

从图 4 和图 5 中可以看出,相比于 FIR 法和 WT 法,本文方法能够更有效地去除强干扰,从而显著提高接收机的捕获概率.如在图 4 中,在 $P_{fa} = 10^{-6}$ 以及 GPS 信噪比(signal to noise ratio, SNR)为 -18 dB 的情况下,对于干噪比(jamming to noise ratio, JNR)为 6 dB 的 BPSK 干扰,利用 FIR 及 WT 方法滤波后基本上很难捕获到信号,但用本文方法滤波后, GPS 信号的捕获概率接近 100% ,在图 5 中同样可得到类似的结果.从信噪比来看,在 GPS 信号捕获概率为 80% 的情况下,对于图 4 中 6 dB 的 BPSK 干扰,本文方法要分别优于 FIR 和 WT 方法约 4 dB 和 6 dB;而对于图 5 中 12 dB 的 BPSK 干扰,本文方法则分别要优于 FIR 和 WT 方法约 5 dB 和 6 dB.

此外,从图 4 和图 5 中还可以看出,对于 ASK 干扰,3 种方法滤波后的 GPS 信号捕获性能相近;而对于 BPSK 和 FSK 干扰,本文方法滤波后的 GPS 信号捕获性能则明显要好于 FIR 和 WT 方法.

(II) 干扰去除后的 GPS 信号跟踪性能仿真

GPS 信号在捕获成功后,会进入跟踪阶段以获取伪距及解调导航星历数据等.由于信号跟踪测距误差主要取决于码环(delay locked loop, DLL)^[18],因此在图 6 与图 7 中针对表 1 所给的不同类型干扰,在分别采用本文方法、FIR 方法和 WT 方法的情况下,对滤波后的 GPS 信号码环跟踪性能进行了仿真.其中, d_{EL} 为码环早和迟相关器间距,在图 6 和图 7 中, d_{EL} 均取为 0.1 码片;而码环的环路跟踪误差则采用早相关与迟相关非相干幅值相减归一化型鉴别器^[18]进行测量.

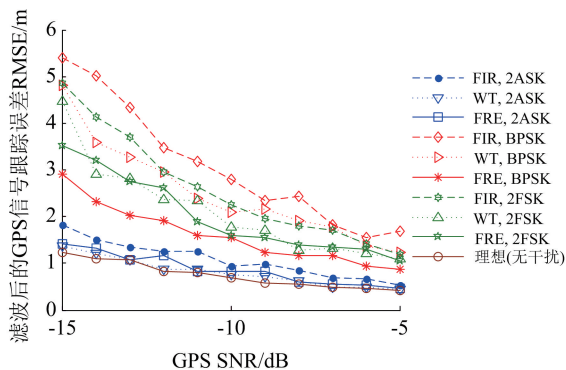


图 6 本文方法与 FIR 方法及 WT 方法去除 6 dB 不同干扰后的 GPS 信号跟踪精度比较

Fig.6 Tracking accuracy comparisons of the separated GPS signal after the mitigation of different strong interference with different methods (JNR=6 dB)

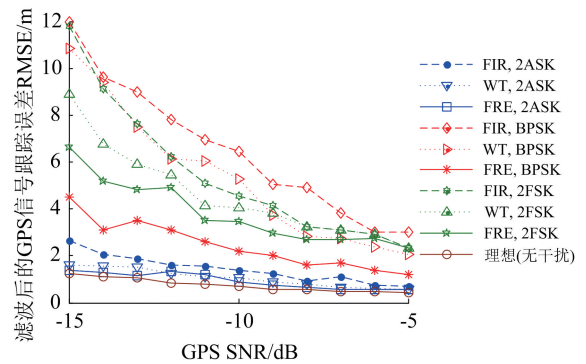


图 7 本文方法与 FIR 方法及 WT 方法去除 12 dB 不同干扰后的 GPS 信号跟踪精度比较

Fig.7 Tracking accuracy comparisons of the separated GPS signal after the mitigation of different strong interference with different methods (JNR=12 dB)

