

信息物理系统基于空间相关性的数据传输*

陈悦¹, 罗俊海²

(1. 乐山职业技术学院, 四川 乐山 614000; 2. 电子科技大学 电子工程学院, 成都 611731)

摘要: 提出一种基于节点间空间相关性的数据传输策略 SCDF, 根据空间相关性, 各个节点间感知自身到目的节点的距离, 并依此来计算传输概率值。节点传输概率即感知节点把消息传递给目的节点的可能性, 它是消息传递时选择下一跳的重要依据。仿真模拟实验表明, 与现有的直接传递 (DD) 算法和 Epidemic 算法相比, SCDF 能以较低的数据传输开销和传输延迟获得较高的数据传输率。

关键词: 信息物理系统; 空间相关性; 数据传输; 物联网

中图分类号: TP393

文献标志码: A

文章编号: 1001-3695(2012)04-1512-02

doi:10.3969/j.issn.1001-3695.2012.04.086

Spatial correlation data forwarding scheme for cyber-physical systems

CHEN Yue¹, LUO Jun-hai²

(1. Leshan Vocational & Technical College, Leshan Sichuan 614000, China; 2. School of Electronic Engineering, University of Electronic Science & Technology of China, Chengdu 611731, China)

Abstract: This paper proposed a novel data gathering and forwarding scheme named spatial correlation data forwarding (SCDF). SCDF introduced a small overhead to gain the relative distance from a node to a sink node and then to calculate the node forwarding probability, which gave a guidance to message transmission based on spatial correlation between nodes. Simulation results show that SCDF does not only achieve a relatively low data forwarding energy consumption, but also get the higher message forwarding ratio with lower transmission overhead and data forwarding delay than direct delivery (DD) and Epidemic algorithms.

Key words: cyber physical systems; spatial correlation; data forwarding; Internet of things

信息物理系统 (cyber-physical systems, CPS) 是一个在环境感知的基础上, 深度融合了计算、通信和控制能力的可控、可信、可扩展的网络化物理设备系统, 它通过计算进程和物理进程相互影响的反馈循环实现深度融合和实时交互来增加或扩展新的功能, 以安全、可靠、高效和实时的方式监测或者控制物理实体^[1,2]。CPS 的每个感知节点感知收集周围物理环境信息, 然后通过计算获取一条最佳的传输路径, 将数据成功传送到目的终端; 另一方面, 终端也可通过一定的路由方式, 将控制信息传输给每个感知节点。CPS 网络中有大量的感知节点进行数据信息的采集和生产, 这些节点或是固定或是移动, 同时它们具有自我判断、自我控制以及自我调节的能力, 因此在 CPS 中, 感知节点的随机部署、节点分布密度低, 以及感知节点的随机移动和能量耗尽等情况都可能导致网络有别于传统的全连通网络^[3,4]。这种网络环境中, 传统的通信模式无法有效运行, 因此本文提出一种基于节点间空间相关性的数据传输策略 SCDF。

1 系统模型

1.1 系统模型

节点移动模型描述了节点的移动模式, 包括位置、速度和方向等特征的变化。CPS 移动模型是以刻画节点相遇特征为核心的, 表征了节点的随机性、时间相关性、空间相关性和地理

受限性等特性, 这是因为在 CPS 中的数据传输依赖于节点移动带来的相遇机会, 而节点的相遇概率、相遇时间分布和相遇空间分布是由节点的移动模型决定的。

1.2 模型假设

本文假设初始状态时, N 个感知节点随机分布在一个 $M \times M$ 的二维正方形区域 A 内, 并且 (为简单起见) 假设网络中唯一的目点也部署在区域 A 内且固定不动。所有感知节点和目的点的通信半径均为 R 。此外, 假设该 CPS 具有如下性质:

