基于压缩感知的土壤呼吸监测传感网动态采样调度策略

王国英^{①③}, 江雨佳^①, 莫路锋^{①③*}, 孙艳丰^②, 周国模^④

① 浙江农林大学信息工程学院, 临安 311300

② 北京工业大学计算机学院, 北京 100124

③ 西安交通大学电子与信息工程学院, 西安 710049

④ 浙江农林大学浙江省森林生态系统碳循环与固碳减排重点实验室, 临安 311300

* 通信作者. E-mail: molufeng@gmail.com

收稿日期: 2013-04-01; 接受日期: 2013-08-07

国家重点基础研究发展计划 (批准号: 2011CB302705)、国家自然科学基金 (批准号: 61190114, 61303236)、浙江省自然科学基金 (批准号: LY12F02016) 和浙江省科技计划项目重大科技专项 (批准号: 2012C13011-1) 资助项目

摘要 土壤呼吸监测传感网通常部署在野外, 无法直接访问电力, 并且土壤呼吸测量过程比较复杂、能耗较高, 因此在满足重建准确度要求的前提下, 希望以尽量少的采样次数来进行测量. 利用土壤呼吸真实物理过程中时间序列的相关性, 可以采用压缩感知理论来实现采样调度. 本文提出了一种基于压缩感知的分段动态采样调度策略: 利用前期测量数据进行分析得到的先验知识, 对测量时间区间的序列数据进行分段线性拟合, 依据分段数据子序列的线性程度度量指标, 动态确定各段的采样率. 土壤呼吸测量仪在进行监测时, 按照分段动态变化的采样率构造压缩感知采样和重建所需的测量矩阵. 实验结果表明, 相比平均采样率相同的固定采样策略, 本文提出的分段动态采样策略能够得到更好的重建质量, 即如果以确定的重建误差阈值作为需求, 则本文的动态采样策略具有更小的采样率. 虽然计算动态采样率会带来一定的计算开销, 但减少的采样次数可以节省更多的能量. 本文提出的基于压缩感知的分段动态采样策略, 虽然在土壤呼吸监测传感网的应用场景中进行了实验和分析, 但其思想对于其他类似应用的采样调度和节能问题也具有借鉴价值和潜在的可用性.

关键词 采样 调度 压缩感知 无线传感网 土壤呼吸

1 引言

陆地生态系统是全球碳库的主体, 它的碳循环由地上和地下两个部分组成. 地上部分主要指植物通过光合作用吸收 CO_2 ; 地下部分包括根系呼吸、土壤微生物呼吸和土壤动物的异养呼吸, 吸收和排放 CO_2 , 即称为土壤呼吸^[1]. 地面碳积累是陆地生态系统结构和功能综合作用的产物, 全球陆地土壤的碳库量约为 $1300\sim 2000\text{Pg}(\text{C})$, 是全球陆地植被碳库 $500\sim 600\text{Pg}(\text{C})$ 的 $2\sim 3$ 倍和大气碳库 $750\text{Pg}(\text{C})$ 的 2 倍多, 在全球碳素收支平衡中具有极为重要的地位^[2]. 土壤呼吸是影响大气 CO_2 浓度变化的一个重要因素, 通过土壤呼吸排放到大气中的 CO_2 大约为 $68\text{Pg}(\text{C})\cdot\text{a}^{-1}$, 约占大气 CO_2 总量的 10% ^[3].

引用格式: 王国英, 江雨佳, 莫路锋, 等. 基于压缩感知的土壤呼吸监测传感网动态采样调度策略. 中国科学: 信息科学, 2013, 43: 1326–1341, doi: 10.1360/112013-118

因此, 客观评价土壤呼吸碳通量对全球碳收支的贡献度以及了解和深入分析土壤呼吸对全球气候变化的影响尤其重要, 已经成为全球碳循环研究的核心问题之一。

由于土壤呼吸监测网络通常部署在野外环境, 无法直接访问电力, 因此必须解决整个系统的高效节能问题, 以便使监测网络能够以无人管理的方式运行尽量长的时间. 无线传感网中能耗的主要来源包括指令处理 (processing)、感知 (sensing) 和无线通信 (radio) 等操作^[4]. 本文主要关注其中第二个能耗来源: 感知. 为了收集到土壤呼吸碳通量随时间变化的详细信息, 理想的采样策略是以非常高的频率进行测量. 然而, 由于土壤呼吸测量是一个耗时耗能的过程, 是土壤呼吸监测节点能耗的重要原因. 因此, 本文的主要目标是土壤呼吸监测节点设计合适的采样调度策略, 减少土壤呼吸监测节点的采样次数以节省能量消耗.

用尽量少的采样次数进行土壤呼吸测量, 并且利用采集到的测量数据对未采样时刻的土壤呼吸估算达到需要的准确度, 这个目标通常可以采用插值或拟合的方法实现. 本文采用压缩感知理论来达到这个稀疏采样目标: 利用土壤呼吸碳通量时间序列的相关性, 可以对物理世界真值进行稀疏化; 利用稀疏采样数据, 可以恢复出满足准确度需要的土壤呼吸碳通量时间序列数据.

在利用压缩感知理论进行土壤呼吸监测节点的稀疏采样和数据恢复时, 需要确定两个矩阵: 表达基矩阵 Ψ , 能够对土壤呼吸真实值进行稀疏化; 测量矩阵 Φ , 用来表示采样策略, 通常可以采用一定采样率的随机或均匀采样.

测量矩阵中体现的采样率决定了节点进行土壤呼吸测量的次数, 即决定了相比稠密测量时节点的节能效果, 同时也直接影响着数据恢复的准确程度. 直观地理解, 采样率越低, 则此采样调度策略的节能效果越好, 但可能会导致恢复数据的误差较大. 因此在采样率和恢复数据的准确性之间存在矛盾, 设计采样调度策略时要根据对恢复数据的准确性需求来对二者进行权衡.

本文提出基于压缩感知的动态采样节能策略 SDSS: 根据数据随时间变化的情况, 采用不同的采样率动态构造测量矩阵, 即数据变化剧烈时提高采样率, 数据变化平缓时降低采样率, 从而实现在节能效果与恢复数据准确性之间进行灵活的动态折中. 由于无法提前知道即将采集到的真实数据, 本文研究了利用从历史数据观察得到的先验知识来进行分段采样的采样调度策略; 并通过我们的土壤呼吸监测传感网中的数据对提出的动态采样调度策略进行分析.

虽然本文提出的动态采样调度策略是基于土壤呼吸监测传感网, 相关性能分析也是利用此应用场景的数据进行的, 但本文提出的分段动态采样调度策略是通用的, 可广泛适用于其他具有先验规律的稀疏采样应用场景.

