的所有用户请求文件共 j 次的概率为:

$$P_{kn}^j = \frac{(\omega_{kn})^j e^{-\omega_{kn}}}{j!}。$$

6、根据权利要求 1 所述的分层网络体系结构中的协作缓存方法, 其特征在于, 步骤 S4 中, 所述内容放置矩阵 m 为:

$$m = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1N} \\ m_{21} & m_{22} & \cdots & m_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1} & m_{K2} & \cdots & m_{KN} \end{bmatrix}$$

式中, m_{kn} 表示每个簇内第 k 个车辆节点缓存第 n 个文件的文件段的数量; m 的值满足以下约束条件:

$$\sum_{n=1}^N m_{kn} s_0 \leq M_k, \forall k, \sum_{k=1}^K M_k \leq M$$

式中, M_k 表示每个车辆节点可以缓存的最大容量, M 表示总文件的大小。

7、根据权利要求 6 所述的分层网络体系结构中的协作缓存方法, 其特征在于, 步骤 S5 中, 优化内容放置矩阵的建模为:

$$\begin{aligned} \min & E_C + E_t^{MBS} + E_t^V \\ \text{s.t. C1: } & \sum_{n=1}^N m_{kn} s_0 \leq M / K, \forall k \\ & \text{C2: } 0 \leq m_{kn} \leq l_n, \forall k, n \end{aligned}$$

式中, 限制条件 C1 和 C2 是对缓存容量的限制, s_0 是每个文件段的大小, M 是总文件的

大小; E_C 是内容缓存的能量消耗, $E_C = e_{ca} \sum_{k=1}^K \sum_{n=1}^N m_{kn} s_0$; E_t^V 是各车辆节点传输的能耗,

$E_t^V = e_t^v \sum_{k=1}^K \sum_{n=1}^N (1 - P_{kn}^0) (m_{kn} s_0 + (l_n - m_{kn}) s_0)$; P_{kn}^0 表示用户在第 k 个节点范围内请求文件 n 次数为

零的概率; E_t^{MBS} 是 MBS 的传输能耗, $E_t^{MBS} = e_T^{MBS} B$, e_T^{MBS} 表示 MBS 传输的每比特数据的能量消耗, B 表示 MBS 通过编码组播传输的总比特数; N 为流行文件集的文件总数。