从图 6 和图 7 可以看出,经本文方法滤波后, GPS 信号的跟踪性能也会得到显著提高.在跟踪精度上,本文方法同样要优于 FIR 和 WT 方法.如在图 6 中,对于 6 dB 的 BPSK 强干扰,在 GPS 信噪比为 -13 dB 时,经过 FIR 方法滤波后,其跟踪误差 RMSE(root mean square error)约为 4.5 m;经过 WT 方法滤波后,其跟踪误差 RMSE 约为 3.5 m;而用本文方法滤波后, GPS 信号的跟踪误差则降为 2 m,跟踪精度得到了显著提高.对于其他情况,也可得到类似的结果.同时,从图 6 和图 7 可以看出,随着 GPS 信噪比的提高,3 种方法滤波后的 GPS 信号跟踪精度都会提高,并且 3 种方法跟踪性能的差距也逐渐变小,但在 GPS 信号信噪比较低的情况下,本文方法滤波后的 GPS 信号跟踪精度要好于 FIR 和 WT 方法.

此外,图 6 和图 7 同样表明,对于 ASK 干扰,3 种方法滤波后的 GPS 信号跟踪精度性能相近;而对于 BPSK 和 FSK 干扰,在 GPS 信号信噪比较低时,本文方法滤波后的 GPS 信号跟踪精度性能则要优于 FIR 和 WT 方法.

(III) 存在多种干扰信号时滤波后的 GPS 信号跟踪捕获性能仿真

为进一步考察所给方法的性能,对存在多种干扰时,利用 3 种方法滤波后的 GPS 信号捕获与跟踪性能也进行了仿真验证,结果如图 8 和图 9 所示.其中 BPSK、2ASK 与 2FSK 干扰的 JNR 分别设为 10 dB、 7 dB 和 5 dB;而在图 9 中, d_{EL} 同样取为 0.1 码片.

从图 8 和 9 可以看出,此时 GPS 信号的跟踪捕获性能相对于单干扰时的性能都出现了下降,但本

文方法滤波后的性能仍要好于 FIR 方法;相较于 WT 方法,本文方法性能则持平或略差于 WT 方法,这主要是由于此时本文方法对多干扰循环频率识别能力变差造成的.如在图 9 中,当同时存在 BPSK 和 2FSK 干扰时,在 GPS 信噪比为 -13 dB 的情况下,本文方法滤波后的 GPS 信号跟踪精度约为 7.5 m,而对于 FIR 和 WT 方法滤波后的 GPS 信号跟踪精度则分别为 16 m 和 7 m,这也表明本文方法对于多干扰的情况仍需研究与改进.

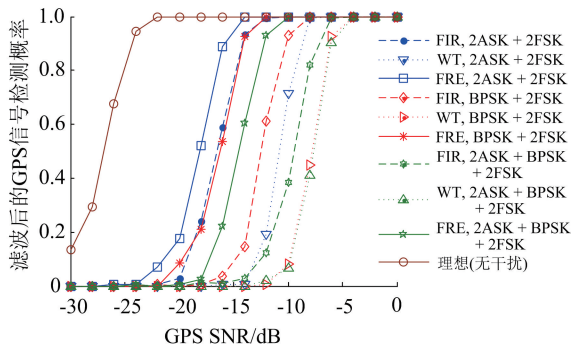


图 8 存在多种干扰时所给方法与 FIR 方法及 WT 方法去除不同干扰后的 GPS 信号捕获概率比较 ($P_{fa} = 10^{-6}$)

Fig.8 Acquisition probability comparisons of the separated GPS signal after the mitigation of multiple strong interference with different methods and the given JNRs ($P_{fa} = 10^{-6}$)

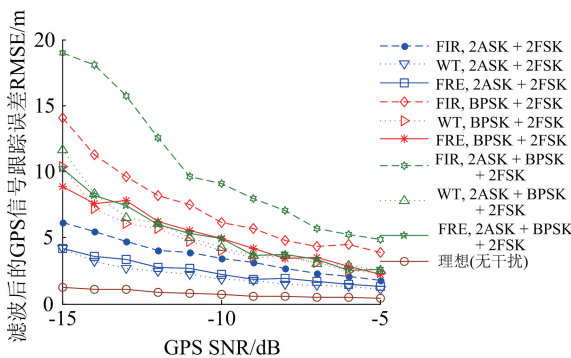


图 9 存在多种干扰时所给方法与 FIR 方法及 WT 方法去除不同干扰后的 GPS 信号跟踪精度比较

Fig.9 Tracking accuracy comparisons of the separated GPS signal after the mitigation of multiple strong interference with different methods and the given JNRs