本文后续内容安排如下: 第 2 小节介绍土壤呼吸监测传感网系统的结构和我们自主研制的土壤呼吸测量仪的采样过程; 第 3 小节对基于压缩感知的土壤呼吸传感网节点采样问题进行了分析和建模; 第 4 小节提出了基于先验知识的分段动态采样调度策略; 第 5 小节介绍了实验数据和实验过程设计, 通过实测数据对提出的分段动态采样调度策略进行了测试; 通过实验结果的比较, 对分段动态采样调度测量的性能进行了分析. 最后一节对全文进行了总结并分析了可继续研究的方向.

2 土壤呼吸监测传感网系统

对土壤呼吸碳通量的测量, 利用现有设备 (例如: 美国 LI-COR 公司的 LI-8100 系列开路式土壤呼吸测量系统^[5]) 可以测量获得较准确的单点土壤碳通量数据, 但不能直接测量区域土壤呼吸碳通量. 由于 LI-8100 设备价格昂贵, 在同一时刻多点同时采样不太实际, 通常只能通过选择一个特征点进行测量作为区域代表数据^[6], 或者选择少量特征点在不同时刻进行测量并通过插值等方法获得除采样点

图 1 土壤呼吸监测传感网结构

Figure 1 Architecture of a soil respiration monitoring sensor network

以外位置的数据 [7], 来实现对区域土壤呼吸碳通量的估算. 然而, 由于受到土壤孔隙大小和类型、土壤温度和湿度、风速以及土气之间的 CO_2 浓度梯度等多种因素的影响, 导致土壤呼吸碳通量的空间和时间异质性很大, 采用选取少量特征点进行依次监测的方法难以实现对区域土壤呼吸碳通量动态的准确监测.

我们自主研制了支持无线传感网通信的土壤呼吸测量仪 (见图 1), 利用无线传感网的优势, 可以满足区域土壤呼吸碳通量准确估算在区域覆盖、同步采样、持续监测等方面的需求.

土壤呼吸监测传感网系统如图 1 所示, 其中部署着两种节点: 普通传感网通信节点和自主研制的土壤呼吸测量仪节点. 由于土壤呼吸测量仪的成本较高, 不适宜大量使用, 而在大范围的区域监测应用中, 仅靠土壤呼吸测量仪无法保证网络的强连通性; 另外由于土壤呼吸测量过程需要消耗比较高的能量, 如果同时还负责网络的组网路由和数据收集功能会大大降低土壤呼吸测量仪的生命期, 因此采用土壤呼吸测量仪与普通传感网通信节点相结合的网络部署方式, 土壤呼吸测量仪只作为数据收集树的叶子节点, 将测量到的土壤呼吸碳通量数据经由普通传感网通信节点传送至 sink 节点.

在土壤呼吸测量仪中, 为了测量一次土壤呼吸碳通量数据, 需要感知土壤温度、湿度、密闭气室内的气压和 CO_2 浓度. 其中温度、湿度、气压测量的能耗较低; 而 CO_2 浓度变化的测量是个较复杂的操作过程, 需要消耗比较高的能量. 每次测量需要先封闭气室持续测量 3 分钟, 期间每 3 秒测量一次气室内的 CO_2 浓度, 再自动打开气室与外界通风 1 分钟以使得后续测量结果能反映真实的土壤呼吸过程. 在持续测量的 3 分钟内, 共测得 60 个 CO_2 浓度数据, 对它们进行线性拟合并计算斜率得到测量期间的气室内 CO_2 浓度变化速率, 然后结合土壤温度、湿度、密闭气室内气压等参数因子, 利用碳通量计算公式得到土壤呼吸通量数据. 因此土壤呼吸测量能耗较大, 通过压缩感知理论来减少采样次数可有效延长设备的监测生命期.

3 基于压缩感知的土壤呼吸传感网采样调度问题描述

3.1 压缩感知

密集的周期性采样策略可以获得大量的数据, 但这样是否最有利于认识真实物理过程? 此问题的

答案表面看来应该是肯定的, 但事实并非如此, 因为数据量增加并不一定意味着信息量的增加. 相反, 过量冗余的含噪数据会掩盖有效数据蕴含的主要结构 (主成份), 同时也增加了采样难度和采样代价.

压缩感知 (compressive sensing) 主要借助于数据稀疏特征和原始数据的低秩特性, 在小于 Nyquist 采样率的条件下, 获取信号的少量离散样本, 然后通过非线性重建算法重建信号 [8,9] 这个理论已经被应用到数据压缩 [10]、信道编码 [11]、模拟信号感知 [12]、路由 [13]、数据收集 [14] 等各个方面.

对于一个大小为 N 的向量 x 表示的离散信号 ($\|x\|_0 \ll N$), 对 x 的测量可以抽象为一个 $M \times N$ 的矩阵 Φ (通常称为测量矩阵), 得到测量结果为大小为 M 的向量

$$y = \Phi x, \quad (1)$$

那么, 至少需要多少次测量才能重建信号 x ? 由线性代数知识可知, 为使 (1) 式的解存在且唯一, 须选择 $M \geq N$, 即需要至少进行 N 次观测. 但是, 如果 x 足够稀疏 ($\|x\|_0 \ll N$), 则理论上存在减少观测次数 M 的可能性.

实际中, x 却很可能不是稀疏的, 但有可能在另一个域内能够被稀疏化表达. 具体来说, 利用某些 $N \times N$ 的矩阵 Ψ , x 可以被写为:

$$x = \Psi s, \quad (2)$$

其中, s 是在 Ψ 域内的 $N \times 1$ 稀疏向量, $\|s\|_0 = K$, 且 $K \ll N$. 矩阵 Ψ 也被称为表达基. 因此采样向量可以写成:

$$y = \Phi \Psi s. \quad (3)$$

因而, 压缩感知的研究和应用, 主要包括三个方面的问题:

- (1) 采样前, 根据 x 的特性, 设计稀疏化 x 所需的表达基矩阵 Ψ ;
- (2) 采样时, 对尽量小的 M , 设计 $M \times N$ 测量矩阵 Φ ;
- (3) 重建信号时, 对给定的测量 y 和已知的矩阵 Φ 和 Ψ , 根据恢复算法 A , 确定 s , 重建原始信号 $x = \Psi s$.

对于第一个问题, 最重要的就是选择的表达基矩阵 Ψ 能够将 x 很好的稀疏化. 对于平滑数据, 采用小波基作为表达基矩阵通常可以达到近似的稀疏化: 其展开系数大多数绝对值较小.

