

无线传感器网络数据融合技术

康健^{1,2} 左宪章² 唐力伟¹ 张西红³ 李浩⁴

(军械工程学院火炮工程系 石家庄 050003)¹ (军械工程学院电气工程系 石家庄 050003)²

(军械工程学院计算机工程系 石家庄 050003)³ (63880 部队 洛阳 471003)⁴

摘要 数据融合技术是无线传感器网络的一个关键技术,能减少传感器节点间的传输量,从而明显提高网络感知效能,延长网络生命周期,减小时间延迟。通过对尚处于研究阶段的数据融合技术进行详细分析与研究,阐述了数据融合技术的重要性,并分类介绍了现有的主要数据融合方法,最后指出该研究领域当前面临的挑战以及需要进一步研究的方向和有前景的研究课题——压缩感知。

关键词 数据融合,无线传感器网络,相关性,能量,压缩感知

中图分类号 TP393 **文献标识码** A

Survey on Data Aggregation of Wireless Sensor Networks

KANG Jian^{1,2} ZUO Xian-zhang² TANG Li-wei¹ ZHANG Xi-hong³ LI Hao⁴

(Department of Guns Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)¹

(Department of Electrical Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)²

(Department of Computer Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)³

(Unit 63880, Luoyang 471003, China)⁴

Abstract Data aggregation is a very important technique used to reduce the communication overhead and energy expenditure of sensor nodes during the process of data collection in a wireless sensor networks (WSN). The researches on data aggregation of WSN were presented. This paper introduced the function and significance of data aggregation, specified the existing and typical classification of data aggregation, pointed out the challenges and the promising research subjects for this research domain, especially compressive sensing.

Keywords Data aggregation, Wireless sensor networks, Dependency, Energy, Compressive sensing

无线传感器网络(wireless sensor networks, WSN)是当前国际上备受关注的、由多学科高度交叉形成的新兴前沿研究热点领域。WSN通过各类集成化的微型传感器协作地实时监测、感知和采集各种环境或监测对象的信息,并通过随机自组织无线通信网络以多跳中继方式将所感知信息传送到用户终端。这样,通过它就可以实现“无处不在的计算”的理念。WSN已成为C4ISR(command, control, communication, computing, intelligence, surveillance, reconnaissance and targeting)系统不可或缺的一部分。

WSN的发展受到包括能量供应、存储数据量、数据处理能力、数据传输速率、同步率、系统鲁棒性等诸多条件的限制和挑战。其中,有效的、不可再生的能量供给是WSN的最大挑战。在WSN中,能量主要消耗在3个方面:数据传输、信号处理和硬件操作。其中,数据传输消耗的能量最多。因此,能量和能力都有限的传感器节点如何实现复杂的数据监测和信息报告是WSN中需要解决的首要问题。由于WSN是面向应用的、以数据为中心的网络,因此可以利用网内数据处理

(in-network processing)技术,即采用数据融合(data aggregation或data fusion)技术来解决上述问题。

1 数据融合的重要性

数据融合是WSN中非常重要的一项技术,也是目前的一个研究热点^[1]。该技术可以通过一定的算法将传感器节点采集到的大量原始数据进行各种网内处理,去除其中的冗余信息,只将少量的有意义的处理结果传输给汇聚节点。采用数据融合技术能够大大减少WSN中需要传输的数据量,降低数据冲突,减轻网络拥塞,从而有效地节省能源开销,起到延长网络寿命的作用。

文献[2]通过理论分析和仿真测试两种手段对数据融合技术的影响进行了分析讨论。理论结果表明,采用数据融合技术和不采用数据融合技术消耗的能量之比的最小值为

$$\lim_{k \rightarrow \infty} \frac{N_D}{N_A} = \frac{1}{k} \quad (1)$$

其中, N_D 为数据融合情况下数据的传输次数; N_A 为不进行

到稿日期:2009-06-04 返修日期:2009-07-23 本文受国家自然科学基金(No. 60672143),军队科研计划项目资助。

康健(1981-),男,博士生,主要研究方向为兵器性能检测、故障诊断、网络测试等,E-mail:kangjian2001404@yahoo.com.cn;左宪章(1963-),男,硕士,教授,主要研究方向为无损检测、故障诊断等;唐力伟(1961-),男,博士,教授,主要研究方向为兵器性能检测与故障诊断等;张西红(1964-),男,博士,副教授,主要研究方向为无线网络、战术互联网等;李浩(1981-),男,硕士,助理研究员,主要研究方向为通信与仿真等。

