















```

21      $NU(x) \leftarrow \emptyset;$ 
22     else
23         delete the replica of  $c_k$  carried by  $u_x$ ;
24          $d(x) \leftarrow d(x) - \{c_k\};$ 
25 return  $D_{imp}^D$ ;

```

以任意路边单元  $u_x \in DU$  为例,其冗余检查步骤如下:

检查  $NU(x)$  中是否存在尚未满足的请求:若存在,则表明该副本不冗余,保留该副本,并将其他路边单元的  $NU(y)$  中能够被  $u_x$  满足的请求删除;否则,即  $|NU(x)|=0$ ,表明  $u_x$  能够响应的请求已经被其他路边单元满足,则  $u_x$  中的  $c_k$  是冗余的,被清理。

清理冗余后得到更新的分布式存储方案,记为  $D_{imp}^D$ 。在图3实例中,执行冗余副本清理算法, $u_7$  中的  $c_1$  被清理,得到清理后的二部图如图4(c)所示。

在冗余副本清理过程中,可能产生新的可用存储单元,因此,控制中心执行第2轮分布式存储,旨在充分利用存储资源来满足更多的请求。具体地,以尚未满足的请求为左顶点,以空余存储单元为右顶点,构建新的初始二部图,再次执行前述二部图最大匹配以及冗余副本清理的步骤,直到二部图无边,控制中心得到最终的分布式存储方案,记为  $D$ 。在图3实例中,第2轮分布式存储将  $c_7$  分配到  $u_7$  的第2个存储单元,如图4(d)所示。

综上所述:在分布式存储机制 DSS 中,路边单元作为存储设备,周期性地从控制中心获取数据,并将副本发送给请求的车载节点。DSS 首先在请求分割规则的约束下,将原始问题转化成为二部图最大匹配问题,然后通过冗余副本清理算法解决规则造成的资源浪费,最终通过多轮数据分配,在多项式时间内计算得到一个响应率较高且资源利用充分的分布式存储方案。

## 4 实验结果及分析

### 4.1 网络环境配置

本文对 DSS 的实验验证在机会网络环境平台(opportunistic networking environment platform,简称 ONE<sup>[16]</sup>)上进行,ONE 被广泛认为是机会通信的标准模拟平台,虽然暂不支持诸如车流、变道等精确的交通模型,但 ONE 集成了多种典型的车辆移动模型和机会路由协议,这些组件简化了 DSS 中数据请求和响应过程的实现,因此,本文选择 ONE 平台进行模拟实验。

实验参数见表1。根据 IEEE 802.11p,车载节点和路边单元的通信半径是 200m,数据传输速率为 5Mbps。在基于地图的最短路径移动模型中,车辆从规定范围内随机选择速度,并以该速度沿着最短路径驶向目的地<sup>[17]</sup>。路边单元的存储空间大小和车载节点的请求数量对服务质量的影响将在第4.3节中讨论。

**Table 1** Simulation environment configurations

**表 1** 仿真环境配置

参数	值
网络范围	网格状 $15000 \times 15000 \text{m}^2$ , 每个网格大小 $500 \times 500 \text{m}^2$
路边单元数量, $N_u$	$30 \times 30$ , 部署在路段交叉口
车载节点数量, $N_v$	1 000
仿真时长	6 000s
存储周期, $\tau$	600s
通信半径	200m
数据传输速率	5Mbps
移动模型	基于地图的最短路径模型
速度范围	10~20m/s
数据数量, $N_c$	3 000
数据大小, $s$	10MB
路边单元存储空间大小	50MB
请求数目	在 $t$ 时间内一个车载节点最多请求 5 个数据
车载节点请求数据	请求发起时刻随机, 请求内容随机



由于经典的 Spray and Wait(S&W)方法和新近提出的 CCR 方法具有较好的分布式存储性能,本文选择这两种对比方法来测试 DSS 的效果.具体来讲,S&W 适用于机会网络,通过调节数据副本在移动节点中的存储来提升传输效率;而 CCR 是车载网络中容量限制的分布式存储机制,通过预测网络容量优化数据存储.由于车载网络是机会网络的典型应用,S&W 和 CCR 均适用于以车载节点为存储设备的车载环境.此外,考虑到针对部署有路边单元的车载网络存储机制的研究较少,选择时延优先的贪心算法作为二部图匹配算法的对比,称为 Greedy.在 Greedy 算法中,控制中心将车载节点  $v_i$  所请求的数据分配给它在下一个存储周期内将遇到的第 1 个路边单元( $g(i)$ 中的第 1 个元素),若该路边单元的存储空间已满,则分配给  $g(i)$ 中的下一个路边单元.由于 Greedy 算法先将数据分配给较早相遇的路边单元,因此有助于缩短响应时延.