对于第二个问题, 由于在第三个任务中需要用到测量矩阵 Φ , 因此 Φ 不能随意的选择, 需要具有有限等距性 (restricted isometry principle, RIP) [11]. 目前采用的测量矩阵有以 Bernoulli 矩阵为代表的高斯随机测量矩阵、Fourier 矩阵等.

对于第三个问题, 由于 $M \ll N$, (3) 式是一个非确定的线性系统, 求解这个系统是近年来广泛研究的主题. 第一类方法是选择具有最小 l_0 范式的 s :

$$\min_{s \in \mathbb{R}^N} \|s\|_0 \quad \text{s.t.} \quad y = \Phi \Psi s. \quad (4)$$

对此进行直接求解是很困难的 [11,15]. 若 N 比较大, 通常是不可解的. 但是存在使用平滑 l_0 范式的快速方法, 例如 SL0 方法 [16]. 第二类方法用 l_1 范式最小化问题来代替 l_0 最小化问题来降低复杂度, 被称为基追踪 (basis pursuit, BP) [17]:

$$\min_{s \in \mathbb{R}^N} \|s\|_1 \quad \text{s.t.} \quad y = \Phi \Psi s. \quad (5)$$

它可以使用线性规划 (LP) 方法解决. 存在能用多项式时间解决上述问题的算法, 包含内部点方法 (interior-point), 也有针对大规模系统的更快的算法 [18,19]. 除了线性规划 LP 方法, 常用的算法将包括

Matching Pursuit (MP), 例如 OMP^[18] 和 ROMP^[19]. 它们被认为比 LP 方法更快, 但是估计质量更差, 特别当信号不是足够稀疏时.

如果 M 满足:

$$M \geq C\mu^2(\Phi, \Psi)K \log N, \quad (6)$$

则使用上述的任何一种恢复算法 A , 从 M 个测量重建一个 K 稀疏的信号的可能性很高. 其中 C 是一个正的常量, N 是信号的维数, $\mu(\Phi, \Psi)$ 是两个矩阵 Φ, Ψ 的相干性. 给定一对 $\mathbb{R}^N$ 的 (Φ, Ψ) , 相干可以定义为:

$$\mu(\Phi, \Psi) = \sqrt{N} \max_{1 \leq i, j \leq N} |\langle \phi_i, \psi_j \rangle| \in [1, \sqrt{N}], \quad (7)$$

这里 ϕ_i, ψ_j 分别是 Φ, Ψ 的列向量. 因此, 在 Φ 和 x 确定的情况下, 必须仔细选择 Ψ , 它既需要在 Ψ 域稀疏表达 x ; 同时还需要使 $\mu(\Phi, \Psi)$ 也尽量小.

3.2 时域采样调度问题建模

在真实物理世界中, 采样点的土壤呼吸碳通量在时间维度上应该是连续的. 但是如果时间单位相对于土壤呼吸变化的时间尺度来说是足够小, 则可以近似认为这个变化过程是离散的. 同样, 在现实中, 无论采样频率多高, 测量设备的操作都是离散的, 能够测量到的土壤呼吸碳通量数据在时间上显然也是离散的. 因此, 我们选择采用离散时间模型来对土壤呼吸监测传感网节点的采样调度问题进行建模^[20]:

(a) 采样位置在一段时间内的真实物理世界土壤呼吸过程的时序数据用 $X = \{x_t\}, (t = 1, 2, \dots, N)$ 来表示; t 表示时间序列时刻, 一共 N 个时刻.

(b) 节点的采样调度策略 π 用一个时间序列 $T^\pi = \{t_1, t_2, \dots, t_n\}, (t_i \in \{1, 2, \dots, N\}, 1 \leq i \leq n)$ 来表达, 表示在这些时刻进行采样.

(c) 当节点按照采样调度策略 π 进行多次测量, 并且假设测量是没有误差和噪声, 则得到这样的采样数据序列: $X^\pi = \{x_{t_1}, x_{t_2}, \dots, x_{t_n}\}$, 即真实物理世界时序数据中的部分值构成的序列.

(d) 为了解真实的土壤呼吸过程, 需要按按照采样调度策略 π 进行多次测量, 然后利用采样得到的部分真实值来重建整个土壤呼吸过程的碳通量时序数据. 即, 根据采样序列 X^π , 采用估计函数 λ , 生成原始序列的估计 $\hat{X}^\lambda = \{\hat{x}_t\}, (t = 1, 2, \dots, N)$, 其中, 如果 $t \in T^\pi$, 则 $\hat{x}_t$ 即为 x_t , 否则 $\hat{x}_t$ 为利用估计函数得到的 $\hat{x}_t^\lambda(x^\pi)$.

(e) 基于以上描述, 采样调度策略的设计目标就是选择最佳采样策略 π 和估计函数 λ , 以便最小化重建的土壤呼吸数据序列 $\hat{X}^\lambda$ 与原始真实物理世界土壤呼吸数据序列 X 之间的估计误差, 并且平均采样率不能超过一个设定的界限, 即

$$\min_{\pi, \lambda} Err(X, \hat{X}^\lambda(X^\pi)), \text{ s.t. } n/N \leq \alpha, \quad (8)$$

这里 Err 是一个特定的误差度量, α 是设定的最大采样率阈值.

3.3 基于压缩感知的采样调度问题建模

在土壤呼吸监测传感网应用中, 采用压缩感知理论来设计节点的采样调度策略. 即用压缩感知理论来对 3.2 小节中的原始数据序列 X 、采样调度策略 π 、采样数据序列 X^π 、重建的估计序列 $\hat{X}^\lambda$ 、误差度量 Err 等因素进行设计.

(a) 原始数据序列 X : 仍然采用一个 $N \times 1$ 的向量 $X = \{x_t, t = 1, 2, \dots, N\}$ 来表示, 对应压缩感知方程式 (1) 中的 x .

(b) 采样调度策略 π : 采用压缩感知方程式中的测量矩阵 $\Phi_{M \times N}$ 来表示采样调度策略. 矩阵的行数 M 即为采样的次数; 矩阵的列表示采样的时机, 每行表示一次采样的时机设计. 如果矩阵 Φ 的第 m 行第 n 列的位置为 1, 则表示第 m 次测量在时刻 n 时发生.

(c) 采样数据序列 X^π : 由采样调度策略 $\pi = \Phi_{M \times N}$ 可知, 如果 Φ 中 (m, n) 位置元素为 1, 则第 m 次采样在时刻 n 发生, 测量到的数据为 x_n , 全部 M 次测量到的采样数据序列即为 ΦX , 对应为压缩感知方程式 (1) 中的 y .

(d) 重建数据序列 $\hat{X}^\lambda$: 根据给定的采样数据序列 y 和矩阵 Φ 、 Ψ , 对压缩感知方程式按照某种求解方法得到 s 后, 则重建数据序列即为 Ψs .