5 结论

为消除 GPS 信号接收中所受到的时频重叠强干扰,本文给出了一种利用循环谱分析并结合频移滤波来消除强干扰的方法.通过对不同干扰场景下利用本文方法去除干扰后的 GPS 信号跟踪捕获性能的考察以及与 FIR 方法和 WT 方法滤波后性能

的比较,验证了本文方法的有效性与优势.此外,由于在干扰去除中没有涉及具体的弱卫星导航信号的循环谱分析,因此本文方法同样适用其他卫星导航系统信号抗时频重叠干扰.当然,由于干扰信号结构及类型的复杂性,本文方法对于一些更复杂的干扰如多干扰的情况等,仍有待进一步研究与考察.

参考文献(References)

[1] Navstar GPS space segment/navigation user interface. IS-GPS-200H-003 [S]. Navstar GPS Joint Program Office, December 2015; 16-21.

[2] IOANNIDES R T, PANY T, GIBBONS G. Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques [J]. Proceedings of the IEEE, 2016, 104(6): 1174-1194.

[3] RAASAKKA J, OREJAS M. Analysis of notch filtering methods for narrowband interference mitigation [C]//2014 Position, Location and Navigation Symposium. Monterey, USA: IEEE/ION Press, 2014: 1282-1292.

[4] AMIN M G, WANG X, ZHANG Y D, et al. Sparse arrays and sampling for interference mitigation and DOA estimation in GNSS [J]. Proceedings of the IEEE, 2016, 104(6): 1302-1317.

[5] AMIN M G, BORIO D, ZHANG Y D, et al. Time-frequency analysis for GNSSs: from interference mitigation to system monitoring [J]. IEEE Signal Processing Magazine, 2017, 34(5): 85-95.

[6] AVASTA S, PRESTI L L, RAO M. Interference mitigation in GNSS receivers by a time-frequency approach [J]. IEEE Transactions on Aerospace and Electronic Systems, 2013, 49(1): 415-438.

[7] MUSUMECI L, DOVIS F, CURRAN J T. A comparative analysis of adaptive notch filtering and wavelet mitigation against jammers interference [J]. Navigation, 2016, 63(4): 533-550.

[8] SHI B B, CHENG Z, QIAN L J, et al. Cyclostationarity based cascaded space-time anti jamming processor designed to appeal to GPS receivers [J]. Journal of Xidian University, 2010, 37 (4): 743-750.

[9] GARDNER W A. Cyclic Wiener filtering: theory and method [J]. IEEE Transactions on Communications, 1993, 41(1): 151-163.

[10] ZHANG J, WONG K M, LUO Z Q, et al. Blind adaptive FRESH filtering for signal extraction [J]. IEEE Transactions on Signal Processing, 1999, 47(5): 1397-1402.

(下转第 754 页)

FCD 大数据并行处理的动态任务调度算法

陈 锋^{1,2}, 张 智¹, 李琴剑¹, 陈宇强², 陈国良³

(1.中国科学技术大学信息学院,合肥 230027;2.安徽中科龙安科技股份有限公司,安徽合肥 230088;
3.中国科学技术大学计算机学院,安徽合肥 230027)

摘要: 浮动车数据(floating car data, FCD)技术是大规模城市路网交通流实时采集的有效方法.城市交通的动态诱导和控制需要对海量 FCD 进行快速处理.鉴于此,提出了 FCD 并行计算的动态任务调度方法.针对 FCD 数据包计算时间的不确定性和动态性,根据计算节点的处理能力进行数据包的动态分割,在处理过程中,采用动态任务分配策略以实现计算节点的同步.该方法在龙芯国产大数据一体机平台上进行了实现,并采用现场 FCD 数据进行了实验验证,结果表明,该方法较轮询和 Min-Min 调度算法,显著地提高了并行处理的性能.

