

# 信息物理社会融合系统： 一种以数据为中心的框架

翟书颖<sup>1</sup>, 郭斌<sup>2</sup>, 李茹<sup>1</sup>, 王庭良<sup>1</sup>, 於志文<sup>2</sup>, 周兴社<sup>2</sup>

1. 西北工业大学明德学院, 陕西 西安 710129

2. 西北工业大学计算机学院, 陕西 西安 710129

## 摘要

信息物理社会融合系统连接信息空间、物理空间和人类社会, 成为一种新的人在回路的计算系统。以数据为中心刻画信息物理社会融合系统的新特征, 包括数据时空交织性、跨空间数据融合及人机智能协作性。构建了一个统一的系统框架, 由资源管理、协作感知、数据预处理、数据融合、隐私保护等部分组成, 并就其在城市管理、公共安全、智能交通、商业智能等重要领域的应用和未来研究挑战进行探讨。

## 关键词

信息物理社会融合系统; 人在回路; 数据驱动架构; 跨空间协同感知

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2017063

## *Cyber-physical-social system: a data-centric framework*

ZHAI Shuying<sup>1</sup>, GUO Bin<sup>2</sup>, LI Ru<sup>1</sup>, WANG Tingliang<sup>1</sup>, YU Zhiwen<sup>2</sup>, ZHOU Xingshe<sup>2</sup>

1. Northwestern Polytechnical University Ming De College, Xi'an 710129, China

2. School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China

## *Abstract*

Cyber-physical-social system(CPSS) connects the cyber space, the physical space, and the social space, which has become a promising human-in-the-loop computing system. The unique features of CPSS from the data-centric level were characterized, including the interweaving of spatio-temporal data, the fusion of cross-space data and the collaboration of human and machine intelligence. Furthermore, a general framework for data-centric CPSS was built, with the underpinnings such as resource management, collaborative sensing, data pre-processing, data fusion, and privacy protection. Finally, the major application areas of data-centric CPSS (e.g., civic management, public safety, intelligent transportation, and business intelligence) and its future research directions were presented.

## *Key words*

cyber-physical-social system, human-in-the-loop, data-driven framework, cross-space collaborative sensing

## 1 引言

信息物理融合系统(cyber-physical system, CPS)能将物理世界和虚拟世界通过计算和物理过程连接起来。目前,在信息物理融合系统中除了包括信息和物理特征,一些与社会相关的要素不断被加入进来,这样的系统被称为信息物理社会融合系统(cyber-physical-social system, CPSS)<sup>[1]</sup>。

信息物理社会融合系统的定义为:在传统CPS关注计算、通信和控制融合的基础上,考虑人类和社会特征,形成“人在回路(human-in-the-loop)”<sup>[2]</sup>的感知与计算系统;信息物理社会融合系统中“数据”成为连接信息、物理和社会空间的重要载体,物理和社会空间数据通过网络传输到信息空间,而信息空间经过计算和决策对物理和社会空间进行反馈。

人机共生<sup>[3]</sup>是一个信息物理社会融合系统,它可以通过身体和脑部传感器检测人的认知行为,推断用户的意图。另外,城市大数据系统也是一个典型的信息物理社会融合系统,它集成和融合信息空间(如政府和组织机构的数据)、物理空间(如监控摄像头、智能手机等)和社会空间(如群智感知<sup>[4]</sup>和移动社交感知<sup>[5]</sup>)的信息来实现对城市动态的高质量感知。然而,这种通过数据驱动的信息物理社会融合系统还没有得到深入研究。

本文以数据为中心刻画信息物理社会融合系统的新特征,并且构建了一个统一的系统框架。此外,还研究了跨空间协同感知和数据融合等信息物理社会融合系统面临的挑战,并且定义了未来的重点研究方向。

## 2 信息物理社会融合系统的新特征与挑战

信息物理社会融合系统的目标是实现人、机、物和谐计算环境,是未来计算技术追求的愿景,其中人类社会、信息空间、物理世界三者的全面连通与融合正在成为新一代计算技术的重要特征与主要趋势。近年来新发展的一些热点研究方向很好地体现了CPSS理念。