数据融合情况下数据的传输次数; d 为传感器节点到汇聚节点的距离, k 代表数据源节点的数量。该式表示,网络中数据源节点越多(k 越大),数据融合技术节约的能量就越显著;同时,网络的规模越大(d 越大),数据融合技术节约的能量也越显著。这说明了数据融合对节省传感器节点能量起着重要的作用。

总的来说,WSN中数据融合的作用主要体现在节省能量、延长网络生命周期、提高数据收集效率、增强数据准确性以及获取综合性信息等方面。

2 数据融合方法分类及研究现状

数据融合技术涉及到检测技术、信号处理、通信、模式识别、决策论、不确定性理论、估计理论、最优化理论等众多学科领域。根据不同的融合规则,可分为不同的类型,例如有损融合、无损融合、依赖于应用的数据融合、独立于应用的数据融合、基于分布式数据库的数据融合^[3]、基于中心的数据融合等。下面从如何有效降低数据传输量和能量方面对WSN中现有的数据融合方法进行分类,介绍其研究进展情况。

2.1 基于生成树的数据融合

在WSN中,汇聚节点在收集数据时,是通过反向组播树的形式从分散的传感器节点逐步将监测数据汇集起来的。因为如果每个节点都生成监测报告并把数据直接发送给汇聚节点,则网络会拥塞,不仅耗能过多,而且会增加监测延时,降低数据包的传送速率。

现有的基于生成树的数据融合算法采用的主要是近源汇集(center at nearest source, CNS)、最短路径树(shortest paths tree, SPT)、贪婪增量树(greedy incremental tree, GIT),以及对它们进行的改进算法。文献[4]提出了DCTC(dynamic convey tree-based collaboration)算法。通过目标附近的节点协同构建动态生成树,协同组节点把测量数据沿确定的生成树向根节点传输,在传输过程中,汇聚节点对其子生成树节点的数据进行数据融合。文献[5]提出了MFST(minimum fusion steiner tree)算法,用于在WSN中以数据融合方式进行高效节能的数据收集。文中考虑了数据传输开销和数据融合开销,并且根据节点产生的数据量来选择融合数据点。文献[6]提出了AFST(adaptive fusion steiner tree)算法,该算法对MFST算法进行了改进,不仅要优化数据传输路由,而且在节点转发数据时,动态决定是否进行数据融合来进一步减少总的能量开销。

另外,在构造生成树时可以考虑采用基于MA(mobile agent)的协同信号和信息处理模式,以取得更好的融合性能。

基于生成树的数据融合方法虽然能够实现传输的数据进行融合的功能,但是有些问题仍需关注。首先是生成树的构建和融合时机的确定问题。Steiner树比较适合事件驱动的应用。另外,基于WSN的特点,有效的数据融合生成树的构造算法应该具有时间复杂度低以及分布式的特点。其次,应该考虑数据融合生成树的维护问题。当节点失效时,需要对生成树进行维护。生成树的维护可采用周期性更新和事件触发两种方式。其实,这个问题就是如何平衡网络能量消耗。而且,维护这些数据融合结构的能耗开销应尽可能小。

2.2 消除时空相关性的数据融合

在WSN中,同一地点布置了许多传感器,这些传感器之

间的相对距离较近,其观察值在时间和空间上存在相关性,导致数据冗余现象。因此,如何消除多传感器间的时空相关性,以降低网络传输数据及节省网络能量和开销,值得研究。而时间和空间上的这种相关性也是设计数据融合方法的主要依据。

现有文献对多传感器间时空融合的研究可分为两类:一类是时间或者空间融合模型;另一类是时空融合模型。

典型的时间融合模型是TiNA(temporal coherency-aware in-network aggregation)^[7],它利用传感器节点采集数据的时间一致性进行网内融合。其基本思想是,只有当前采集的数据与上一次采集的数据的差值大于某个用户指定的容忍限度时,节点才进行数据发送。这种方法对于监测数据波动较小的应用十分有效,能够显著地减少网络中的数据传输量。然而,当监测数据波动较大时,TiNA的作用就不是非常明显了;而且TiNA对于节点存储空间的要求比较高,尤其当网络规模较大时,转发节点需要保存大量的额外信息。

空间融合模型^[8]应用于WSN事件监测。该模型完全用节点间距离远近来衡量节点观测值间的相关程度,而实际上离事件源距离相同的节点的观测值间的相关性更高。

时空融合模型是消除时空相关性的数据融合研究的发展趋势,其典型的如文献[9,10]所述。此类模型主要是对时间融合和空间融合进行组合排序,融合性能不但取决于进行时间与空间融合的先后次序,而且还与具体采用何种融合方法有关,这些都需要做进一步的定量分析。例如文献[9]中采用D-S证据理论进行空间融合,计算量过于巨大,对处理能力有限的感知节点来说负担过重;此外,计算所造成的延时也将严重影响系统的实时性和同步性,限制了其在数据融合领域更加广泛的应用。因此,这些方面均可作为时空融合模型下一步研究的重点。