关键词: 浮动车数据;大数据;并行计算;动态任务划分;动态任务调度

中图分类号: TP391 **文献标识码:** A **doi:** 10.3969/j.issn.0253-2778.2018.09.005

引用格式: 陈锋,张智,李琴剑,等. FCD 大数据并行处理的动态任务调度算法[J]. 中国科学技术大学学报,2018,48(9):718-722.

CHEN Feng, ZHANG Zhi, LI Qinjian, et al. Dynamic task scheduling algorithm of parallel computing for FCD big data[J]. Journal of University of Science and Technology of China, 2018,48(9):718-722.

Dynamic task scheduling algorithm of parallel computing for FCD big data

CHEN Feng¹, ZHANG Zhi¹, LI Qinjian¹, CHEN Yuqiang², CHEN Guoliang¹

(1. Department of Automation, University of Science and Technology of China, Hefei 230027, China;

2. Anhui LoongSon Science and Technology Co., Ltd, Hefei 230088, China;

3. School of Computer Science and technology, University of Science and Technology of China, Hefei 230027, China)

Abstract: FCD (floating car data) technique is new way of collecting real-time traffic flow from large-scale urban networks. It is necessary to implement rapid processing of FCD big data for the dynamic guidance and control of urban traffic. A dynamic task scheduling algorithm is proposed for parallel computation of FCD. To address the uncertainty and dynamics of FCD package processing, FCD packages are partitioned dynamically. The load balance among computing nodes can be achieved using the dynamic task allocation strategy. The algorithm is developed on LoongSon big data integrated machine platform and evaluated using field FCD. The experimental results indicate that the proposed algorithm has significantly higher parallel processing performances compared to the polling scheduling algorithm and Min-Min scheduling algorithm.

Key words: floating car data; big data; parallel computing; dynamic task partition; dynamic task scheduling

收稿日期: 2018-03-27; 修回日期: 2018-04-27

基金项目: 国家重点研发(973)计划(2017YFC0840206),安徽省重大科技专项(17030901007)资助.

作者简介: 陈锋(通讯作者),男,1966年生,博士/副教授.研究方向: 研究神经网络与模式识别. E-mail: chenfeng@ustc.edu.cn

0 引言

大范围交通流采集是实现城市交通诱导和主动控制的前提.传统的交通流采集方法如地感线圈、地磁传感器、微波和视频等,是固定型的,仅提供道路的截面数据,无法反映路网的交通流状况.浮动车数据技术是大范围交通流采集的有效手段^[1-2].该技术利用一定数量的、在城市路网上运行的配备车载 GPS 装置的车辆,形成移动车辆传感网络.各浮动车以一定的时间间隔,将其经/纬度坐标、时间、速度和方向等信息传输至远程监控中心.通过对 FCD 的处理能够获取大范围城市路网的实时交通状况,为城市交通的诱导、控制、管理和规划提供科学的决策支持^[3-4].目前 FCD 的研究主要集中在交通状态估计、出行诱导、旅行时间预测和交通事件检测等方面^[5-9].

为了较准确地对城市路网交通流进行建模,需要占路网运行车辆总数一定比例的车辆作为浮动车进行数据采集.一般地,浮动车数量约为机动车总数的 5%.考虑到通信代价,浮动车的采样间隔为 15 秒左右,其传输至远程中心的时间间隔约为 3~5 分钟,需要处理的 FCD 的数量巨大.以 1 万辆出租车作为浮动车为例,每小时的 FCD 数据包大小约为 0.15 G 字节.

城市交通的管控和服务需要对 FCD 大数据进行快速处理和分析,相关的研究包括:Deng 等基于 Hadoop-GIS 方法,给出了浮动车数据处理框架^[10].Hadoop-GIS 由两个主要工具组成:Hadoop 的空间框架和地理处理工具,使用空间查询语言 HiveQL 和 Hive 中的类 SQL 语言,将查询透明转换至 MapReduce,从而实现大规模 FCD 的有效处理.为了实现城市交通优化控制,Zhang 等研究了在云计算环境下的浮动车数据处理^[11],通过实验验证了 Hbase 和 MapReduce 能够提供大数据流的实时计算.上述方法由于没有考虑不同 FCD 数据的处理时间差异,难以充分发挥 Hadoop 并行处理机制的效率.