实验评估参照 4 个标准,即:数据响应率、平均响应时延、副本数量和传输开销.数据响应率表示被响应的数据请求在所有请求中所占的比率,响应率越高,则存储机制的性能越好.平均响应时延即获得响应的请求的响应时延的平均值,一个请求的响应时延是指从请求被发出到被响应之间的时间跨度,平均响应时延短则表示该机制适用于对时延有严格要求的车载应用.副本数量直接体现了整个网络中存储资源的消耗,而传输开销则表示在实验时间内传输的数据包个数,反映了网络通信资源的消耗.

### 4.2 实验结果及分析

在表 1 所示的实验参数下,各个分布式存储方案的性能如图 5 所示.

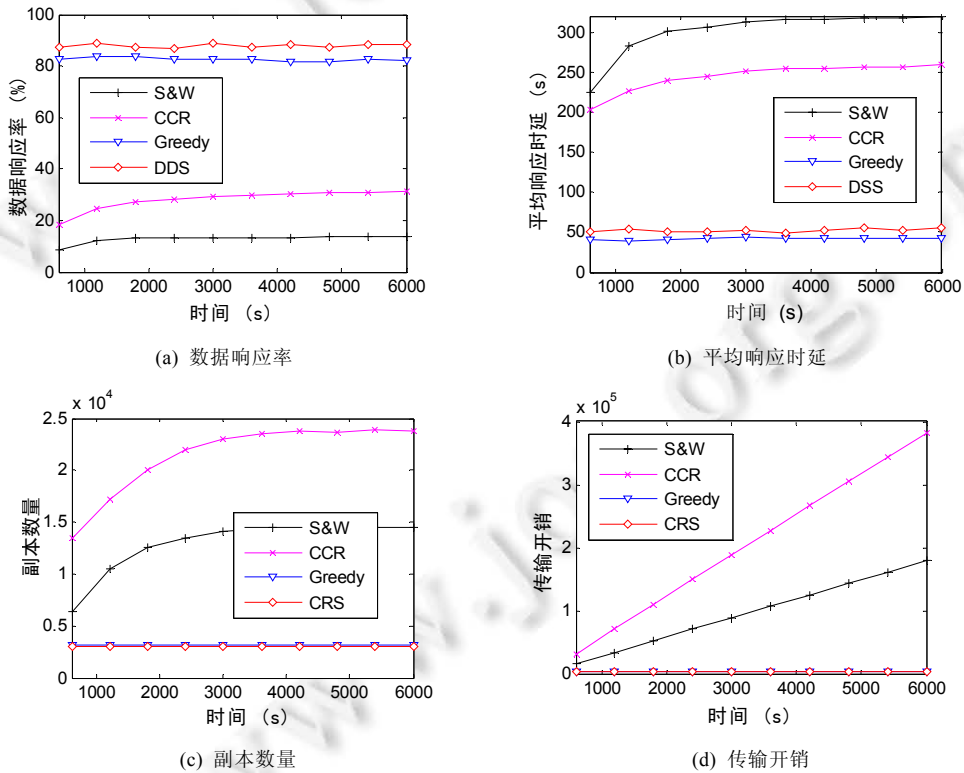


Fig.5 Experimental results

图 5 实验结果

在图 5(a)中,CCR 因为支持副本个数的动态调整,其响应率约是 S&W 的两倍.尽管如此,S&W 和 CCR 的数据响应率明显低于其他两种方案,这是由于 S&W 和 CCR 以车载节点作为存储设备,而车载节点之间的相遇时间短且相遇概率小,极大地限制了数据的传输.DSS 的响应率约为 88%,高于 Greedy.DSS 使用二部图最大匹配算

法来提升满足的请求数量,实验结果表明,DSS 的数据存储比 Greedy 更高效。

如图 5(b)所示:由于 Greedy 优先将数据分配给较早遇到的路边单元,故具有最短的平均响应时延,DSS 和 Greedy 的时延差距只有 10s,与 600s 的存储周期相比,DSS 和 Greedy 的时延差距几乎可以忽略不计。而 S&W 和 CCR 由于使用车载节点存储数据,时延远远大于其他两种方案。