2.3 路由驱动型数据融合

路由驱动型数据融合算法主要基于两类路由协议:定向扩散路由协议和分层聚簇路由协议。

2.3.1 定向扩散路由型数据融合

定向扩散协议(directed diffusion, DD)^[11]是一个以数据为中心的路由和融合协议,为WSN提供有效的能量通信。在定向扩散协议中,数据融合包括路径建立阶段的兴趣(interest)融合和数据传输阶段的数据融合。然而,毕竟DD主要解决的是WSN中的路由问题,其中涉及的数据融合技术相对较为简单。同时,DD中的数据融合是基于其提出的属性值而对设计的,与具体应用关系非常紧密,是以数据为中心的一种数据融合技术。另外,在DD中,数据传输路由的确定还取决于节点到sink的时延,时延小的路径将被确定为数据传输路由。但这种方法并不利于邻近节点的数据融合,因此可采用改进的GIT算法。

2.3.2 分层聚簇路由型数据融合

低能耗自适应聚类分层(low-energy adaptive clustering hierarchy, LEACH)^[12]协议是典型的WSN路由协议,通过定期选举一些簇头节点,形成基于层次结构的路由机制。该路由协议可以支持簇头节点中的数据融合,能够体现出尽早进行数据融合的原则和优势。不过,LEACH仅强调了数据融合的重要性,并未给出具体的融合方法。而且,仅仅在簇头实施数据融合,簇内节点直接将原始数据送到簇头的策略也没

有充分利用相邻节点间的数据相关性来减少簇内节点传输的能耗。另外,LEACH 设定的假设太多。如:假设所有节点开始时具有相同的能量,而且各簇首消耗的能量基本相同。这样,若该协议用于节点能量不均匀的网络中,则需要进行调整,例如采用基于能量阈值方法^[13,14]来选择簇首,进行融合算法的设计。

PEGASIS(power-efficient gathering in sensor information systems)^[15]对 LEACH 的数据融合方式进行了改进,采用链式结构,具有更好的融合性能。然而,PEGASIS 中链的长度与节点数量有关,而 WSN 的节点数量规模很大,因此在数据融合时会产生很大的延迟。为了减少单链上的数据融合时延,可以使用多链结构。

PEDAP(power efficient data gathering and aggregation protocol)^[16]协议进一步发展了 PEGASIS 协议,其核心思想是把 WSN 的所有节点构造成一棵最小汇集树(minimum spanning tree)。节点不管在每一轮内接收到多少个来自各子节点的数据包,都将压缩融合为单个数据包,再进行转发,以最小化每轮数据传输的总能耗。然而,PEDAP 存在难以及时排除死亡节点(非能量耗尽)的缺点。

PEDAP-PA(power efficient data gathering and aggregation protocol-power aware)在 PEDAP 的基础上进一步提升了能量有效性,并在节能的基础上考虑了能量平衡特性,保证节点间的能量差不会随着网络运行的时间延续而变得巨大,也不会导致某些节点过早地失效,从而提高了每个节点在整个网络中的利用率,且能延长大部分节点的生命期。

对于如何降低或平衡融合节点的能量消耗速度可参考文献^[17,18]中所论述的方法。

总的来说,路由驱动型数据融合算法与具体路由的建立关系密切,还涉及到融合节点的选择、融合时机的选择及如何进行数据融合处理等问题。在路由驱动型数据融合算法中仍然有很多关键问题未解决,如融合增益与能量和通信服务质量的平衡、支持拓扑频繁改变的数据融合算法、安全数据融合等问题。

2.4 基于预测的时域数据融合

在基于预测的时域数据融合中,通过对传感器节点采集的时间序列数据进行分析,建立了能够反映时间序列中所包含的动态依存关系的数学模型,并对观察值进行预测。只有当预测值与实际值之间的误差超过给定阈值时,才传输数据。通过预测,减少不必要的数据传输,从而降低网络能耗,延长网络的生命周期。

现有文献采用的预测融合方法主要有概率模型^[19]、自回归滑动平均模型(autoressive moving average, ARMA)^[20]以及自适应的双重预测策略^[21]等。其中,自适应的双重预测策略在自适应预测模型选取时,由于考虑了时间序列、先验知识、预测精度以及可用的计算资源等因素,能实现在线的自适应的实时预测融合,并具有较好的可扩展性。