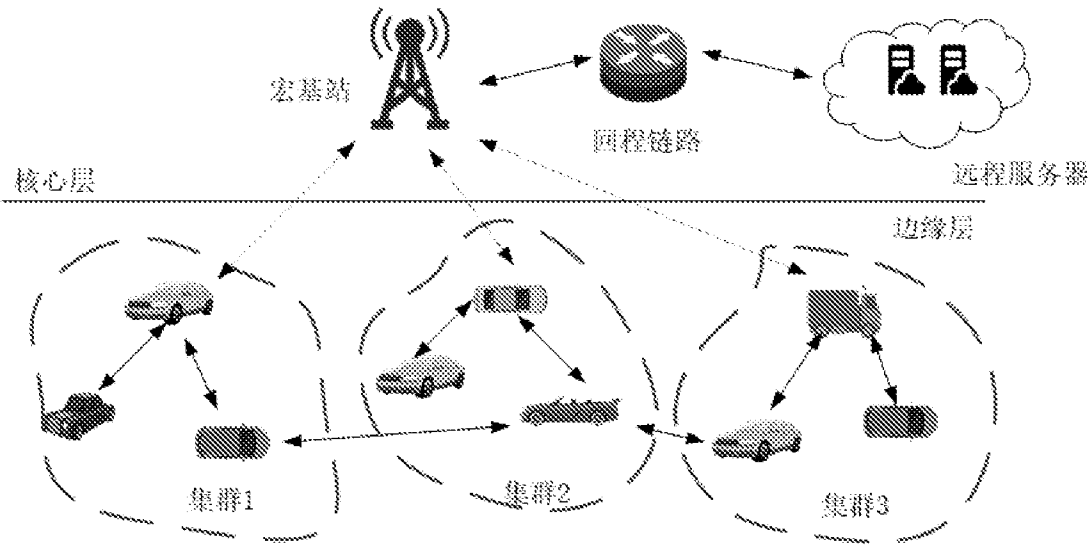


图 1

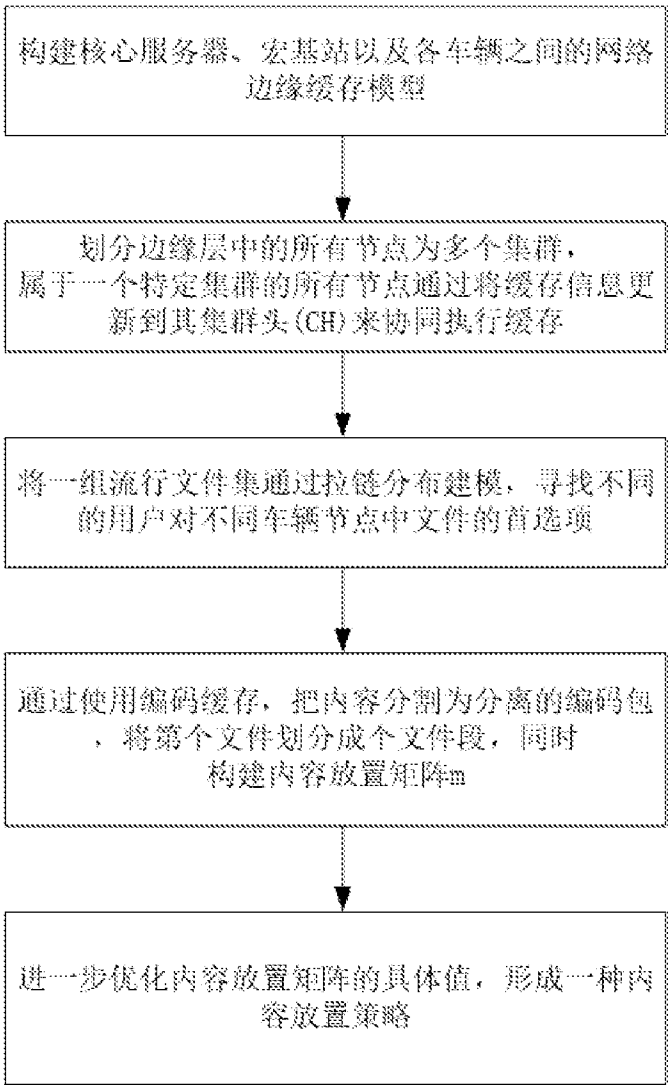


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/126959

A. CLASSIFICATION OF SUBJECT MATTER

H04W 28/14(2009.01)i; H04W 4/44(2018.01)i; H04W 4/46(2018.01)i; H04W 4/06(2009.01)i; H04L 67/568(2022.01)i;
H04L 67/5683(2022.01)i; H04W 40/32(2009.01)i; H04L 45/00(2022.01)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)

H0W; 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)

CNXTX; VEN; CNKI: 编码缓存, 分层, 协作, 边缘, 车辆, 集群, 多播, cooperat+, hierarchical, edge, cod+, cache, vehicle+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 114666843 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 24 June 2022 (2022-06-24) claims 1-7	1-7
X	CN 111741495 A (HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN (SHENZHEN INSTITUTE OF SCIENCE AND TECHNOLOGY INNOVATION, HARBIN INSTITUTE OF TECHNOLOGY)) 02 October 2020 (2020-10-02) description, paragraphs [0008]-[0105], and figure 1	1-7
A	CN 112839082 A (XIDIAN UNIVERSITY et al.) 25 May 2021 (2021-05-25) entire document	1-7
A	CN 113543071 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS et al.) 22 October 2021 (2021-10-22) entire document	1-7

☐ 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

01 December 2022

Date of mailing of the international search report

06 January 2023

Name and mailing address of the ISA/CN

**China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China**

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/126959

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	114666843	A	24 June 2022	None	
CN	111741495	A	02 October 2020	None	
CN	112839082	A	25 May 2021	None	
CN	113543071	A	22 October 2021	None	

国际检索报告

国际申请号

PCT/CN2022/126959

A. 主题的分类 H04W 28/14(2009.01)i; H04W 4/44(2018.01)i; H04W 4/46(2018.01)i; H04W 4/06(2009.01)i; H04L 67/568(2022.01)i; H04L 67/5683(2022.01)i; H04W 40/32(2009.01)i; H04L 45/00(2022.01)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) H0W; 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)) CNTXT;VEN;CNKI:编码缓存,分层,协作,边缘,车辆,集群,多播,cooperat+, hierarchical, edge, cod+, cache, vehicle+				
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 114666843 A (南京邮电大学) 2022年6月24日 (2022 - 06 - 24) 权利要求1-7	1-7		
X	CN 111741495 A (哈尔滨工业大学深圳哈尔滨工业大学深圳科技创新研究院) 2020年10月2日 (2020 - 10 - 02) 说明书第[0008]段至第[0105]段,图1	1-7		
A	CN 112839082 A (西安电子科技大学 等) 2021年5月25日 (2021 - 05 - 25) 全文	1-7		
A	CN 113543071 A (南京邮电大学 等) 2021年10月22日 (2021 - 10 - 22) 全文	1-7		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。				
<table><tr><td>* 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</td><td>“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</td></tr></table>			* 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
* 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 2022年12月1日		国际检索报告邮寄日期 2023年1月6日		
ISA/CN的名称和邮寄地址 中国国家知识产权局(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 马晔 电话号码 (86-512)88996063		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/126959

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	114666843	A	2022年6月24日	无	
CN	111741495	A	2020年10月2日	无	
CN	112839082	A	2021年5月25日	无	
CN	113543071	A	2021年10月22日	无	