a) CPS 中所有节点的运动规律服从随机路点 (random waypoint, RWP) 移动模型^[5]。RWP 描述为: 开始时节点在运动空间随机选择起始点, 并在当前位置停留一段随机时间 $T_p \in [T_{\min}, T_{\max}]$, T_p 表示移动的时间。如果 $T_p = 0$, 表示连续移动, 然后在运动空间内随机选取一个位置作为目标位置, 并以随机选定的速率 $V \in [V_{\min}, V_{\max}]$, 匀速向该目标位置直线移动。节点到达目的地后, 再随机暂停一段时间 T_p , 然后重复上述过程。该模型中, 每个节点的速率和移动方向的选择与其他节点相互独立, 最大速率 $V_{\max}$ 和停留时间 T_p 是两个关键参数。如果 $V_{\max}$ 小而 T_p 长, 则网络的拓扑相对稳定; 如果 $V_{\max}$ 大而 T_p 短, 拓扑则有较高动态性。

b) 通过全球定位系统 GPS, 各个感知节点可获知任意时刻自己的当前位置。

收稿日期: 2011-09-01; 修回日期: 2011-10-09 基金项目: 国家自然科学基金资助项目 (61001086); 中央高校基本科研业务费资助项目 (ZYGX2011X004)

作者简介: 陈悦 (1980-), 女, 四川成都人, 助教, 硕士, 主要研究方向为信息处理与英语翻译 (chenyue0832@163.com); 罗俊海 (1980-), 男, 四川成都人, 副教授, 博士 (后), ACM 和 CCF 会员, 主要研究方向为新型无线通信体系和 CPS。

2 SCDF 策略

在一个特定的监测区域通常会布置很多 CPS 节点,这些节点感知和传输的数据具有空间相关性,即物理位置上相邻的节点采集到的监测数据往往比较相似或存在某种函数关系。假设(如图1所示),在检测区域空间位置为 $s = (\langle x, y \rangle) = (\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \dots, \langle x_i, y_i \rangle)$ 的 N 个节点, CPS 节点通过对 $s_1, s_2, \dots, s_k$ 监控区域进行了 k 次独立的采样(如区域温湿度),测量取得了样本 $(s_1, s_2, \dots, s_{ki}, \hat{f}_k)$, $k = 1, 2, \dots$ 后,可得测量样本的模型为

$$\begin{aligned} \hat{f}(x_i, y_i) &= \sum_i w_i h_i(x_i, y_i) = f(x_i, y_i) + \varepsilon_i \\ \hat{f}(s_i) &= \sum_i w_i h_i(s_i) \approx f(s_i) \end{aligned} \quad (1)$$

其中: $w = (w_1, w_2, \dots, w_k)^T$ 为测量函数系数。根据测量的随机误差,本文利用均方根误差(RMSE),可得

$$\begin{aligned} w^* &= \arg \min_w \sqrt{\frac{1}{m} \sum_{k=1}^m (f_k - \hat{f}_k)^2} = \\ &\arg \min_w \sqrt{\frac{1}{m} \sum_{k=1}^m (f_k - \sum_i w_i h_i(x_k, y_k))^2} \end{aligned} \quad (2)$$

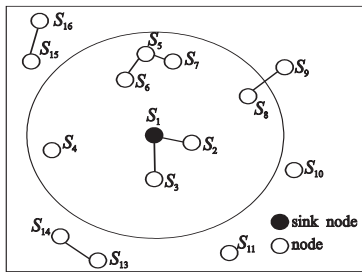


图1 区域节点分布示意图

基于线性回归执行最优化^[6,7],定义矩阵 H ,即 $h(x, y) =$

$$\begin{pmatrix} h_1(x_1, y_1) & h_2(x_1, y_1) & h_k(x_1, y_1) \\ h_1(x_2, y_2) & h_2(x_2, y_2) & h_k(x_2, y_2) \\ \vdots & \vdots & \vdots \\ h_1(x_i, y_i) & h_2(x_i, y_i) & h_k(x_i, y_i) \end{pmatrix}, \text{根据矩阵论可以得到}$$

$$\begin{aligned} w &= h(s)^T h(s)^{-1} h(s)^T f = \\ &h(x, y)^T h(x, y)^{-1} h(x, y)^T f \end{aligned} \quad (3)$$