- 人计算(human computation)<sup>[6]</sup>: 变“人让机器算”为“机器让人算”,通过设计各种游戏/应用,让用户无意识地贡献人类的识别能力。

- 混合智能(hybrid intelligence): 近期,《中国工程院院刊》发表人工智能2.0专题<sup>[7]</sup>,指出“人机混合增强智能”是人工智能2.0的5个重要发展方向之一。人类智能与机器智能具有较大的互补性,机器的优势在于快速、低成本地对信息进行存储、比较、排序和检索,而人脑的优势在于联想、推理、分析、归纳等,所以实现人机智能的协同能极大地提高系统的性能。

- 群智感知计算(mobile crowd sensing and computing)<sup>[4]</sup>: 利用大量普通用户使用的移动设备作为基本感知单元,通过物联网/移动互联网进行协作,实现感知任务分发与感知数据收集利用,最终完成大规模、复杂的城市与社会感知任务。群智感知计算系统是典型的人机物融合系统,提高了城市级别感知的时空覆盖面,并通过融合人的参与,实现对大规模数据的高效传输和处理。

- 城市计算(urban computing)<sup>[8]</sup>: 城市计算就是用城市中的大数据解决城市本身面临的挑战,通过对多种异构数据的整合、分析和挖掘,提取知识和智能,并用

智能创造“人—环境—城市”三赢的结果。

以上均从不同侧面体现了信息、物理与社会空间融合计算的目标，与CPSS密切相关。随着相关研究方向的不断发展，CPSS的内涵和外延也得到不断丰富和扩展。总体而言，相比一般大数据系统研究而言，以数据为中心的CPSS体现出如下新特征和挑战。

- 数据时空交织性：CPSS在传统信息空间基础上，融入物理空间和社会空间信息，其数据具有较为明显的时间和空间特征。此外，由于人类行为在不同空间之间进行不断切换，实现三元空间数据的时空交织和复杂关联。其挑战在于如何融化空间隔离，发现不同空间数据之间的多维语义关联，为协同感知和数据融合提供重要基础。

- 跨空间数据融合：不同空间的感知能力和感知数据体现出差异性和互补性特征，如物理传感器可以感知用户在现实世界的位置、轨迹、活动等信息，而社交网络则体现用户的在线交互和社会关系。其挑战在于如何通过跨空间感知和计算能力的融合，实现对感知目标的前景刻画。

- 人机智能协作性：CPSS数据具有海量、跨空间、多模态、复杂关联等特征，如

何对其进行高效处理和准确理解成为新的挑战，而单纯利用机器智能则很难实现该目标。在CPSS中，更关注人类和机器智能协作完成复杂的数据理解任务。需探索群智感知、人机智能融合等关键理论与方法，实现人和机器智能的深度融合和高效计算。

### 3 以数据为中心的信息物理社会融合系统

数据驱动的模式将来有望成为一种新的软件开发模式，这也将指导信息—物理—社会融合应用和服务的变革。本文利用从多种数据源中得到的跨空间、多模态数据，更准确地刻画感知目标，而3个空间数据的融合将促进城市中大量新颖应用和服务的产生。