在此类融合方法中,如何根据采集的数据类型进行数据样本及预测融合算法的选择是问题的关键所在。再者,目前进行的主要是同步融合,关于异步融合的研究报道很少,值得关注。

2.5 基于分布式压缩的数据融合

通过对传感器节点采集到的数据进行分布式压缩实现融

合的功能主要有以下几种思路:分布式信源编码、报头压缩、数据包合并、基于小波变换的数据压缩融合、压缩感知。

2.5.1 分布式信源编码

以 Slepian-Wolf 编码为基础的分布式信源编码(distributed source coding)技术^[22]采用分布式方式,在每个信息源点进行数据压缩,使所有信息源点输出的信息消息量最少,而且互不相关。所有互不相关的数据无需网内汇聚处理,直接传送到 sink, sink 根据数据之间的相关性恢复出所有原始数据,并进行相应的后处理。路由算法可以采用简单的 SPT 算法。为了进一步减少网络中数据传输的开销,可采用节点最佳速率分配算法^[23],即离 sink 越远的、数据传输路径越长的节点产生的信息量越少;相反,离 sink 越近的、数据传输路径越短的节点产生的信息量越多。尽管这种方法可以在理论上达到数据采集的网络能耗开销最小化,但在实际应用中,需要事先了解全网所有节点数据之间的相关性,需要所有节点数据成功送达 sink,这对于节点数量众多、传输可靠性不高的 WSN 来讲,很难达到。由于联合考虑数据相关性和数据路由的最小能耗数据采集问题是 NP 完全问题^[23],最佳路由将介于 SPT 和旅行商路径(traveling salesman path, TSP)之间。研究者或是在特定假设条件下寻求最优解,或是提出启发式次优算法。因此,如何在实际的 WSN 中应用分布式源编码技术仍待解决。

2.5.2 报头压缩

报头压缩是基于报头信息存在大量冗余这一事实。首先是包里的冗余,即不同协议的报头携带了相同信息或可推论的信息;其次是连续包之间的冗余,即同一个包流的前后包的对应字段间只有微小的变化。根据报头中各个字段的变化方式,可将其分为不变字段和变化字段。这样,就可以根据具体情况设计相应的压缩算法,从而减小传输的数据量,提高信道利用效率。

文献^[24]提出了一种自适应的 TCP/IP 报头压缩算法,并提出基于信道状态进行自适应调整以进一步提高压缩算法的性能。这种算法适用于无线 IP 网络中对 TCP/IP 报头的压缩。文献^[25]提出了自适应的报头压缩机制 AAHC(advanced adaptive header compression)。AAHC 机制在获得较高压缩效率的同时,可保证报头压缩的抗差错鲁棒性,其性能优于传统的 ROHC(robust header compression)机制和 CRTP(compressed RTP)机制。但是,该机制没有考虑时延的影响。

2.5.3 数据包合并

数据包合并是 WSN 中一种有效的数据融合算法。数据包合并的主要思想是当某个节点收到多个子节点发来的数据包时,将它们合并成一个大的数据包,然后将合并后的数据包发送到父节点。在 WSN 中,数据字段相对较短,而控制字段相对较长。数据包合并能够有效地降低包头的开销。

典型的数据包合并算法包括数据漏斗(data funneling)^[26]以及 AIDA(application-independent data aggregation)^[27]等。数据漏斗实质上是一种基于簇的数据融合,簇头节点负责合并簇内节点的数据包。然而,数据漏斗要求节点具有自身的位置信息,并且有可能产生漏斗效应(funneling effect)^[28]。AIDA 是一种与应用无关的数据融合算法,实质上是在 MAC 层与网络层之间加入了一个数据融合层进行数

据包合并的操作。通过数据包合并, AIDA 能够有效地减少网络中的数据传输量, 降低无线信道中发生冲突的可能性。然而, AIDA 与应用相互独立, 无法利用高层次的语义信息对数据作进一步的压缩, 因此其融合度相对较低。

2.5.4 基于小波变换的数据压缩

小波变换已经被广泛应用于图像、视频等数据的压缩中。实验结果表明, 小波变换具有压缩率较高、图像恢复质量好、速度快等优点, 十分适合数据的高保真压缩。另外, 小波分解和重构算法是循环使用的, 易于硬件实现。因此可把小波变换应用到 WSN 数据压缩融合中。

从现有的文献来看, 如文献[29], 采用基于小波变换的压缩算法进行传感器网络数据融合能够取得良好的性能。但是, 对其运算开销有待进一步研究。这是一个关键性的问题, 如果解决不好, 就无法进行实际应用, 并且它还涉及到对时延性能的影响问题。