图 5(c)和图 5(d)表明:S&W 和 CCR 的副本数量和传输开销远远高于 DSS 和 Greedy,其中,CCR 的副本数量最大。这是由于 S&W 中副本数量是不变的,而 CCR 会根据预测的网络容量来决定副本数量。此外,与 Greedy 相比,DSS 减少约 3%的副本数量,而两者的传输开销基本相同。

在本文提出的 DSS 机制中:

- 一方面,基于二部图最大匹配的分布式存储方法的计算复杂度与二部图节点数和边数相关,其时间复杂度为  $O(L \cdot |E|)$ 。由于  $L$  表示车辆的数据请求集合,假设在一个存储周期内任意车辆请求不多于  $N_{rq}$  个不同数据,则左顶点个数不大于  $N_{rq} \cdot N_v$ ;而边集  $E$  由车辆与路边单元的相遇情况决定,假设在一个存储周期内任意车辆最多遇到  $N_{uc}$  个不同路边单元,记路边单元的存储单元个数的最大值为  $N_{us}$ ,即  $N_{us} = \text{MAX}(w(j))$ ,则边数不大于  $N_{rq} \cdot N_v \cdot N_{uc} \cdot N_{us}$ 。综合可得,该方法的时间复杂度表示为

$$O(N_{rq}^2 \cdot N_v^2 \cdot N_{uc} \cdot N_{us});$$

- 另一方面,由算法 1 可知,冗余副本清理算法的时间复杂度为  $O(N_c \cdot N_u^2 \cdot N_v)$ 。

综上所述,DSS 机制实现了在多项式时间内求解分布式存储问题。

整体来讲,与 CCR 机制相比,DSS 在大幅提高数据响应率的同时,将平均响应时延缩短至大约 20%;与贪心算法 Greedy 相比,DSS 的数据响应率提高了约 13%,而平均响应时延只是略有延长。总而言之,DSS 有效地提高了数据响应率,同时保持了较短的响应时延。

### 4.3 参数分析

本节讨论路边单元的存储空间大小和车辆的请求个数对存储机制性能的影响,为节省空间,此处仅分析数据响应率和平均响应时延这两项核心指标。

#### 4.3.1 存储空间大小的影响分析

令路边单元的存储空间从 10MB~100MB 递增,考虑到 S&W 和 CCR 不使用路边单元,此处只比较 Greedy 和 DSS 的实验结果。如图 6 所示:随着存储空间扩大,Greedy 和 DSS 的响应率都先有所提高,然后平缓下降,同时响应时延不断缩短。主要原因是:(1) 存储资源匮乏限制了分布式存储方案 DSS 的优势发挥;(2) 当路边单元有足够的存储空间,就能存储更多的数据副本来满足请求并缩短响应时延;(3) 若路边单元的存储空间过大,则导致 V2U 通信时的带宽竞争加剧,最终使得响应率降低。

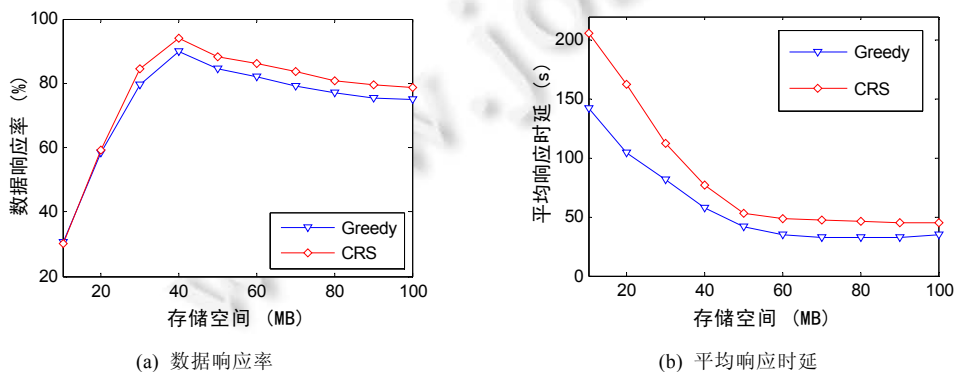


Fig.6 Results with different storage sizes of a roadside unit

图 6 路边单元存储空间大小的影响

### 4.3.2 请求个数的影响分析

令每个车载节点在一个存储周期内发送的请求个数从 1~10 依次递增,实验结果如图 7 所示.当请求不太多时,各个存储机制的性能基本保持稳定;在请求增加到一定程度后,网络的服务质量逐渐下降.车辆发送的请求越多,系统就需要越多资源来满足请求.因此,当请求个数适当时,整体性能保持不变.然而,当请求个数达到网络的饱和值,资源匮乏将极大地限制车载网络的整体性能.