(e) 重建误差度量 Err : 对重建质量的评价, 可以采用重建数据序列 Ψs 与原始数据序列 X 中对应各元素的平均误差 $\|\Psi s - x\|_1 / N$ 或均方差 $\|\Psi s - x\|_2 / N$ 等作为重建质量度量指标.

4 基于先验知识的分段动态采样调度策略

利用压缩感知理论来进行土壤呼吸监测传感网的采样调度, 必然需要面临第 3.1 小节中描述的压缩感知研究和应用中的几个问题. 我们重点研究其中的第二个问题, 即测量矩阵 Φ 的设计.

如 3.3 小节所描述, 我们设计的测量矩阵 $\Phi_{M \times N}$ 的行数 M 表示采样次数, 每一次的采样时刻由该行非零元素的列号 n 表示. 由于土壤呼吸测量仪的每次采样只能测量一个时间点的数据, 因此测量矩阵 Φ 的每一行有一个非零元素, 每一列有不超过一个非零元素. 因此, 测量矩阵的设计包括两个方面: (1) 行数 M , 即采样次数; (2) 各行非零元素的列号 n , 即每次采样的时间点.

本文主要针对测量矩阵设计中的前一个方面, 即采样次数 M 的确定. 对于真实物理世界中的原始数据序列 X , 确定采样次数 M 可等同于确定采样率 M/N . 如果数据序列 X 中元素随时间线性变化, 则通过适当的表达基矩阵 Ψ 的线性变换可以得到良好的稀疏化效果, 因此采用较低的采样率就可以得到良好的重建效果. 然而真实物理世界的的数据序列 X 通常并不是线性变化的, 因此需要提高采样率来采集到丰富的数据变化信息, 以便重建数据序列能够达到需要的整体准确度.

虽然 X 整体不是线性变化的, 但如果对数据序列 X 进行进一步分解, 则可能在 X 的某些部分具有近似线性变化的局部特征. 如果在整个测量过程中采用固定的采样率来设计测量矩阵, 则在近似线性变化部分可能得到较好的重建效果, 但在非线性变化部分的重建效果则比较差; 如果为了提高非线性变化部分的重建效果而提高采样率, 则在线性变化部分的采样数据序列就存在一定的冗余. 因此本文研究动态采样调度策略, 即根据数据序列 X 变化的趋势, 在不同时间区间采用动态变化的采样率来构建测量矩阵: 在线性变化部分降低采样率, 在非线性变化部分提高采样率, 以便在满足准确性要求的前提下, 尽可能的降低采样率. 由于土壤呼吸的测量是耗时耗能的, 因此降低采样率能够降低整个监测系统的能量消耗、延长系统的野外持续工作时间.

研究表明土壤呼吸随时间的变化存在规律性: 白天土壤呼吸比较慢, 而晚上土壤呼吸则相对较快. 因为土壤呼吸的一个主要原因是植物根系的呼吸作用, 所以土壤呼吸的变化受植物生理过程的影响^[21]. 在中午阳光最好的时间植物进行光合作用固定碳, 数小时后碳才输送到根部, 从而通过根际呼吸在夜间释放出来^[21,22], 根际呼吸大约滞后于光合作用 7~12 小时^[23]. 另外中午气温高于土壤温度, 气体压强也大, 抑制土壤 CO_2 的扩散释放, 所以这期间土壤呼吸值较低^[24]. 植物根系呼吸作用的强

弱也与所处地带和季节存在关系. 在夏季植物生长旺季, 土壤呼吸速率在夜间出现高峰值, 夜间土壤呼吸速率大于昼间土壤呼吸速率; 而冬季植物生长缓慢的季节, 夜间土壤呼吸速率略大于昼间土壤呼吸速率, 没有非常明显的高峰和低谷 [25].

由于土壤呼吸随时间的变化以天为周期存在规律性和相似性, 季节引起的植物生长状况和日出日落时间引起的一天内温度变化趋势在相邻几天内差别不大, 可以依据前几天的土壤呼吸数据来估计接下来的土壤呼吸变化趋势. 因此本文提出基于先验知识的动态采样调度策略: 根据对近期实测或重建历史数据的观察和分析, 将一天的采样时间进行分段, 各段按照历史数据曲线变化的情况而采用不同的采样频率: 在数据序列非线性程度高的时间段内, 提高采样率; 在非线性程度低的时间段内, 则降低采样率.

为了对采样时间进行分段, 可以按照历史数据的变化趋势, 进行分段线性拟合和分割, 以最小化各分段内的数据非线性程度为目标, 以便在保证重建质量的前提下最大程度地降低采样率. 但是, 考虑到土壤呼吸监测传感网的实际情况, 各测量节点除了测量外还具有通信功能, 这要求测量节点与其他节点的通信时机需要进行相互协调. 如果按照数据变化趋势进行直接分段的方法会导致同一节点的不同分段长度不尽相同, 并且由于土壤的空间异质性, 不同采样位置对采样时间进行分段的结果很可能不同. 这将使得各个节点的测量很难与节点间的通信相互协调.

因此本文采用将采样时间均匀分段的方法. 一方面, 各分段的大小相同, 从而物理世界原始数据序列 X 也被分割成相等数目元素的子序列 X_i , 对应第 3.3 小节的描述, 设分段后的每个子序列 X_i 的元素数目为 N , 则各段可以采用统一 $N \times N$ 大小的表达基矩阵 Ψ ; 另一方面, 相同的时间分段方法有利于不同节点之间的通信同步, 各节点可以采用相同的通信调度方法, 例如在分段采样周期开始时传输前一周期的测量数据.

Algorithm 1 Segmented dynamic sampling scheduling algorithm based on prior knowledge (SDSS)

Require: X_{P_i}, N, SR_{base}

```

1:  $X_{P_i}$  = 当前节点最近  $P_d$  天采样重建的土壤呼吸数据序列 ( $1 \leq P_i \leq P_d$ ), 每天数据序列包含  $N_d$  个元素;
2:  $X$  = 前  $P_d$  天的数据序列  $X_{P_i}$  按每天对应时间求平均值得到的待分析数据序列;
3: for  $i = 1$  to  $d = \lceil N_d/N \rceil$  do
4:    $X_i$  = 将  $X$  均匀分段为大小为  $N$  的子序列;
5:    $R^2_i$  = 子序列  $X_i$  用最小二乘法进行线性拟合时的决定系数;
6: end for
7:  $R^2_{mean} = \sum_{i=1}^d R^2_i/d$ , 所有子序列  $X_i$  进行线性拟合时决定系数  $R^2_i$  的均值;
8:  $SR_{base}$  = 设定的基准采样率, 代表子序列线性拟合时的决定系数等于  $R^2_{mean}$  时的该分段的采样率;
9: for  $i = 1$  to  $d$  do
10:   $SR_i = SR_{base} + (R^2_{mean} - R^2_i) \beta$ , 计算结果为各分段的采样率;
11:   $\Phi_i$  = 依据采样率  $SR_i$  构造适合当前分段的动态测量矩阵.
12: end for
13: return  $\Phi_i, (1 \leq i \leq d)$ ;

```

在均匀分段的基础上, 本文提出基于先验知识的分段动态采样调度策略如算法 1 所示. 利用该算法得到分段动态测量矩阵 Φ_i 后, 土壤呼吸测量仪将根据时钟处于不同的时间分段时选择对应时段的动态测量矩阵作为采样调度策略进行测量.