2.5.5 压缩感知

压缩感知或压缩传感 (compressive sensing, CS) 是近年来发展迅速的一个理论分支, 是一种新的数据获取理论。该理论指出: 对可压缩的信号可通过远低于 Nyquist 标准的方式进行数据采样, 仍能够精确地恢复出原始信号^[30,31]。在 CS 理论框架下进行数据融合研究, 能够显著降低数据传输量, 大大提高网络感知效能。

文献[32]提出了分布式 CS 的稀疏模型及重建算法。通过多个传感器同时测量某一现象, 基站收到所有的测量值并运行算法联合解码传感器所获取的信号, 重建传感器所监测的现象。这样, 能够在能量和带宽两个性能评价指标上取得较好的效果。文献[33]提出了一种分散压缩的方法, 指出可以采用 CS 理论来实现从分布式数据中提取有关信息并以无线的方式传输到远程目的地。并用 WSN 监测河水的例子阐述了这一问题。文献[34]提出了一种在压缩域中使用改进的采样模式的图像融合算法, 即使用一种简单最大选择融合策略对输入图像的压缩测量矩阵进行融合。文献[35]进行了基于智能传感器网络的分布式感知研究, 主要关注特征层的融合体系结构及分布式感知的情景评估结构的设计等, 并提出了将定位、关联、融合联合起来的策略。但是, 该文献却没有考虑将 CS 理论应用其中。如果将 CS 理论, 特别是将分布式 CS 理论应用到该文献中, 会取得更好的效果。

3 数据融合面临的挑战及发展趋势

3.1 面临的挑战

由于国内外的 WSN 研究日趋热烈, 在数据融合方面取得了很多研究成果。但是, 在 WSN 中进行数据融合, 除了其本身固有的资源和能量制约, 仍面临如下的一些挑战: (1) 对连续数据流的处理。在周期性监测应用中, 需要传感器节点周期性地传输数据, 相邻轮次的数据采集具有一定的相关性, 需要利用历史信息等减少传输的数据量。(2) 无线多跳网络。传感器节点需要协作进行数据传输。(3) 资源开销。在许多 WSN 的应用场合中, 数据融合的能量开销是不能忽略不计的, 如在音视频传感器网络应用或加密传输的 WSN 应用中, 数据融合开销和数据传输开销非常接近, 而且这类应用还将越来越普遍。(4) 安全问题。无人值守和恶劣的应用环境特点要求所采用的数据融合技术是安全的, 因为攻击者能够通

过对传感器节点进行攻击, 获得对捕获节点的完全控制权, 使得捕获节点能发送伪造的数据以改变整个融合结果。

3.2 发展趋势

(1) 时间序列预测融合

时间序列预测融合的关键在于对周期性时间序列数据的分析及预测融合模型的选择。如何考虑采集数据的非线性、随机性以及突变情况, 以及预测样本数目有限即小样本情况下融合模型的建立问题, 是进一步的研究方向。而采用灰色理论和支持向量机相结合的方法有望解决这个问题。因为灰色理论把一切随机变量、随机过程均看作灰色变量、灰色过程, 具有处理随机过程、突变的能力。而支持向量机较好地解决了局部极小点、维数灾难、小样本学习以及过拟合等传统学习算法所不可回避的难题, 并且具有很强的泛化能力。

针对异步采样数据的多传感器数据时间序列预测融合问题, 可研究采用基于分布式的预测融合算法, 大大降低融合中心的计算负荷, 并充分有效地利用系统资源, 以达到较高的预测融合精确度。

(2) 分层聚类数据融合

对于多传感器节点协作式数据传输仍需进行分层聚类数据融合算法或体系结构的研究。难点在于如何进行分层、自适应动态聚类以及设计多节点传输场景下的数据融合数学模型。可采用模糊聚类或基于核函数最大间隔聚类算法进行簇头选择, 并采用相关优化算法对融合过程进行优化, 提高融合计算的准确性, 降低能耗和执行时间。同时, 考虑 QoS 性能指标, 进行相关性能参数的平衡或折中, 以确保网络 QoS 的实现。

(3) 数据融合的影响

数据融合最主要的优点就是会减少信息开销, 从而提高能量效率和延长网络的生命周期。因此, 在数据融合算法设计及应用中, 需考虑数据融合可能带来的节能增益和能量开销。研究传感器能量、计算能力和传输能力等各种约束下的最优融合, 并综合考虑数据融合和数据传输开销等对数据融合算法的影响, 以获得网络传感器的最佳配置, 是数据融合研究的发展趋势。这就要求融合算法慎重地考虑其对准确度、等待时间、鲁棒性等因素的影响, 并且联合 MAC、路由协议等进行数据融合模型的整体设计。