其中:假设 $\forall_n i \neq j, \varepsilon_i$ 服从均值为零、方差为 σ^2 的独立、正态分布,即

$$E\{\varepsilon_i\} = 0, \text{var}\{\varepsilon_i\} = \sigma^2 \quad i = 1, 2, \dots$$

对于位于 $s_i(x_i, y_i)$ 的节点,有

$$\rho(i, j) = K_\theta(d(i, j)) = \frac{E[f_i f_j]}{\sigma^2} \quad (4)$$

其中: $d(i, j) = \|s_i - s_j\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ 表示节点在 s_i 和 s_j 位置之间的欧式距离,联合高斯随机变量(joint Gaussian random variable) $K_\theta(\cdot)$ 为相关函数, s_i 和 s_j 分别为节点 n_i 和 n_j 的坐标 $s_i(x_i, y_i)$ 和 $s_j(x_j, y_j)$,根据节点空间相关性,可以得出任意节点的相关系数,即

$$\rho(i, j) = K_\theta(d(i, j)) = \begin{cases} 0 & d = \infty \\ \frac{E[S_i S_j]}{\sigma_s^2} & \text{else} \\ 1 & d = 0 \end{cases} \quad (5)$$

其中: $K_\theta(d) = e^{(-d/\theta_1)\theta_2} = K_\theta(d(i, j)) = e^{(-d(i, j)/\theta_1)\theta_2}$, $\theta_1 > 0$, $\theta_2 \in (0, 2]$ 。当参数 $\theta_2 = 1$ 时,模型变成指数形式。在节点 n_i 与 n_j 之间,参数 θ_1 控制着节点间距离 $d(i, j)$ 和相关系数 $\rho(i, j)$ 。由于 CPS 节点观测的空间相关性依赖于它们之间的距离,所以 θ_1 是很重要的参数^[8,9]。由式(5)可知, $\rho(i, j)$ 表示节点间空间相关性程度。当节点 i 在一段时间内多次接近目的点,或者始终在目的点附近徘徊,该节点与目的节点间空间相关性的值将变大,表明它有更将可能性将消息传递给目的节点;若节点始终活跃在远离目的点的区域, $\rho(i, j)$ 值将非常小甚至趋近于 0,同时表明该节点数据传输的可能性小。

在 CPS 环境中,根据任意节点间位置空间相关性,各个节点间感知自身到目的节点的距离,并依此来计算相关性系数,得到节点间传输概率值。节点传输概率即感知节点把消息传递给目的节点的可能性,它是消息传递时选择下一跳的重要依据。

3 实验仿真分析

3.1 实验参数

模拟实现了 SCDF、直接传递算法(direct delivery, DD)和 Epidemic 算法^[10],分别按照平均传输成功率、平均传输延迟和网络寿命的仿真性能指标进行分析比较,参数设置如表1所示。

表1 实验参数

parameter	default value	parameter	default value
network size	400 m × 400 m	pause time/s	0 ~ 40
number of sensor node	50	size of each message/bit	200
transmission radius/m	3	message generation rate	0.02/s
speed of sensor node V/m/s	3	position of sink node	(200, 200)

3.2 算法分析比较

如表2所示,将数据传输算法(SCDF)、DD 和 Epidemic 算法分别应用于实验环境,就平均传输成功率、数据传输开销以及平均延迟三个性能指标进行多次仿真实验分析。

表2 算法性能对比分析

性能指标	SCDF	DD	Epidemic
average forwarding ratio/%	92.8	63.6	70.3
average delay/s	402.9	1 582.6	1 500.1
average copies of each message	7.1	1.0	8.9

从表2可以看出,在仿真环境参数下,SCDF 的性能最好,具有最高的数据传输成功率以及最低的消息平均传输延迟,而且对网络平均传输开销控制得比较好。由于 DD 转发数据策略,因此副本数始终为 1,具有非常低的平均传输开销,但由于节点间相遇的概率比较小,所以具有较大的消息延迟和较低的消息传输成功率。而 Epidemic 的消息由于能被不同节点传输和转发,具有较高的传输成功率,因此 Epidemic 的总体传输成功率比直接传递要大些,而传输延迟要小些。