为了降低 FCD 处理延迟,Lu 等设计了一种基于多服务器的 FCD 分布式处理系统^[12],以实现旅行时间计算,根据 FCD 的位置信息进行原始 FCD 数据的分组与分类处理,以降低服务器从外存中装载或卸载区域地图的时间.其存在的问题包括:地图被划分为一定数量的网格,根据网格的经纬度范围

静态分配 FCD 数据,以实现负载均衡.由于浮动车的位置可能分布在不同的网格,因此增加了通信开销;文献[13]侧重于研究根据历史 FCD 在异构的计算机平台上进行道路状况建模,采用的任务调度策略为计算机数量确定的情况下,按 CPU 的计算能力分配任务,实现负载均衡,该方法没有考虑 FCD 数据的实时性要求.上述方法属于任务的静态分配模式,没有对 FCD 数据的差异进行分析,而相同数量的 FCD 数据包由于噪声的存在导致其处理时间不同.另外,基于 Internet 的并行计算通信延迟难以满足实际应用的实时性要求.

任务调度策略是并行处理的关键技术,典型的任务调度算法包括轮询调度算法、Min-Min 算法等.轮询调度算法直接将待执行的任务按照任务流顺序依次下发给各节点,空闲节点接收到任务后开始执行;否则放入节点的任务队列中,等待节点空闲时调度.该算法的优点是实现简单,服务器开销很小,在任务量不多的情况下系统效率较好.Min-Min 算法将预期完成时间最早的任务映射到性能最好的服务器上执行^[14-15],理论上任务完成效率较高.轮询和 Min-Min 调度算法在处理简单任务时效率较好,但在处理复杂任务时会导致服务器节点间负载均衡,降低并行处理效率^[16].

本文对 FCD 的特性进行了深入研究,设计了 FCD 并行处理机制,提出了 FCD 并行计算的动态任务划分与调度方法.针对 FCD 数据包计算时间的不确定性和动态性,根据计算节点的处理能力进行数据包的动态分割.在处理过程中,采用动态分配策略以实现计算节点的同步和负载均衡.本文所提方法在龙芯国产大数据一体机平台上进行了实现,并采用现场 FCD 数据对本文方法进行了实验测试,实验结果表明,本文方法显著地提高了并行处理的性能.

1 FCD 并行处理

1.1 FCD 的特性分析

在浮动车形成的无线传感网络中,每个装载 GPS 的浮动车按一定采样间隔对其所在的路段的交通状况进行感知,产生 FCD 记录.一个 FCD 记录主要包括:车载终端 ID、经纬度坐标、瞬时速度、方向、时间、载客状态等字段.各浮动车以一定的时间间隔向远程监控中心传送 FCD 数据包.监控中心对浮动车上传的 FCD 进行地图匹配、速度计算,根据

交通流模型,对路段交通流密度、平均速度和行程时间进行估计,从而获取城市路网的实时交通状态,为驾驶员提供优化的行驶路径和行程信息,也为交通管理者提供了有效的分析数据。

我们分析了龙芯 3B2000CPU 的指令系统,根据 WCET(worst-case execution time)原则研究了 FCD 对处理时间的要求,对计算节点 CPU 的 FCD 处理所涉及的各种基本操作以及函数的处理时间进行统计,发现不同的 FCD 记录之间处理时间最大差异约为 3.5 倍,其原因如下:

(I)动态采样间隔.由于噪声干扰,本质上 FCD 的采样间隔是不确定的;

(II)FCD 记录处理时间的不确定性.由于各浮动车的速度不同和采样间隔的动态性,为计算路段平均车速而进行的路径搜索范围不同,导致不同的 FCD 记录处理时间有较大的差异。