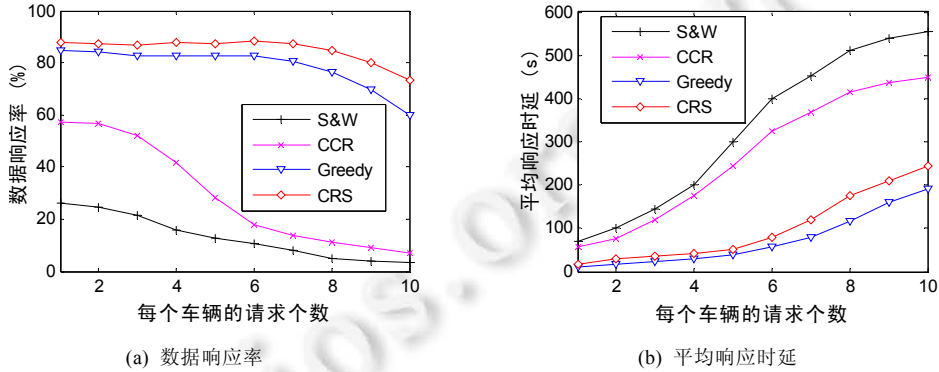


Fig.7 Results with different numbers of requests

图 7 请求个数的影响

综上所述,DSS 在不同的存储空间大小和不同的请求个数的条件下,都具有较高的响应率和较低的响应时延,进一步验证了其良好的可扩展性和通用性.

## 5 总结

本文提出了基于二部图匹配的分布式存储机制,以路边单元作为存储设备,旨在充分利用网络资源,有效地提高智慧交通应用中的数据接收率.首先设定请求分割规则,根据车载节点的数据请求、路边单元的存储空间大小以及车辆与路边单元的预期相遇情况,将初始的分布式存储问题转化成为二部图最大匹配问题,在多项式时间内计算得到初始分配方案;然后设计冗余副本清理算法,针对不同路边单元存储同一数据时可能产生的冗余副本问题,依据响应因子对各副本排序,逐个进行冗余检查和清理,从而提高存储资源利用率.实验结果表明:与传统的车载网络存储机制相比,DSS 在有限的存储资源下,显著提高了数据响应率并保持较短的响应时延.

尽管如此,路边单元的无线信道竞争仍然是 DSS 实施的瓶颈,解决无线通信资源的调度问题将进一步提高整体性能<sup>[18]</sup>.此外,车载节点之间的协同下载<sup>[19]</sup>有助于改善数据传输,将车辆协同下载技术与路边单元的分布式存储机制相结合,是未来的研究方向.

**致谢** 在此,我们向对本文的工作给予支持和建议的同行,尤其是北京航空航天大学计算机学院熊璋教授、蒲菊华副教授领导的讨论班上的同学和老师表示感谢.

### References:

- [1] Piao J, McDonald M, Hounsell N. Cooperative vehicle in frastructure systems for improving driver information services: An analysis of COOPERS test results. *IET Intelligent Transport Systems*, 2012,6(1):9-17. [doi: 10.1049/iet-its.2010.0169]
- [2] Mershad K, Artail H. A framework for secure and efficient data acquisition in vehicular ad hoc networks. *IEEE Trans. on Vehicular Technology*, 2013,62(2):536-551. [doi: 10.1109/TVT.2012.2226613]
- [3] Wang SG, Lei T, Zhang LY, Hsu CH, Yang FC. Offloading mobile data traffic for QoS-aware service provision in vehicular cyber-physical systems. *Future Generation Computer Systems*, 2016,61:118-127. [doi: 10.1145/2491288.2491312]