在算法 1 中, 采用决定系数 R^2 作为各数据子序列线性程度的评价指标. 决定系数 R^2 通常被用来评价拟合结果的估计值与对应的实际数据之间的拟合程度, 其取值范围为 0~1 之间的数值, 当线性拟合的 R^2 值等于 1 或接近 1 时, 说明数据之间的线性程度非常高, 反之则线性程度较低. 第 10 行中的 β 为增量调节因子, 其值越大, 非线性程度不同的分段间的采样率差越大. 由第 10 行中 SR_i 的计算公式可以看出, 各分段子序列的 R^2_i 越大, 说明该分段的土壤呼吸变化的线性程度越强, 因此计算得到较小的采样率, 反之则得到较大的采样率.

算法 1 中 SR_{base} 为设定的基准采样率, 代表子序列线性拟合时的决定系数等于 R^2_{mean} 时的该分段的采样率, 它影响所有时段的动态采样率, 从而决定着整个土壤呼吸监测传感网系统的数据重建质量. 因此它的取值取决于应用系统对数据重建质量的需求: 当要求得到较低的重建误差时, 应选用较高的基准采样率, 否则选用较低的基准采样率. 在实际应用中, 可在系统运行初期使 SR_{base} 的取值从高到低来进行尝试, 计算不同 SR_{base} 取值时的重建数据误差 [参见 3.3 小节的 (e)], 选择满足重建质量需求的最低基准采样率取值稍大一些的采样率值作为系统正常运行期的 SR_{base} .

算法 1 中 N 为数据序列分段的大小, 是动态采样率的分段粒度. 它既对数据重建质量有影响, 也影响着动态采样调度策略的运算量, 参见 5.5 小节实验结果的分析与讨论.

算法 1 使用分段时前 P_d 天的土壤呼吸数据序列 $X_{P_i} (1 \leq P_i \leq P_d)$, 导致计算结果具有一定的时效性. 因此, 在长期土壤呼吸监测过程中, 土壤呼吸测量仪需要定期 (通常可设定为三天到一周左右) 重复执行上述算法来更新各分段采样率. 在部署初期没有前期数据用于分析土壤呼吸数据的变化趋势时, 采用各分段相同采样率的静态调度策略.

5 模拟实验

如 3.1 小节所述, 压缩感知的重建质量受三个因素的影响: 测量矩阵 Φ 、表达基矩阵 Ψ 、恢复算法 A. 下面分别介绍本文实验中这三方面因素的设计和选择, 以及实验的数据和方案.

5.1 实验数据

为了获得真实物理世界原始数据序列, 我们采用自主研发的土壤呼吸测量仪部署在室外密集测量了 10 天. 如第 2 节所述, 设备每 3 分钟测量一次土壤呼吸碳通量, 然后自动开箱通风 1 分钟, 即每 4 分钟得到一个土壤呼吸碳通量数据. 则 10 天共测量到 3600 个原始土壤呼吸数据作为本文实验所采用的数据集. 根据实验评估和分析的需要, 实验中将整个数据集数据平均分成多个子序列, 每个子序列作为一个实验数据进行采样和重建.

如第 4 节所述, 实验中将采样时间均匀分段, 整个数据序列 X 数据平均分成多个子序列, 每个子序列作为一个单位进行采样和重建. 按照第 3.3 小节的模型, 设每个子序列中的元素数目记为 N , 则子序列的数目 $d = 3600/N$, 将第 i 个子序列记为 $X_i (1 \leq i \leq d)$.

5.2 测量矩阵 Φ 的构造

如第 4 节所述, 测量矩阵的设计包括确定行数 M 、各行非零元素列号这两个方面. 第 4 节提出的分段动态采样调度策略, 解决的都是第一个方面的问题: 确定第 i 个时间分段内对子序列 X_i 的采样率 SR_i , 从而这个时间区间的采样次数 $M = (SR_i \times N)$, 则测量矩阵即为 $(SR_i \times N)$ 行 N 列的矩阵.

对于第二个方面: 确定各行非零元素列号即 M 次采样各自的具体测量时机, 这个问题不属于本文的研究内容, 在实验中选择两种简单但满足 RIP 约束的方案: 周期性采样 PS 和伪随机采样 RS. 周期性采样是指节点按照 N/M 的周期来测量 M 次, 这种测量矩阵记为 Φ_P ; 伪随机采样是 M 次采样的间隔时间按照均匀随机概率分布, 这种测量矩阵记为 Φ_R .

5.3 表达基矩阵 Ψ 和恢复算法 A 的选择

由于物理世界中土壤呼吸的变化是平滑的, 利用相邻采样数据之间的相关性, 原始数据序列 X_i 可以被稀疏表达. 对于稀疏化数据向量的表达基矩阵 Ψ , 本文采用两种方案^[20]:

(1) 如下差分矩阵:

$$M_D = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -\gamma \end{pmatrix}, \quad (9)$$

其中最后一个元素 $\gamma, 0 < \gamma < 1$, 以保证 M_D 是可逆的, 实验中 γ 设为 0.001. 把 X_i 投影到 M_D 上, $s = M_D X_i$ 是一个包含许多 0 元素和小绝对值元素的向量. 如果是这样, 可以认为原始信号 x 在 M_D 域被近似稀疏表达为 $x = M_D^{-1}s$, 因此, 我们使用 M_D^{-1} 作为本文实验所采用的一个表达基, 记作 Ψ_D .

(2) 由于 Haar 小波变换 M_H 可以用来稀疏化平滑数据, 因此也使用 M_H^{-1} 作为本文实验所采用的表达基矩阵, 记作 Ψ_H .

恢复算法采用第 3.1 小节提到的算法 SL0 和 BP(使用 LP). 基于 Wu 的研究比较^[20], 本文实验在表达基用 Ψ_D 时恢复算法 A 采用 SL0¹⁾, 表达基用 Ψ_H 时采用 LP. LP 的 Matlab 代码从 SparseLab²⁾中获得.