(4) 安全数据融合

为了抵抗在融合过程中节点被捕获的攻击以及量化融合结果中的不确定性, 需同时考虑数据机密性和数据完整性, 通过设计伪造数据入侵攻击检测策略、逐跳加密数据融合、端到端加密数据融合、传感器节点失效恢复机制等来保证数据融合过程中的安全性。还可采用基于信任的融合构架, 通过用节点的信誉度来衡量节点的可信赖程度, 解决融合过程中存在的安全问题。

(5) 基于 CS 的数据融合

CS 理论的出现为 WSN 数据融合的研究提供了一个很有发展前景的新课题。WSN 数据融合中, 在信号采样时就通过压缩融合大大降低了采集的数据量, 从而使 WSN 的节能功效提高了一个或几个数量级。这对 WSN 而言是实实在在的好处。有关这方面的研究, 国外刚刚起步, 而国内还没有开展相关研究。

基于 CS 的数据融合, 通过设计测量矩阵实现前端传感

器节点处的采样数据压缩功能,可大大降低采集的数据量,并在数据传输过程中进行融合处理,可有效解决周期性的大量连续数据流的处理问题,这具有深远的理论和现实意义。其中,构建硬件容易实现的测量矩阵和快速稳定的重建算法是将基于CS的数据融合推向实用化的关键,也是基于CS的数据融合的主要研究内容。

结束语 数据融合技术是为适应WSN以数据为中心的应用而产生的,主要关注如何对采集到的或接收到的其它传感器节点发送的多个数据进行融合处理,消除冗余信息,然后再传输处理后的数据,其重点在于减少需要传输的数据。关于数据融合的研究虽然已经取得了很大进展,但是研究工作尚处于起步阶段,大量问题还没有涉及到。WSN数据融合是与应用高度相关的,应根据不同的应用需求采用合适的数据简化技术,并依据此技术的特点,采用新兴的交叉学科知识来设计数据融合算法。因此,对于WSN数据融合算法的设计与分析既具有很高的研究价值和创新性,又存在一定的挑战性。

