

浅析大数据技术 及其在土木工程中的应用

叶金明

河南理工大学, 焦作

摘要 | 大数据技术相比于传统的数据分析技术可以实现对海量数据的快速处理和价值挖掘, 当前已经应用于医疗、能源等行业。为探寻大数据技术在土木工程中的应用前景, 首先需要把握其应用现状。本文在 Web of Science 数据库、Engineering Village 数据库以及中国知网等国内外有代表性的文献数据库中检索相关文章并做归纳分析, 总结大数据技术在土木工程中的应用现状, 并预测在土木工程中应用大数据技术的发展趋势, 为在土木工程中进一步应用大数据技术提供参考依据。

关键词 | 大数据技术; 土木工程; 大数据应用; 数据处理; 价值挖掘

Copyright © 2022 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). <https://creativecommons.org/licenses/by-nc/4.0/>



1 前言

近几年, 随着互联网的快速发展, 在互联网上发布、传递和共享信息产生了大量的数据, 同时, 传感器和物联网在交通、环境等方面的大规模使用也使得数据呈现暴发式增长。另外, 通过实施企业信息化, 大量的管理数据也积累在企业管理信息系统中。这些数据不仅包含结构化数据, 还包含半结构化数据

作者简介: 叶金明, 河南理工大学在读研究生, 研究方向: 信息技术在土木工程中的应用。

文章引用: 叶金明. 浅析大数据技术及其在土木工程中的应用 [J]. 土木工程进展, 2022, 4 (3): 44-51.

<https://doi.org/10.35534/ace.0403005c>

和非结构化数据。对这些数据进行分析和处理,可以得到大量有价值的信息。利用传统数据处理方法处理这些海量数据时往往效率较低,而且半结构化和非结构化数据的存在也给数据处理带来困难。

这些在有限时间内用传统数据处理方法无法处理完的数据,通常被称作大数据。虽然对大数据的分析要比对传统数据的分析困难,但大数据蕴含巨大价值,对大数据进行有效挖掘可以得到传统数据无法分析得到的价值,因此专门用于大数据处理的大数据技术已应用在医疗行业、能源行业、通信行业、零售业和科学研究等。

为了应对大数据的挑战和机遇,各国政府都提出了相应的应对措施。2012年3月美国政府启动大数据发展计划,该计划斥资2亿多美元,利用大数据技术在环境、生物医学等领域进行突破。2015年9月我国国务院印发《促进大数据发展行动纲要》,提出形成政府数据统一共享交换平台的目标,并对促进大数据产业健康发展、规范利用大数据保障数据安全等方面做出部署。

土木工程支撑基础设施及民众住房建设,很明显,土木工程中会产生大数据,且土木工程中涉及大量决策,可以预计,大数据技术在土木工程中具有良好的应用前景。本文旨在总结大数据技术在土木工程中的应用现状并预测其发展趋势。为此,本文首先阐述大数据技术的关键概念,然后基于文献调研分析大数据技术在土木工程中的应用现状和发展趋势。

2 大数据及大数据技术简介

2.1 大数据及大数据技术

当前大数据并没有统一的定义,不同的定义试图通过描述大数据的特征给出定义。Grobelnik认为大数据的特征为:体量大(volume)、多样(variety)和高速(velocity)。国际数据公司IDC则认为大数据具有以下四个特性。

体量大:大数据处理的数据集体量大,一般为TB或PB($1\text{PB}=2^{10}\text{TB}$)级别,超出传统数据处理方式的处理能力。

多样:数据来源广泛,包含结构化、半结构化和非结构化数据。

高速：对于大数据的处理应当快速、实时。

价值：通过挖掘大数据，可以得到隐藏在数据中的价值，这些价值是挖掘传统数据所不能得到的。

专门用于处理大数据的技术就是大数据技术。IDC 将大数据技术定义为“大数据技术描述了新一代的技术和架构体系，通过高速采集、发现或分析，提取各种各样的大量数据的经济价值”，说明大数据的高速和价值特征需要通过大数据技术实现。

2.2 大数据处理环节及处理环节中用到的技术

如果不对大数据进行分析和处理，那么它只是一堆数据。只有对它进行分析和处理，才能挖掘出潜藏在其中的价值。一般地，大数据处理需经过产生大数据、获取大数据、存储大数据和分析大数据等 4 个环节。

为实现大数据价值，需要在大数据处理的各环节应用相应的技术。各环节所使用的技术举例如下。

- (1) 产生大数据：管理信息系统、社交网络、传感器、智能仪表等；
- (2) 获取大数据：基于领域知识的搜索技术、自然语言处理技术等；
- (3) 存储大数据：NoSQL、分布式数据库等；
- (4) 分析大数据：Map Reduce 计算模型、Dryad 计算模型等。