5.4 实验方案

本文的实验主要是为了分析第 4 小节所描述的分段动态采样调度策略基于周期性采样和伪随机采样得到的测量矩阵 Φ (Φ_P 和 Φ_R), 在表达基矩阵 Ψ (Ψ_D 和 Ψ_H)、恢复算法 A (SL0 和 LP) 等因素确定时, 对土壤呼吸测量数据的采样率动态变化情况及重建效果分析, 实验方案见算法 2.

如第 2 小节所描述, 设备每小时可采集 15 个数据, 因此实验中子序列的长度 N 的取值范围 {30, 60, 90, 120, 180, 360} 分别对应时间分段周期为 2, 4, 6, 8, 12, 24 小时. 实验中采用平均误差来评价重建效果, 其计算方法如算法 2 第 16 行所示, 其中 R 是随机实验次数, 在实验中对包含随机因素的调度机制, R 设置为 20.

5.5 实验结果及分析

我们首先分析利用动态采样策略计算动态采样率的结果. 在分段动态采样策略中, 采用实验数据

1) SL0. <http://ee.sharif.ir/~SLzero/>

2) SparseLab. <http://sparselab.stanford.edu/>

中前三天的数据, 按照算法 1, 将前三天数据按时间对应关系进行平均, 结果如图 2 所示. 然后进行分段线性拟合, 拟合结果如图 3 所示. 然后求得各分段的决定系数 R^2 , 如图 4 所示. 再按照算法 1 中的第 10 行计算各分段的动态采样率, 结果如图 5 所示, 其中基准采样率 $SR_{\text{base}}=10\%$, 增量调节因子 $\beta=0.2$.

Algorithm 2 Experimental procedure experiment

```

1:  $\Phi\_method = \{\Phi_P, \Phi_R\}; \langle \Psi\_method, A\_method \rangle = \{\langle \Psi_D, SL0 \rangle, \langle \Psi_H, LP \rangle\};$ 
2:  $N = \{30, 60, 90, 120, 180, 360\};$ 
3: for each  $\langle \psi\_method, a\_method \rangle \in \langle \Psi\_method, A\_method \rangle$  do
4:   for each  $\phi\_method \in \Phi\_method$  do
5:     for each  $n \in N$  do
6:       将数据集  $X$  分成大小为  $n$  的子序列  $X_i, (1 \leq i \leq d = \|X\|_0 / n);$ 
7:        $\psi =$  按照  $\psi\_method$  方法构建  $n \times n$  的表达基矩阵;
8:       for  $i = 1$  to  $d$  do
9:          $SR_i =$  按照动态采样调度策略得到  $X_i$  的采样率;
10:         $\phi =$  按照  $\phi\_method$  方法构建的  $\lceil SR_i * n \rceil \times n$  测量矩阵;
11:         $y = \phi X_i;$ 
12:         $s = a\_method(y, \phi, \psi);$ 
13:         $\widehat{X}_i = \Psi s;$ 
14:         $Err_{(\phi\_method, \psi\_method, i)} = \|\widehat{X}_i - X_i\|_1$ 
15:      end for
16:       $ErrAvg_{(\psi\_method, \phi\_method, n)} = (1/R) \sum_{r=1}^R (1/d) \sum_{i=1}^d Err_{(\phi\_method, \psi\_method, i)} / n;$ 
17:    end for
18:  end for
19: end for
20: return  $ErrAvg;$ 

```

图 2 土壤呼吸碳通量数据

Figure 2 Data of soil respiration carbon flux

图 3 数据分段线性拟合结果

Figure 3 Results of segmental linear fitting of data

 图 4 各分段线性拟合的决定系数 R^2

 Figure 4 R^2 of segmental linear fitting

图 5 各分段的动态采样率

Figure 5 Dynamic sampling rates of data subsequences

观察图 4 中的各分段的拟合决定系数 R^2 , 其中比较大的几段是 $N=30$ 时的第 5 段 (8 点到 10 点)、 $N=30$ 时的第 9 段 (16 点到 18 点) 和第 10 段 (18 点到 20 点)、 $N=60$ 时的第 5 段 (16 点到 20 点), 图 3 中也可以看出, 对应时间段的土壤呼吸碳通量数据的时序变化的确线性程度较好. 另外, 观察图 5 中这几个分段对应的动态采样率, 都在 5% 以下, 比基准采样率降低了一半左右. 图 5 中各分段的动态采样率最高的是 $N=60$ 时的第 4 段 (12 点到 16 点), 对应的图 4 中该分段的 R^2 为 0.14 左右. 图 4 中 R^2 最小的是 $N=360$ 时的整个分段 (0 点到 24 点), 但这段并没有对应很高的采样率, 而是采用选择基准采样率 (见图 5), 因为一天只有一个分段, 没有其他线性化程度好的分段来进行平衡, 此时即等同于没有分段的固定采样率时的情形.

图 6 展示了本文提出的分段动态采样调度策略 (SDSS) 与固定采样率的静态采样策略 (CSS) 两种情形下的重建整体平均误差. 静态采样策略的固定采样率和动态采样策略的基准采样率都分别选择

图 6 动态采样率与固定采样率的重建性能比较

Figure 6 Reconstructive performance comparison of dynamic and static sampling rates. (a) $\Phi_P\Phi_D$; (b) $\Phi_R\Phi_D$; (c) $\Phi_P\Phi_H$; (d) $\Phi_R\Phi_H$

10%(10CSS、10SDSS)、20%(20CSS、20SDSS)、30%(30CSS、30SDSS). 由算法 1 中第 10 行 SR_i 的计算公式可以看出, 采用分段动态采样时, 各分段的动态采样率 SR_i 的平均值与基准采样率 SR_{base} 相等. 因此图 6 中的 10CSS 和 10SDSS 两种情况的平均采样率都是 10%, 即进行采样所消耗能量相同.

从图 6 中可以看出, 在测量矩阵 Φ 和 Ψ 分别选取不同构建方法时, 相同平均采样率的分段动态采样调度策略都比静态采样调度策略得到的重建误差更小, 即本文提出的动态采样策略对土壤呼吸监测传感网的采样调度问题具有良好的应用效果. 另外, 从图中可以看出, 无论测量矩阵构造方法选择 Φ_P 还是 Φ_R , 表达基矩阵使用 Ψ_D 比 Ψ_H 的重建误差要小, 说明 Ψ_D 更适合用于土壤呼吸碳通量数据的处理. 尤其是在子序列的元素数目比较少 (例如 30) 时, 使用 Ψ_H 会导致重建误差偏大. 总体来说, 使用 Ψ_H 时的重建误差大约是使用 Ψ_D 的两三倍甚至更多. 在表达基矩阵 Ψ 确定的情况下, 测量矩阵选择 Φ_P 和 Φ_R 的差别不是很大. 这也正是本文重点研究动态采样调度方法而没有把测量矩阵构造方法当做研究内容的原因之一.