- [4] Kenney JB. Dedicated short-range communications (DSRC) standards in the United States. Proc. of the IEEE, 2011,99(7): 1162–1182. [doi: 10.1109/JPROC.2011.2132790]
- [5] Saad W, Han Z, Hjorungnes A, Niyato D, Hossain E. Coalition formation games for distributed cooperation among roadside units in vehicular networks. IEEE Journal on Selected Areas in Communications, 2011,29(1):48–60. [doi: 10.1109/JSAC.2011.110106]
- [6] La C, Michiardi P, Casetti C, Chiasserini C, Fiore M. Content replication in mobile networks. IEEE Journal of Selected Areas in Communications, 2012,30(9):1762–1770. [doi: 10.1109/JSAC.2012.121021]
- [7] Vahdat A, Becker D. Epidemic routing for partially connected adhoc networks. Technical Report, CS-200006, Durham: Duke University, 2000.
- [8] Spyropoulos T, Psounis K, Raghavendra C. Spray and wait: An efficient routing scheme for intermittently connected mobile networks. In: Proc. of the ACM SIGCOMM Workshop on Delay-Tolerant Networking (WDTN). Philadelphia: ACM Press, 2005. [doi: 10.1145/1080139.1080143]
- [9] Jaho E, Koukoutsidis I, Stavrakakis I, Jaho I. Cooperative content replication in networks with autonomous nodes. Computer Communications, 2012,35(5):637–647. [doi: 10.1016/j.comcom.2011.12.006]
- [10] Cavaglione L, Cervellera C. An optimized content replication and distribution framework for vehicular networks. Journal of Intelligent Transportation Systems, 2011,15(4):179–192. [doi: 10.1080/15472450.2011.620474]
- [11] Kapadia S, Krishnamachari B, Ghandeharizadeh S. Static replication strategies for content availability in vehicular ad-hoc networks. Mobile Network Applications, 2009,14(5):590–610. [doi: 10.1007/s11036-008-0120-y]
- [12] Kumar N, Kim J. Probabilistic trust aware data replica placement strategy for online video streaming applications in vehicular delay tolerant networks. Mathematical and Computer Modelling, 2013,58(1):3–14. [doi: 10.1016/j.mcm.2012.07.021]
- [13] Zhuo X, Li Q, Gao W, Cao G, Dai Y. Contact duration aware data replication in delay tolerant networks. In: Proc. of the IEEE Int'l Conf. on Network Protocols (ICNP). Vancouver: IEEE Press, 2011. 236–245. [doi: 10.1109/ICNP.2011.6089057]
- [14] Wu Y, Zhu Y, Zhu H, Li B. CCR: Capacity-Constrained replication for data delivery in vehicular networks. In: Proc. of the IEEE Int'l Conf. on Computer Communications (INFOCOM). Turin: IEEE Press, 2013. 2580–2588. [doi: 10.1109/INFOCOM.2013.6567065]
- [15] Rezgui J, Cherkaoui S. About deterministic and non-deterministic vehicular communications over DSRC/802.11p. Wireless Communications & Mobile Computing, 2014,14(15):1435–1449. [doi: 10.1002/wcm.2270]
- [16] Keranen A, Ott J, Karkkainen T. The ONE simulator for DTN protocol evaluation. In: Proc. of the Int'l Conf. on Simulation Tools and Techniques (SIMUTools). Rome: ACM Press, 2009. [doi: 10.4108/ICST.SIMUTOOLS2009.5674]
- [17] Harri J, Filali F, Bonnet C. Mobility models for vehicular ad hoc networks: A survey and taxonomy. IEEE Communications Surveys & Tutorials, 2009,11(4):19–41. [doi: 10.1109/SURV.2009.090403]
- [18] Malandrino F, Casetti C, Chiasserini C, Fiore M. Content download in vehicular networks in presence of noisy mobility prediction. IEEE Trans. on Mobile Computing, 2014,13(5):1007–1021. [doi: 10.1109/TMC.2013.128]
- [19] Trullols-Cruces O, Fiore M, Barcelo-Ordinas J. Cooperative download in vehicular environments. IEEE Trans. on Mobile Computing, 2012,11(4):663–678. [doi: 10.1109/TMC.2011.100]



唐晓岚(1987—),女,河北邢台人,博士,讲师,CCF 会员,主要研究领域为车载网络,无线传感器网络,城市计算.



陈文龙(1976—),男,博士,副教授,CCF 会员,主要研究领域为互联网体系结构,物联网,高性能路由系统.



洪东惠(1993—),女,硕士生,主要研究领域为车辆自组织网络.



蒲菊华(1976—),女,博士,副教授,CCF 会员,主要研究领域为车辆自组织网络,数据挖掘,智慧城市,城市计算.