大数据处理中用到的上述技术，一部分属于传统数据处理技术，例如管理信息系统、自然语言处理技术等；另一部分属于大数据技术，例如 NoSQL、分布式数据库等。在分析大数据环节所用到的方法一般称为大数据分析方法。

2.3 大数据分析方法

传统的数据分析方法，例如聚类分析、相关分析、回归分析、数据挖掘、机器学习等，在处理大数据时会遇到效率低下等挑战，在大数据时代需要在这些方法的算法等方面做出调整，才能满足要求。

另外，也产生了一些专门针对大数据的数据分析方法，包括 Bloom Filter、Hashing、索引、并行计算等。这些大数据分析方法使得短时间内分析大规模数

据集、得到可辅助决策的数据分析结果成为可能。

2.4 大数据分析方法与传统数据分析方法的区别

大数据分析方法与传统数据分析方法的数据处理速度不同。传统数据分析方法在处理数据量特别大的数据时,存在处理速度慢的问题。得益于计算机硬件和数据处理架构、方法的改进和提高,大数据分析方法使得快速处理大量数据成为可能。

大数据分析方法与传统数据分析方法的分析对象不同。由于传统数据采集技术的不足,往往很难获得全面的数据。因此绝大多数传统数据分析方法的分析对象是随机抽样选择的样本数据。随着互联网、检索技术、传感器的大规模应用,获得一个事件多维度的数据已经不是技术上的难题。这为大数据的产生创造了客观条件,也决定了大数据分析方法的分析对象不是随机抽取的样本,而是数据的全体,即大数据分析方法是从数据全体中挖掘有价值的信息。

3 大数据技术在土木工程中的应用

为把握当前大数据技术在土木工程中的应用现状及其发展趋势,笔者采用系统地检索并分析相关文献的方式。为此,选取了国内外有代表性的文献数据库进行检索。

首先对 Web of Science 的核心合集数据库进行了检索。在检索时对目标文献限制如下:文献主题或标题包含“Big Data”,研究领域为“Construction & Building Technology”(Web of Science 对文献研究领域进行划分,其中与土木工程有关的只有“Construction & Building Technology”),不对文献发表时间做限制。通过对 Web of Science 核心合集数据库进行检索,共得到 14 篇文献。

其次对 Engineering Village 的核心期刊数据库进行了检索。在文献检索时对目标文献限制:文献标题或摘要包含“Big Data”。检索得到 6224 篇文献。这些文献被 Engineering Village 划分为多个研究领域,其中与土木工程相关的研究领域有“Buildings and Towers”“Accidents and Accident Prevention”“Engineering Research”“Structural Design”和“Engineering Profession”,将文献限制为这 5

个领域后得到 568 篇文献。

最后对中国知网数据库进行了检索，在文献检索时对目标文献限制如下：文献篇名、关键词或摘要包含“大数据”。检索得到文献 39681 篇。这些文献被中国知网划分为多个研究领域，其中与土木工程相关的研究领域有“建筑施工安全”“灾害及其防治”“铁路线路工程”“道路工程”“桥涵工程”“隧道工程”“航道工程”“通航建筑物及助航设备”“港口工程”“其他科学技术在建筑中的应用”和“建筑工程”，将文献限制为这 11 个领域后得到文献 238 篇。

3.1 大数据技术用于辅助建筑能耗分析

建筑能耗的一个重要因素是建筑占用，它影响建筑的光照、建筑内部的热交换等。D'Oca 等在开源数据挖掘软件 Rapid Miner 的基础上提出一个数据挖掘框架，以发现办公空间占用模式。该研究工作首先用传感器得到法兰克福一栋办公楼 16 个办公室 2 年内每隔 10 min 的占用数据，然后以这些占用数据为基础，利用决策树挖掘、规则归纳和聚类分析等数据分析方法得到建筑占用模式和时间表，并以 4 种原型工作状况图表示出来。在进行建筑能耗分析时，将这些工作状况作为建筑能耗建模软件（例如 Energy Plus、IDA-ICE 等）的输入，研究占用情况对办公楼的设计、运营和能源使用的影响，最后根据研究结果，采取不同的能源节约策略或推荐合适的建筑设计。