分段子序列长度 N 较大时, 数据间的相关性使得子序列中包含的冗余数据比例较大, 则子序列的稀疏化效果较好, 因此较小的采样率即可较好重建; 相反子序列长度 N 较小时, 则需要较大的采样率来满足重建需要. 因此图 6 中的各个固定采样策略 (CSS) 曲线在采样率较低 (10%) 时, 较小子序列长度 ($N=30$) 的分段方法引起的误差较大; 而采样率较高 (20%, 30%) 时, N 取不同值时的重建误差的差别不是很大. 分段动态采样调整策略在一定程度上可以缓解这种问题, 因为它对变化较复杂的分段提高了采样率, 则能够改善重建效果, 因此图 6 中的各条动态采样策略 (SDSS) 曲线基本上都比较平缓. 由于利用压缩感知重建时求算法 (LP, SL0 等) 的计算复杂度较高, 所以计算量主要受分段大小的影响, 而不是分段数量. 因此虽然 N 较小时分段数量多, 但总体计算量却相对较小. 所以, 重建误差并非

图 7 动态采样率与固定采样率的分段重建性能比较

Figure 7 Segmental reconstructive performance comparison of dynamic and static sampling rates

实际应用中确定分段粗细的唯一依据, 在重建误差差别不大的情况下, 应该尽量选择子序列长度较小的分段方法. 从图 6 中可以看出, 对于本文提出的动态采样调度策略, 在基准采样为 10% 的情况下, 在应用时可选用 $N=120$ 进行分段; 在基准采样率为 20% 和 30% 时, 可采用 $N=30$.

图 6 展示的是动态采样调度与静态采样调度在全局的重建效果差别. 例如在图 6(a) 中, 分段子序列长度 N 为 60、静态采样率和动态的基准采样率都为 10% 时, 二者的全局平均重建误差分别为 0.15 和 0.08 左右. 为了分析一下差别产生的原因, 计算了两种情况下的各分段的数据子序列的平均重建误差, 如图 7 所示.

从图 5 中可以看出, 按 $N=60$ 进行分段时, 在第 1, 4, 6 个时间段内, 算法 1 分配的动态采样率高于基准采样率 10%, 即这几段时间对应的土壤呼吸碳通量数据子序列的非线性程度高, 需要提高采样率才能得到较好的重建效果. 从图 7 中可以看出, 对应的第 1, 4, 6 个时间段内, 使用固定采样率得到的分段重建误差都远远高于使用动态采样率的情形. 在另外的第 2, 3, 5 个时间段内, 图 5 中分配了低于基准采样率的动态采样率, 说明这几段数据序列的线性程度较高, 不需要很高采样率就可以达到较好的重建效果. 即使静态采样策略的高采样率得到了高于动态采样率的重建质量, 但提高并不大, 图 7 中的第 2, 3, 5 时间段固定采样率比动态采样率的重建误差稍小. 从图 7 中的两种采样率对应的分段误差来看, 固定采样率的静态采样调度策略在各个时间分段得到的重建误差相差较大, 这是因为各个时间分段的数据的线性化程度不同, 而采用无差别的采样率使得到的重建效果存在较大差别; 而动态采样调度策略在各个分段得到的重建误差相差不大, 这是因为在各分段采用不同的采样率, 平衡了数据变化带来的重建准确度差别.

6 结论

本文针对土壤呼吸监测传感网的节能采样问题, 在压缩感知理论的基础上, 提出了一种基于先验知识的分段动态采样策略 SDSS, 并通过实验分析了动态采样策略相比固定采样率的采样策略更能适应监测对象的动态变化, 从而达到减少采样率节省能耗的目的. 实验结果表明动态采样调度策略能够适应土壤呼吸物理过程在不同时间段的不同采样率需求, 达到较为平均的分段采样误差和总体较优的全局采样效果. 虽然动态采样策略需要土壤呼吸仪额外执行采样率更新算法, 产生一定的能耗, 但由于土壤呼吸的测量是一个耗时耗能的过程, 相对来说, 减少采样次数带来的能量节省远远超过了更新采样率算法引起的额外能量消耗.

由于土壤呼吸的主要包括根系呼吸、土壤微生物呼吸和土壤动物的异养呼吸, 这些呼吸活动都受到土壤温度和湿度的影响, 因此, 不论是空间梯度还是时间水平, 影响土壤呼吸速率的主要环境因子是土壤湿度和温度^[26]. 从第 2 节的描述可以知道, 土壤呼吸的测量需要采用动态开放箱法等方法, 包括箱体移动、抽气测量等过程, 耗时耗能; 但土壤温湿度的测量相对容易, 所以在后续的研究中, 我们拟分析温湿度与土壤呼吸这几类感知量之间的关联性, 对土壤碳通量动态采样策略优化和调整, 以进一步降低节点的能耗.