如表3所示,研究三种算法的 CPS 的网络寿命(TTL),并从总体方面对算法进行分析。假设每个节点的初始能量相同,每发送或接收特定长度消息的能量消耗采用文献[5]的能量模型计算,并假设当网络中一半以上数量的节点能量耗尽时,系统失效。

表3 算法性能 TTL 对比分析

	SCDF	DD	Epidemic
CPS TTL/days	120	1200	95

从表3可以看到,DD 由于节点消息递交的概率比较小,所以节点的能量消耗极小,TTL 最长;Epidemic 算法只有偶尔将消息递交,因此 TTL 比较长;SCDF 的 TTL 比 Epidemic 稍长,消耗能量很少。可见,SCDF 策略能够更好地平衡网络开销与消息提交成功率,因此获得了较理想的 TTL。

(下转第 1517 页)

方法在诊断精度上略优于现有的联立方程组求解方法,同时在算法复杂度和耗时上明显优于后者,论证了该方法的有效性。但是,该方法的误检率仍然偏高,使得网络维护人员耗费大量代价于实际为正常链路的“故障链路”上。因此,在保持甚至提高检测精度的基础上,研究怎样降低故障链路诊断算法的误检率将是今后工作的重点。

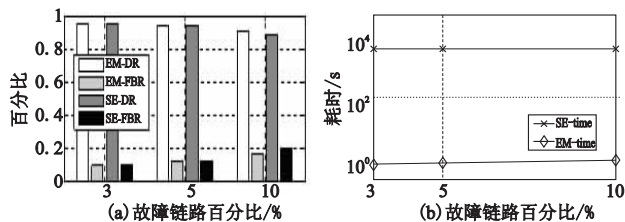


图7 算法性能比较(节点数300, 时隙数30)

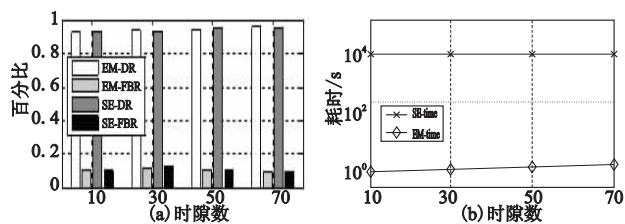


图8 算法性能比较(节点数300, 故障链路比例5%)

参考文献:

- [1] COATES M, HERO A, NOWAK R, *et al.* Internet tomography[J]. *IEEE Signal Process Magazine*, 2002, 19(3): 47-65.
- [2] CASTRO R, COATES M, LIANG Gang, *et al.* Network tomography: recent developments[J]. *Statistical Science*, 2004, 19(3): 499-517.
- [3] PADMANABHAN V N, QIU Li-li, WANG H J. Server-based Inference of internet performance[C]//Proc of IEEE INFOCOM. 2003.
- [4] DUFFIELD N G, PRESTI F L, PAXSON V, *et al.* Network loss tomography using striped unicast probes[J]. *IEEE/ACM Trans on Networking*, 2006, 14(4): 697-710.
- [5] DUFFIELD N G, HOROWITZ J, TOWSLEY D. Multicast-based loss inference with missing data[J]. *IEEE Journal on Selected Areas in Communications*, 2002, 20(4): 700-713.
- [6] GHITA D, ARGYRAKI K, THIRAN P. Network tomography on correlated links[C]//Proc of the 10th Annual Conference on Internet Measurement. 2010: 225-238.
- [7] NGUYEN H X, THIRAN P. Network loss inference with second order statistics of end-to-end flows[C]//Proc of the 7th ACM SIGCOMM Conference on Internet Measurement. New York: ACM, 2007: 227-240.
- [8] GHITA D, NGUYEN H X, KURANT M, *et al.* Netscope: practical network loss tomography[C]//Proc of IEEE INFOCOM. 2010: 1-9.
- [9] 赵佐, 蔡晓东. 基于简单网络断层扫描的失效链路定位研究[J]. *计算机科学*, 2010, 37(1): 108-117.
- [10] 赵佐, 蔡晓东. 基于端到端测量的链路状态概率快速推断方法[J]. *华南理工大学学报: 自然科学版*, 2010, 38(11): 110-115.
- [11] BU Tian, DUFFIELD N, PRESTI F L, *et al.* Network tomography on general topologies[C]//Proc of ACM SIGMETRICS. New York: ACM, 2010: 21-30.
- [12] JI Chuan-yi, ELWALID A. Measurement-based network monitoring and inference: scalability and missing information[J]. *IEEE Journal on Selected Areas in Communications*, 2002, 20(4): 714-725.
- [13] MAO Yong-yi, BANNIHASHEMI A H. Decoding low-density parity-check codes with probabilistic schedule[C]//Proc of IEEE Pacific RIM Conference on Communications, Computers, and Signal Processing. 2010: 119-123.
- [14] GUO Dong, WANG Xiao-dong. Bayesian inference of network loss and delay characteristics with applications to TCP performance prediction[J]. *IEEE Trans on Signal Processing*, 2003, 51(80): 2205-2218.
- [3] 罗俊海, 周应宾, 邓宵博. 物联网网关系统设计[J]. *电信科学*, 2011, 27(2): 105-110.
- [4] BENVENISTE A. Loosely time-triggered architectures for cyber-physical systems[C]//Proc of Conference on Design, Automation and Test in Europe. Leuven, Belgium: European Design and Automation Association, 2010: 3-8.
- [5] 许富龙, 刘明, 龚海刚, 等. 延迟容忍传感器网络基于相对距离的数据传输[J]. *软件学报*, 2010, 21(3): 490-504.
- [6] GUESTIN C, BODIK P, THIBAU R, *et al.* Distributed regression: an efficient framework for modeling sensor network data[C]//Proc of the 3rd International Symposium on Information Processing in Sensor Networks. New York: ACM Press, 2004: 846-851.
- [7] XU Ying-qi, LEE W C. Exploring spatial correlation for link quality estimation in wireless sensor networks[C]//Proc of the 4th Annual IEEE International Conference on Pervasive Computing and Communications. Washington DC: IEEE Computer Society, 2006: 200-211.
- [8] VURAN M C, AKYILDIZ I F. Spatial correlation based collaborative medium access control in wireless sensor networks[J]. *IEEE/ACM Trans on Networking*, 2006, 14(2): 316-329.
- [9] 潘立强, 李建中, 骆吉洲. 传感器网络中一种基于时空相关性的缺失值估计计算[J]. *计算机学报*, 2010, 33(1): 1-11.
- [10] VAHDAT A, BECKER D. Epidemic routing for partially connected Ad-hoc networks, Technical Report CS-200006[R]. Durham, NC: Duke University, 2000.

(上接第1513页)

4 结束语

由于CPS有别于传统系统, CPS各节点既可以作为数据的终端节点,也可以是网络的路由节点;CPS具有无线链路动态特性、时变特性、干扰不规则性和丢失特性;CPS节点移动性带来的链路不稳定、节点能量的有限性、占空比较低、节点之间的连通性是间歇性。同时, CPS节点均具有连网能力, 可多层次多规模连网, 具有自我判断、自我控制以及自我调节的能力, 能够动态重组和重识别。本文提出的基于节点间空间相关性的数据传输策略, 仿真模拟实验表明, 基于节点间空间相关性的方法, 能以较低的数据传输开销和传输延迟获得较高的数据传输成功率以及较长的网络寿命。

参考文献:

- [1] ZIMMER T C, BHAT B, MUELLER F, *et al.* Time-based intrusion detection in cyber-physical systems[C]//Proc of the 1st ACM/IEEE International Conference on Cyber-Physical Systems. New York: ACM Press, 2010: 66-72.
- [2] PAROLINI L, TOLIAZ N, SINOPOL B, *et al.* A cyber-physical systems approach to energy management in data centers[C]//Proc of the 1st ACM/IEEE International Conference on Cyber-Physical Systems. New York: ACM Press, 2010: 168-177.