参考文献

- [1] Nakamura E F, Loureiro P A F, Frery P C. Information fusion for wireless sensor networks: methods, models, and classifications[J]. *ACM Computing Surveys*, 2007, 39(3): 1-55
- [2] Krishnamaehari B, Estrin D, Wicker S. The Impact of data aggregation in wireless sensor networks[C]// *The 22nd International Conference on Distributed Computing Systems Workshops*. Vienna, Austria; IEEE Press, 2002; 575-578
- [3] Madden S R, Franklin M J, Hellerstein J M, et al. TinyDB: an acquisitional query processing system for sensor networks[J]. *ACM Transactions on Database Systems*, 2005, 30(1): 122-173
- [4] Zhang Wen-sheng, Cao Guo-hong. DCTC: dynamic convoy tree-based collaboration for target tracking in sensor networks[J]. *IEEE Transactions on Wireless Communications*, 2004, 3(5): 1689-1701
- [5] Luo Hong, Liu Yong-he, Sajal K D. Routing correlated data with fusion cost in wireless sensor networks[J]. *IEEE Transactions on Mobile Computing*, 2006, 5(11): 1620-1632
- [6] Luo Hong, Luo Jun, Liu Yong-he, et al. Adaptive data fusion for energy efficient routing in wireless sensor networks[J]. *IEEE Transactions on Computers*, 2006, 55(10): 1286-1299
- [7] Sharaf M A, Beaver J, Labrinidis A, et al. TiNA: a scheme for temporal coherency aware in network aggregation[C]// *The 3rd ACM International Workshop on Data Engineering for Wireless and Mobile Access*. San Diego; ACM Press, 2003; 69-76
- [8] Vuran M C, Akan O B, Akyildiz I F. Spatio-temporal correlation: theory and applications for wireless sensor networks[J]. *Computer Networks*, 2004, 45(3): 245-259
- [9] 罗大庸, 张远. 多传感器信息时空融合模型及算法研究[J]. *系统工程与电子技术*, 2004, 26(1): 36-39
- [10] 郭利, 马彦恒, 张锡恩. 一种多传感器数据时空融合估计算法[J]. *系统工程与电子技术*, 2005, 27(12): 2016-2019
- [11] Intanagonwivat C, Govindan R, Estrin D, et al. Directed diffusion for wireless sensor networks[J]. *IEEE/ACM Transactions on Networking*, 2003, 11(1): 2-16
- [12] Heinzelman W B, Chandrakasan A P, Balakrishnan H. An application specific protocol architecture for wireless microsensor networks[J]. *IEEE Transactions on Wireless Communications*, 2002, 1(4): 660-670
- [13] Manjeshwar A, Agrawal D P. TEEN: a routing protocol for enhanced efficiency in wireless sensor networks[C]// *The 15th Parallel and Distributed Processing Symposium*. San Francisco: IEEE Press, 2001; 2009-2015
- [14] Manjeshwar A, Agrawal D P. APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks[C]// *Proceedings of the International Parallel and Distributed Processing Symposium*. New York: IEEE Press, 2002; 195-202
- [15] Lindsey S, Raghavendra C S. PEGASIS: power-efficient gathering in sensor information systems[C]// *Proceedings of IEEE Aerospace Conference*. Big Sky, MT; IEEE Press, 2002; 1125-1130
- [16] Tan H O, Korpeoglu I. Power efficient data gathering and aggregation in wireless sensor networks[J]. *ACM SIGMOD Record*, 2003, 32(4): 66-71
- [17] Abusaimeh H, Yang Shuang-hua. Balancing the power consumption speed in flat and hierarchical WSN[J]. *International Journal of Automation and Computing*, 2008, 5(4): 366-375
- [18] Abusaimeh H, Yang Shuang-hua. Dynamic cluster head for lifetime efficiency in WSN[J]. *International Journal of Automation and Computing*, 2009, 6(1): 48-54
- [19] Chu D, Deshpande A, Hellerstein J M, et al. Approximate Data Collection in Sensor Networks using Probabilistic Models[C]// *The 22nd International Conference on Data Engineering*. Washington, DC, USA; IEEE Computer Society, 2006; 48-59
- [20] Wang Xue, Ma Jun-jie, Wang Sheng, et al. Time series forecasting for energy-efficient organization of wireless sensor networks[J]. *Sensors*, 2007, 7: 1766-1792
- [21] Le B Y A, Santini S, Bontempi G. Adaptive model selection for time series prediction in wireless sensor networks[J]. *Signal Processing*, 2007, 87(12): 3010-3020
- [22] Xiong Zi-xiang, Liveris A D, Cheng S. Distributed source coding for sensor networks[J]. *IEEE Signal Processing Magazine*, 2004, 21(5): 80-94
- [23] Cristescu R, Beferull L B, Vetterli M. On network correlated data gathering[C]// *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*. Hong Kong, China; IEEE Press, 2004; 2571-2582
- [24] 吴亦川, 黄奎, 郑建平, 等. 一种自适应的TCP/IP报头压缩算法[J]. *计算机研究与发展*, 2005, 42(4): 655-661
- [25] 周新运, 孙利民, 皇甫伟, 等. 无线多媒体传感器网络中一种自适应的报头压缩机制[J]. *软件学报*, 2007, 18(5): 1122-1129
- [26] Petrovic D, Shah R C, Ramchandran K, et al. Data funneling: routing with aggregation and compression for wireless sensor networks[C]// *The First IEEE International Workshop on Sensor Network Protocols and Applications*. Anchorage, AK; IEEE Press, 2003; 156-162
- [27] He T, Blum B M, Stankovic J A, et al. AIDA: adaptive application-independent data aggregation in wireless sensor networks[J]. *ACM Transactions on Embedded Computing Systems*, 2004, 3(2): 426-457