参考文献

- 1 Zhang M, Wen X F, Yu G, et al. Effects of CO₂ storage flux on carbon budget of forest ecosystem. *Chinese J Applied Ecology*, 2010, 21: 1201–1209 [张弥, 温学发, 于贵瑞, 等. 二氧化碳储存通量对森林生态系统碳收支的影响. *应用生态学报*, 2010, 21: 1201–1209]
- 2 Pu Y F. Sensitivity test of the distribution of Indian ocean surface CO₂ partial pressure on the physical and biochemical processes by a two-dimensional ocean carbon cycle model. *Climatic Environmental Res*, 2009, 14: 301–308 [浦一芬. 印度洋二维碳循环模式中表层 CO₂ 分压分布对物理和生化过程的敏感性试验. *气候与环境研究*, 2009, 14: 301–308]
- 3 Raich J W, Schlesinger W H. The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus B*, 1992, 44: 81–99
- 4 Halgamuge M N, Zukerman M, Ramamohanarao K, et al. An estimation of sensor energy consumption. *Prog Electromagnetics Res B*, 2009, 12: 259–295
- 5 LI-COR Corporation. LI-8100A Automated Soil CO₂ Flux System, http://www.licor.com/env/products/soil_flux/
- 6 Powers J, Read J, Denslow J, et al. Estimating soil carbon fluxes following land-cover change: a test of some critical assumptions for a region in Costa Rica. *Global Change Biology*, 2004, 10: 170–181
- 7 Mishra U, Torn M, Masanet E, et al. Improving regional soil carbon inventories: combining the IPCC carbon inventory method with regression kriging. *Geoderma*, 2012, 189: 288–295
- 8 Donoho D L. Compressed sensing. *IEEE T Inf Theory*, 2006, 52: 1289–1306
- 9 Candès E J, Wakin M B. An Introduction To Compressive Sampling. *Signal Process Mag*, 2008, 3: 21–30
- 10 Baron D, Wakin M, Duarte M, Sarvotham S, and Baraniuk R. Distributed compressed sensing, *IEEE T Inf Theory*, 2006, 52: 5406–5425
- 11 Candès E J, Tao T. Decoding by linear programming. *IEEE T Inf Theory*, 2006, 51: 4203–4215
- 12 Tian Z, Giannakis G. Compressed sensing for wideband cognitive radios. In: *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2007. 1357–1360
- 13 Quer G, Masiero R, Munaretto D, et al. On the interplay between routing and signal representation for compressive sensing in wireless sensor networks. In: *Proceedings of Information Theory and Applications Workshop (ITA 2009)*, 2009. 206–215
- 14 Luo C, Wu F, Chen C W, et al. Compressive data gathering for large-scale wireless sensor networks. In: *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking*, 2009. 145–156
- 15 Donoho D L, Elad M, Temlyakov V. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE T Inf Theory*, 2006, 52: 6–18
- 16 Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed l0 norm. *IEEE T Signal Process*, 2009, 57: 289–301
- 17 Donoho D L. For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 2006, 59: 797–829
- 18 Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans Inf Theory*, 2007, 53: 4655–4666
- 19 Needell D, Vershynin R. Signal recovery from inaccurate and incomplete measurements via regularized orthogonal matching pursuit. *IEEE J Sel Topics Signal Process*, 2010, 4: 310–316

- 20 Wu X, Liu M. In-situ soil moisture sensing: measurement scheduling and estimation using compressive sensing. In: Proceedings of the 11th International Conference on Information Processing in Sensor Networks (IPSN), 2012. 1–12
- 21 Liu Q, Edwards N T, Post W M, et al. Temperature-independent diel variation in soil respiration observed from a temperature deciduous forest. *Global Change Biology*, 2006, 12: 2136–2145
- 22 Dilkes N B, Jones D L, Farrar J. Temporal dynamics of carbon partitioning and rhizodeposition in wheat plant. *Plant Physiology*, 2004, 134: 706–715
- 23 Tang J, Baldocchi D D, Xu L. Tree photosynthesis modulates soil respiration on a diurnal time scale. *Global Change Ecology*, 2005, 11: 1298–1304
- 24 Cao J, Song L, Jiang G, et al. Diurnal dynamics of soil respiration and carbon stable isotope in Lunan stone forest, Yunnan province. *Carsologica sinica*, 2005, 24: 23–27 [曹建华, 宋林华, 姜光辉等. 路南石林地区土壤呼吸及碳稳定同位素日动态特征. *中国岩溶*, 2005, 24: 23–27]
- 25 Feng W T, Zou X M, Sha L Q, et al. Comparisons between seasonal and diurnal patterns of soil respiration in a montane evergreen broad-leaved forest of Ailao mountains, China. *J Plant Ecology*, 2008, 32: 31–39 [冯文婷, 邹晓明, 沙丽清, 等. 哀牢山中山湿性常绿阔叶林土壤呼吸季节和昼夜变化特征及影响因子比较. *植物生态学报*, 2008, 32: 31–39]
- 26 Raich J W, Wtufekcioglu A. Vegetation and soil respiration: correlations and controls. *Biogeochemistry*, 2000, 48: 71–90

Dynamic sampling scheduling policy for soil respiration monitoring sensor networks based on compressive sensing

WANG GuoYing^{1,3}, JIANG YuJia¹, MO LuFeng^{1,3*}, SUN YanFeng² & ZHOU GuoMo⁴

1 *School of Information Engineering, Zhejiang A&F University, Lin'an 311300, China;*

2 *School of Computer, Beijing University of Technology, Beijing 100124, China;*

3 *School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China;*

4 *Zhejiang Provincial Key Laboratory of Carbon Cycling in Forest Ecosystems and Carbon Sequestration, Zhejiang A&F University, Lin'an 311300, China*

*E-mail: molufeng@gmail.com

Abstract Sensor networks for soil respiration monitoring are usually deployed in the fields where electric or manual intervention cannot be accessed directly, and the measurement of soil respiration is complex and highly energy-consuming. Therefore, we hope to minimize the number of soil respiration measurements on the premise of reconstruction accuracy. Sampling scheduling can be realized using compressive sensing theory on the basis of temporal correlation of the physical process of soil respiration. Here we propose a segmental dynamic sampling scheduling policy based on compressive sensing. Using a prior knowledge obtained by means of analysis on the earlier measurement data, the data serial in measurement period is partitioned and linear fitted. Then the dynamic sampling rate of each segment is determined according to the linear degree of data in the segment, based on which the measurement matrix is constructed for the sampling and reconstructed for compressive sensing process using the soil respiration measuring instrument. The experimental result shows that the proposed segmental dynamic sampling policy can lead to better reconstructive quality than static sampling policy of the same average sampling rate. That is to say, the proposed dynamic sampling policy needs smaller sampling rate if the reconstructive error threshold is given. The reduction of sampling rate can save more power although the calculation of dynamic sampling rate may consume some power. The proposed segmental dynamic sampling policy based on compressive sensing can also be referenced and potentially used by similar applications for the sampling scheduling and power-saving issues, although it is experimented and analyzed in the soil respiration monitoring sensor networks.

Keywords sampling, scheduling, compressed sensing, wireless sensor networks, soil respiration

WANG GuoYing was born in 1977. He received the B.E. degree from Beijing Jiaotong University in 1999, and the M.E. degree from Guangxi University in 2004. He is currently a lecturer at Zhejiang A&F University and a Ph.D. student at Xi'an Jiaotong University. His research interest includes computer networking and wireless sensor networks.

MO LuFeng was born in 1979. He received the M.S. degree from the Peking University in 2004. Currently, he is an associate professor at Zhejiang A&F University and a Ph.D. student at Xi'an Jiaotong University. His research interest includes computer software, sensor networks, pervasive computing, and information security.

ZHOU GuoMo was born in 1961. He received the Ph.D. degree from Zhejiang University in 2006. Currently, he is a professor at Zhejiang A&F University. His research interests include forests carbon monitoring, climate change adaptation and carbon management, and sustainable forest management. Dr. Zhou is a member of Chinese Forestry Association, as well as Vice President of forest management branch, and director of the academic committee of Joint Laboratory on Internet of Things and Global Climate Change.

academic committee of Joint Laboratory on Internet of Things and Global Climate Change.