从“数据为中心”的观点来看，每个信息物理社会融合系统遵循从数据采集、数据处理到数据应用的过程。图1给出了以数据为中心的信息物理社会融合系统的通用框架。

这个框架的输入为信息空间、物理空间和社会空间的数据源，输出为数据在不同领域的应用。在4个垂直层中，左边两个

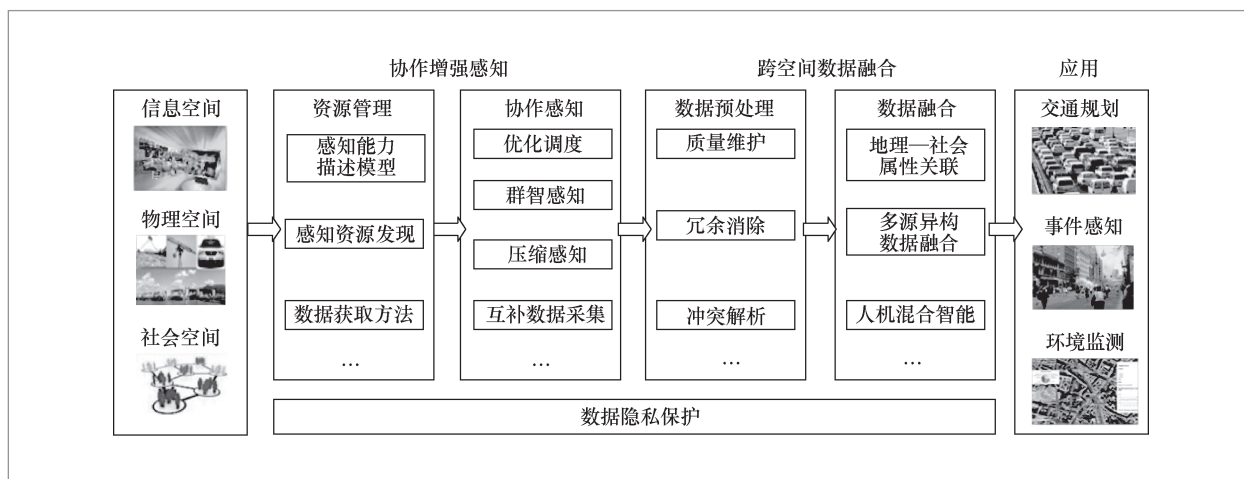


图1 一个数据驱动的信息物理社会融合系统框架

和数据收集相关,称为“协作增强感知”,右边两个则负责数据处理,称为“跨空间数据融合”。其中协作增强感知强调利用不同空间感知能力的差异性,实现对感知对象的增强感知;跨空间数据融合则利用不同空间数据的互补性,通过融合实现对感知对象的高效理解。作为一个融合人类参与的计算系统,信息物理社会融合系统应该在整个过程中保护数据的私密性。接下来将对每个分层的组成和功能进行介绍。

### 3.1 协作增强感知

数据驱动的CPSS连接了线上和线下感知资源,不同资源具有不同的感知和计算能力,为提高感知性能需要对其进行协同。在协作感知中,输入为感知任务和不同空间的感知源,输出为与感知任务相关的数据。图1中的资源管理层和协作感知层表示了协作增强感知的过程。

#### (1) 资源管理层

资源管理层管理来自于信息、物理、社会空间的多种感知源。当输入一个感知任务时,该层能快速根据其需求定位相关资源。该层具有如下功能。

- 感知能力描述模型: 一个统一的模型应该可以描述各种感知资源的能力。不同侧面的感知能力通过该模型进行描述,包括时空覆盖率、采样频率、准确率、误差范围和数据类型等。

- 感知资源发现: 该模块根据任务需求和每个感知资源的描述能力发现相关的感知资源。

- 数据获取方法: 该模块负责从不同数据源中收集数据,如使用软件应用程序编程接口(application programming interface, API)、网页爬虫或普适感知技术。

#### (2) 协作感知层

协作感知层根据动态感知任务需求,

完成资源调度及资源间协作。该层具有以下功能。

- 优化调度: 由于所感知的目标具有动态性,所以当前的感知资源常常需要重新组织。例如,当感知目标移动或者获得了一个相近的目标时,就需要发掘和激活新的可利用的感知资源。