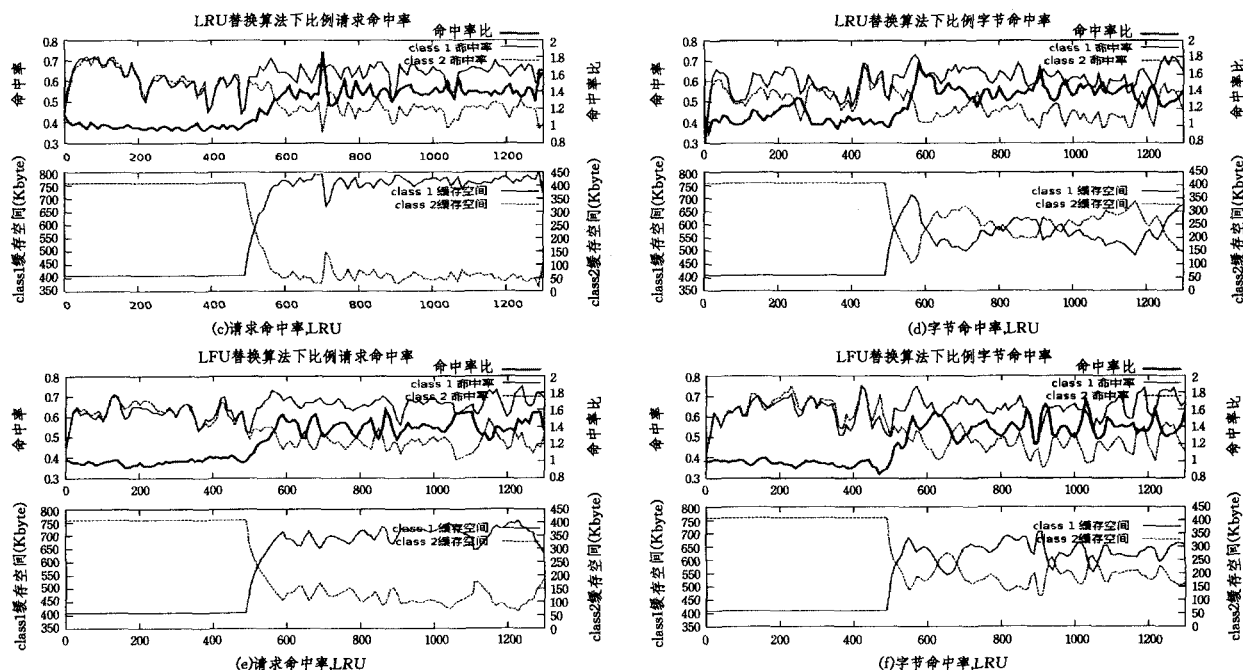


图4 实验结果

5) 同样条件下,若以字节命中率计算,抖动较大。这是由于大文件的丢失(Miss hit)对字节命中率影响较大的缘故,而请求文件的重尾分布,使得大文件被请求的次数和频率都较小,因此无论在何种替换算法下,都不易命中,从而造成了字节命中率的抖动。

结束语 本文对静态 Web 缓存模块进行了研究,建立了比例命中率保证的分离式 Cache 结构,根据系统建模和参数辨识确定被控对象的数学模型,并在此基础上完成了缓存控制器的设计。最后,在 Windows 平台下,对 Apache2 的 mod_mem_cache 模块进行了修改并分别对 GSDF, LRU, LFU 缓存替换算法进行了实验,验证了该系统模型的有效性,实现了 Web 缓存区分服务。

参考文献

[1] Dovrolis C, Stiliadis D, Ramanathan P, et al. Proportional differentiated services; delay differentiation and packet scheduling[C] // Proceedings of the ACM SIGCOMM, 1999
 [2] Lu Y, Abdelzaher T F, Saxena A. Design, Implementation, and

Evaluation of Differentiated Caching Services[J]. IEEE Transactions on Parallel and Distributed Systems, 2004

[3] Lu A, Saxena A, Abdelzaher T F. Differentiated caching services; a control-theoretical approach [C] // 21st International Conference on Distributed Computing Systems, 2001
 [4] Venketesh P, Sivanandam S N, Manigandan S. Enhancing Qos in Web Caching using Differentiated Services [J]. International Journal of Computer Science & Applications, 2006, 3(1)
 [5] Zhang Y, Feng W, Hurley R. Integration of QoS Queuing Schedules to QoS Caching Schemes[C] // Proceedings of the 5th ICIS-COM SAR, 2006
 [6] Barford P, Crovella M. Generating Representative Web Workloads for Network and Server Performance Evaluation[J]. ACM SIGMETRICS Performance Evaluation Review, 1998, 26(1): 151-160
 [7] Cao P, Irani S. Cost-aware www proxy caching algorithms[C] // Proceedings of the 1997 USENIX Symposium on Internet Technology and Systems, December 1997; 193-206

(上接第 35 页)

[28] Ee C T, Bajcsy R. Congestion control and fairness for many-to-one routing in sensor networks[C] // The 2nd International Conference on Embedded Networked Sensor Systems, Baltimore, MD; ACM Press, 2004; 148-161
 [29] Xu Ning, Rangwala S, Chintalapudi K K, et al. A wireless sensor network for structural monitoring[C] // The 2nd International Conference on Embedded Networked Sensor Systems, Baltimore, MD; ACM Press, 2004; 13-24
 [30] Donoho D L. Compressed Sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306
 [31] Candes E J, Wakin M B. An introduction to compressive sampling[J]. IEEE Signal Processing Magazine, 2008, 25(2): 21-30
 [32] Hormati A, Vetterli M. Distributed compressed sensing; sparsity

models and reconstruction algorithms using annihilating filter [C] // IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV; IEEE Press, 2008; 5141-5144
 [33] Haupt J, Bajwa W U, Rabbat M, et al. Compressed sensing for networked data[J]. IEEE Signal Processing Magazine, 2008, 25(2): 92-101
 [34] Wan Tao, Canagarajah N, Achim A. Compressive image fusion [C] // IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, Nevada; IEEE Press, 2008; 1308-1311
 [35] Leung H, Chandana S, Wei Shuang. Distributed sensing based on intelligent sensor networks[J]. IEEE Circuits and Systems Magazine, 2008, 8(2): 38-52