1.2 FCD 并行处理体系结构

在我国,一个大型城市路网规模为数千个交叉口、几万个路段,获取的 FCD 数据是海量的.为满足交通动态诱导和控制的实时性要求,本文以龙芯国产大数据一体机为平台,设计了 FCD 大数据的并行处理结构,如图 1 所示。

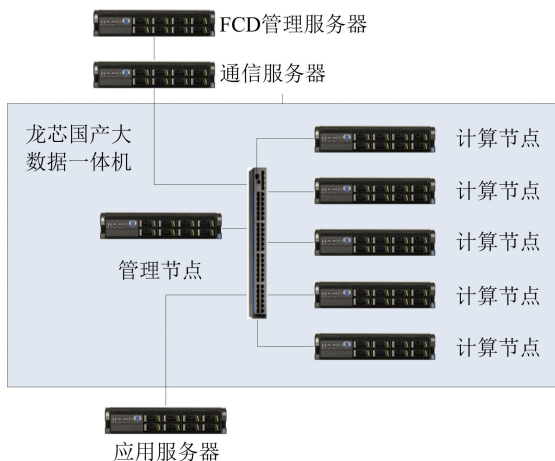


图 1 基于龙芯国产大数据一体机平台的 FCD 并行计算结构

Fig.1 FCD parallel computing architecture based on LoongSon big data integrated machine platform

本系统由 FCD 采集服务器、通信服务器、龙芯国产大数据一体机、应用服务器等组成,龙芯国产大数据一体机由一个管理节点、五个计算节点和万兆交换机组成,计算速度可达 1.5 万亿次/秒,并可扩展至 15 万亿次/秒.其中,FCD 采集服务器实现 FCD 流数据的存储与管理,通信服务器将 FCD 数

据包转换为 XML 文档并传送至一体机管理节点;管理节点解析 XML 文档获取 FCD 数据集,根据车载终端 ID 对 FCD 数据排序组织,基于各计算节点 CPU 的当前状态和处理能力,将 FCD 划分为多个数据包,并动态分配给各计算节点的 CPU.各计算节点 CPU 初始化时装载电子地图,当接收到分配的 FCD 数据后,进行数据预处理以实现噪声过滤,将各 GPS 数据与路段进行匹配处理,完成浮动车行驶轨迹推理和速度计算,当所分配的 FCD 数据包处理完毕后,将计算结果返回给管理节点。

管理节点接收到全部计算节点的处理结果后,计算各路段的平均速度,并根据流量-速度-密度模型推理对应路段的实时交通流密度和流量,获取路网交通流状态,提供给应用服务器处理。

2 FCD 大数据的动态任务划分与调度

任务划分是 FCD 并行处理的关键技术.由于浮动车数据具有松散耦合的特点,因而适于并行处理.在本文的方法中,管理节点将 FCD 流数据按照浮动车 ID 进行排序组织后,根据资源监控模块获取空闲的 CPU 数量,然后进行任务划分与调度.任务调度以实现各计算节点 CPU 的负载均衡为原则.为了减少计算节点之间的通信代价,同一个浮动车的数据分配给同一个计算节点 CPU 处理。

由于环境干扰,原始的浮动车数据集中存在一定数量的野值,导致相同记录数的 FCD 数据包处理时间可能存在较大差异.直接将 FCD 数据包按照空闲 CPU 数量均分的任务划分方法难以实现各计算节点的同步,使得整个系统的处理结果存在较大的延误.针对 FCD 的特点,本文提出了动态任务划分方法,详细描述如下:

设一个浮动车数据计算的最大时间 T_{\max} ,结合空闲的计算节点 CPU 数(令其为 C_{idle})和 FCD 处理的约束时间 T_{res} ,确定 FCD 流数据的初步划分.令 FCD 的大小为 M ,则 FCD 包的初始划分组数 N_0 为

$$N_0 = \text{int}\left(\frac{M \times T_{\max}}{T_{\text{res}}}\right) + 1 \quad (1)$$

式中, $\text{int}()$ 为取整函数.显然,满足在指定的约束时间内完成的条件为 $N_0 \leq C_{\text{idle}}$,因此确定 FCD 流数据的初始划分组数 $N_0 = C_{\text{idle}}$.

由于噪声和误差的存在,不同浮动车数据的处