- 群智感知: 随着智能手机数量的增加,群智感知可以让用户以不同形式参与数据采集,包括离线感知和在线数据生成。

- 压缩感知: 由传感器收集到的数据常常具有空间和时间上的相关性。大量数据可以通过过滤或压缩的方式实现数据存储和处理。

- 互补数据采集: 不同的感知源反映感知目标的不同侧面,因此可以从中采集互补的数据实现对感知目标的完整刻画。

### 3.2 跨空间数据融合

为了更好地识别所感知目标的特征,数据驱动的信息物理社会融合系统从不同的感知源中融合互补数据。系统框架中的数据预处理层和数据分析层则主要负责后台数据的处理和挖掘。

#### (1) 数据预处理层

数据预处理层负责数据清理和质量维护,它应具有以下功能。

- 质量维护: 多种因素可能会影响数据的质量。例如,用户生成的社交媒体数据可能存在虚假,人们会提供一些矛盾的信息,传感器读取的数据可能会出现错误等。因此,需要寻找合适的方法评估数据的质量。

- 冗余消除: 通过群体参与方式贡献的数据往往会产生语义和内容上的冗余。为有效地收集和处理数据,应根据任务需求定义数据的冗余性,进一步消除冗余数据。

- 冲突解析: 传感器读取和获得的信

息有时互相冲突。例如,由于环境的差异,一组位置相近的智能手机在同一时间会得到不同的识别结果。由跨空间、多模态传感器获取的数据在语义层面的不一致则更难被发现和处理。

#### (2) 数据融合层

数据融合层融合来自不同空间的数据,并且学习它们之间的相关性。它主要具备以下功能。

- **地理—社会属性关联:** 跨空间数据往往呈现不同维度的特征,如时间、空间和社会性等,这些属性以人类的行为作为纽带紧密关联起来。对不同空间属性间的关联和相互作用规律进行挖掘具有重要意义,例如结合用户的社会关系,可以提高移动性预测的准确性<sup>[5]</sup>。

- **多源异构数据融合:** 从在线社交媒体或物理传感器获得的数据虽然来自不同数据源,但可能代表同一个对象(如同一事件或个体)。特别是从这些数据源获得的关于某一对象的信息具有互补性,可以帮助系统增强对感知对象的理解。在这方面可以借鉴有关多源数据融合<sup>[9]</sup>和跨媒体挖掘<sup>[10]</sup>等领域的最新进展。

- **人机混合智能:** 人类的参与为CPSS数据的收集和处理带来很多方面的优势,而人类和机器智能如何进行更好的融合是值得研究的新问题。笔者之前的工作讨论了几种可能的人机智能融合方式<sup>[11]</sup>。

## 4 数据驱动的信息物理社会融合系统应用

数据驱动的信息物理社会融合系统将会不断改变城市/社会管理及运行的方式。尽管该方面的研究还处在早期阶段,但已经在一些领域体现了其优势。这些应用涉及的领域包括城市管理、公共安全、

智能交通和商业智能等。

- **城市管理:** 当无线传感器设备遍布于整个城市空间及城市居民身上时,就可以利用收集的大规模数据解决一些日益严重的城市问题。SeeClickFix<sup>[12]</sup>和PublicSense<sup>[13]</sup>通过市民参与,利用其携带的智能移动终端实现对城市市政问题(道路塌陷、井盖丢失、路灯不亮)、环境卫生问题等的及时发现与报告,提高了城市精细化管理的效率和质量。微软亚洲研究院通过城市跨空间数据优化城市规划<sup>[14]</sup>,并实现对PM2.5的细粒度预测<sup>[15]</sup>。

- **公共安全:** 利用信息和物理空间多种信息源,可以对影响公民安全的事件(如恐怖袭击、自然灾害等)进行预警或及时响应。InstantSense<sup>[16]</sup>通过群体参与实现对城市热点和敏感事件的实时感知与多侧面在线呈现。美国警方使用多源感知数据对历史性逮捕模式、发薪日、体育项目、降雨天气和假日等变量进行分析,结合大规模历史犯罪记录对犯罪行为进行预测,实现警力的优化配置,并极大降低了犯罪率。

- **智能交通:** 美国麻省理工学院的“实时罗马”<sup>[17]</sup>项目,利用从罗马市内收集来的移动电话、公交车及出租车信息,分析城市热点地区动态变化规律。Google Waze允许驾驶员及时报告道路交通状况(交通事故、拥堵)和周边信息(加油站打折信息),实现城市实时交通动态的共享。TripPlanner<sup>[18]</sup>通过融合基于位置社交网络数据和城市出租车轨迹数据,实现个性化的旅游路线推荐。

- **商业智能:** 通过融合城市跨空间多源数据(如车辆移动轨迹数据、商业评价网站数据、城市路网数据、社交网络数据、城市地图兴趣地点数据等)可以实现对商业行为的预测与智能推荐,如商业选址<sup>[19]</sup>、房价预测<sup>[20]</sup>等。

## 5 未来研究挑战

作为一个新的研究领域,实现数据驱动的信息物理社会融合系统还面临很多挑战,下面对其进行简要说明。

- **协作增强感知:** 跨空间协作和增强感知的研究仍然处于初级阶段,在动态感知过程中如何及时发现和感知目标关联的感知源,并利用其互补能力,是未来的重要研究方向之一。

- **数据关联和语义表达:** 在CPSS中,虽然数据来源和模态不同,但它们可能描述的是相同的目标,例如文字和图片可能刻画的是同一事件。如何发掘多源异构数据间的关联性,并对其进行统一语义表达是对CPSS数据进行融合的前提和关键。

- **跨空间数据挖掘:** 不同空间数据特征往往反映不同侧面的内容,融合来自不同空间的数据可实现对目标的全面和准确理解。因此跨空间数据挖掘是CPSS主要研究方向之一<sup>[21]</sup>。

- **隐私保护:** 来自多源空间用户的数据相互关联将从不同侧面实现对用户的画像,增加公众对隐私暴露的恐慌。因此,新的隐私保护方式应该解决在跨空间学习和预测过程中的用户隐私暴露问题。

- **新颖的应用领域:** 为了推进数据驱动的CPSS的发展,在已有应用的基础上,应该不断探索其在不同领域的独特优势。

## 6 结束语

信息、物理、社会三元空间相互融合的本质是为人们提供更透明、更智能、更泛在、更绿色、更安全的一体化服务,实现高

度和谐的社会生活环境。“数据”驱动的信息物理社会融合系统集合了信息、物理和社交空间的力量,提高了组织的效率和人们的生活质量。然而,信息物理社会融合系统的相关研究还处于初始阶段,在该领域还有很多的挑战和问题需要开展深入研究,这将为人们带来许多新机遇。

## 参考文献:

- [1] WANG F Y. The emergence of intelligent enterprises: from CPS to CPSS[J]. IEEE Intelligent Systems, 2010, 25(4): 85-88.
- [2] SCHIRNER ET AL G. The future of human in the loop cyber physical systems[J]. Computer, 2013, 46(1): 36-45.
- [3] LICKLIDER J. Man-computer symbiosis[J]. IRE Transactions on Human Factors in Electronics, 1960(1): 4-11.
- [4] GUO B, WANG Z, YU Z W, et al. Mobile crowd sensing and computing: the review of an emerging human-powered sensing paradigm[J]. ACM Computing Surveys, 2015, 48(1): 1-31.
- [5] EAGLE ET AL N. Inferring social network structure using mobile phone data[J]. PNAS, 2007, 106(36): 15274-15278.
- [6] QUINN A J, BEDERSON B B. Human computation: a survey and taxonomy of a growing field[C]// The ACM CHI Conference on Human Factors in Computing Systems, May 7-12, 2011, Vancouver, Canada. [S.l.:s.n.], 2011: 1403-1412.
- [7] PAN Y H. Special issue on artificial intelligence 2.0[J]. Frontiers of Information Technology & Electronic Engineering, 2017, 18(1): 1-2.
- [8] ZHENG Y, CAPRA L, WOLFSON O, et al. Urban computing: concepts, methodologies, and applications[J]. ACM Transactions on Intelligent Systems and Technology, 2014, 5(3): 38.
- [9] CHOO J, BOHN S, NAKAMURA G C, et al.

- Heterogeneous data fusion via space alignment using nonmetric multidimensional scaling[C]// SIAM International Conference on Data Mining (SDM), April 26–28, 2012, Anaheim, USA. [S.l.:s.n.], 2012: 177–188.
- [10] SHI Z. Cross-media knowledge discovery[C]// The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 12–16, 2012, Beijing, China. New York: ACM Press, 2012.
- [11] GUO B, CHEN C, YU Z, et al. Building human-machine intelligence in mobile crowd sensing[J]. IEEE IT Professional, 2015, 17(3): 46–52.
- [12] BERTOT J C, JAEGER P T, MUNSON S, et al. Social media technology and government transparency[J]. Computer, 2010, 43(11): 53–59.
- [13] ZHANG J, GUO B, CHEN H, et al. Public sense: re-ined urban sensing and public facility management with crowdsourced data[C]// 2015 IEEE International Conference on Ubiquitous Intelligence and Computing, August 10–14, 2015, Beijing, China. New Jersey: IEEE Press, 2015: 1407–1412.
- [14] ZHENG Y, LIU Y, YUAN J, et al. Urban computing with taxicabs[C]//The 13th International Conference on Ubiquitous Computing, September 17–21, 2011, Beijing, China. New York: ACM Press, 2011: 89–98.
- [15] ZHENG Y, LIU F, HSIEH H. U-air: when urban air quality inference meets big data[C]// The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 11–14, 2013, Chicago, USA. New York: ACM Press, 2013: 1436–1444.
- [16] CHEN H, GUO B, YU Z, et al. Toward real-time and cooperative mobile visual sensing and sharing[C]// The 35th Annual IEEE International Conference on Computer Communications, April 10–14, 2016, San Francisco, USA. New Jersey: IEEE Press, 2016: 1–9.
- [17] CALABRESE F, RATTI C. Real time rome[J]. Networks and Communication Studies, 2006, 20(3–4): 247–258.
- [18] CHEN C, ZHANG D, GUO B, et al. TripPlanner: personalized trip planning leveraging heterogeneous crowdsourced digital footprints[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(3): 1259–1273.
- [19] KARAMSHUK D, NOULAS A, SCELLATO S, et al. Geo-spotting: mining online location-based services for optimal retail store placement[C]// The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 11–14, 2013, Chicago, USA. New York: ACM Press, 2013: 793–801.
- [20] FU Y, GE Y, ZHENG Y, et al. Sparse real estate ranking with online user reviews and offline moving behaviors[C]// The 2014 IEEE International Conference on Data Mining, December 14–17, 2014, Shenzhen, China. New Jersey: IEEE Press, 2014: 120–129.
- [21] GUO B, YU Z, ZHANG D, et al. Cross-community sensing and mining[J]. IEEE Communications Magazine, 2014, 52(8): 144–152.

## 作者简介

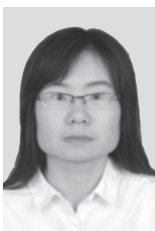


翟书颖 (1981–), 女, 西北工业大学明德学院讲师, 主要研究方向为物联网与大数据分析。

## 作者简介



**郭斌** (1980-), 男, 博士, 西北工业大学计算机学院教授、博士生导师, 主要研究方向为普适计算、群智感知计算。



**李茹** (1987-), 女, 西北工业大学明德学院讲师, 主要研究方向为物联网。



**王庭良** (1979-), 男, 西北工业大学明德学院讲师, 主要研究方向为物联网。



**於志文** (1977-), 男, 博士, 西北工业大学计算机学院教授、博士生导师, 主要研究方向为普适计算与智能感知。



**周兴社** (1955-), 男, 博士, 西北工业大学计算机学院教授、博士生导师, 主要研究方向为信息物理融合计算。

收稿日期: 2017-06-14