Lee 等根据用电设备历史电力消耗数据分析得到电力消耗模式，并以此为基础预测未来一段时间的电力消耗情况。该研究工作以一座建筑中的办公室、会议室、实验室及服务器机房共计 240 间房间为样本，通过专门的传感器每隔一分钟采集照明设备、通风设备、低压设备（如电脑、打印机等）和高压设备（如服务器、实验仪器等）的耗电量数据，数据量达 10GB。然后，通过特征提取、聚类和关联分析等方法得到不同用途的房间中不同类型的设备的电力消耗模式，并预测未来一段时间的电力消耗。当发现对于与预测不符的情况时，针对四种类型设备分别进行当前数据和历史数据的对比，分析可能发生的事件，以便采取措施降低能耗。例如，对比历史数据，某间会议室的低电压设备几小时内的平均耗电量显著增加，可能是由于临时召开会议，大量人员进入。而某间房间

照明及低压设备无电力消耗但通风设备仍在耗电，并与历史数据不符，可能是空调在无人的房间内空转。

3.2 大数据技术用于建筑破坏检测

利用无人机图像可以进行建筑震后快速破坏检测。通过利用无人机可在短时间内拍摄成千上万张图像，提高图像处理速度对于及时的震后建筑破坏评估和灾后救援至关重要。Hong 等提出利用并行计算处理震前地形图和震后无人机图像以加快建筑破坏三维检测速度。该处理过程对震前地形图经过坐标转换和海拔提取等处理后生成震前数字表面模型，同时对震后无人机图像经过连测点提取、相机校正、生成准核线影像、用半全局匹配方法进行密集匹配等处理后生成震后数字表面模型，然后计算震前和震后数字表面模型的不同，进而得到建筑破坏三维检测结果。

4 土木工程中大数据技术应用趋势

考虑到土木工程涉及海量数据，特别是，随着物联网技术的发展，可在土木工程的具体对象中可以布置更多的传感器，以便更方便地采集数据，再通过大数据技术进行分析，即可更好地把握土木工程的相关规律。利用这些规律就可以更好地进行土木工程的建造、维护等各方面的工作，让大数据技术在土木工程中发挥作用。

土木工程是一个很大的领域，对任何研究者来说，全部把握该领域几乎不可能。为了把握土木工程中大数据技术可能的应用点有必要借助现有文献，特别是关于大数据技术应用的、议论性的文献。因为，这些文献描述土木工程中大数据技术的应用，往往揭示了土木工程中应用大数据技术的潜在应用点，反映了作者对大数据技术应用趋势的预测。

为此，笔者分析了议论土木工程中大数据技术应用的可能性或提出土木工程中大数据技术应用系统架构但没有实际应用的文献，对其中大数据技术的应用点进行归纳。大数据技术的应用点主要在土木工程运维阶段，具体应用对象包括建筑、道路、桥梁等。其中，运维阶段相比于设计阶段和施工阶段，其数

据采集更方便且数据量大,涉及的决策也多,具有运用大数据技术的天然优势。由此可见,在土木工程的运维阶段,大数据技术具有最广阔的应用前景,而在其他阶段,也具有很大的应用潜力。

5 结语

本文首先阐述了大数据及大数据技术的相关概念,然后选取有代表性的国内外典型科技文献数据库,系统地检索了与大数据及土木工程相关的文献,并进行了分析。结果表明,当前大数据技术在土木工程中主要应用于辅助建筑能耗分析和建筑破坏检测等方面,应用特征表现为:数据源主要是传感器和图像设备获得的数据,数据量相对较少,较少使用真正的大数据分析方法。随着物联网技术的发展及其在土木工程中的应用,土木工程大数据将迅速发展,大数据技术在土木工程中的应用方兴未艾。本文可为在土木工程中进一步应用大数据技术提供参考。

参考文献

- [1] 孟小峰,慈祥. 大数据管理:概念、技术与挑战[J]. 计算机研究与发展, 2013, 1(1): 146-169.
- [2] 张引,陈敏,廖小飞. 大数据应用的现状与展望[J]. 计算机研究与发展, 2013, 50(2): 216-233.
- [3] D'Oca S, Hong T. Occupancy schedules learning process through a data mining framework[J]. Energy and Buildings, 2015, 88: 395-408.
- [4] Hong Z, Tong X, Cao W, et al. Rapid three-dimensional detection approach for building damage due to earthquakes by the use of parallel processing of unmanned aerial vehicle imagery[J]. Journal of Applied Remote Sensing, 2015, 9(1).

Analysis of Big Data Technology and its Application in Civil Engineering

Ye Jinming

Henan Polytechnic University, Jiaozuo

Abstract: Big data techniques can rapidly process and mine value from mass data compared to traditional data analysis methods in many domains, such as finance and energy industry. Before we ascertain the prospect of big data techniques in civil engineering, the status quo of big data techniques that are used in civil engineering should be investigated. Articles related to big data and civil engineering are analyzed after being searched in the major literature databases including Web of Science, Engineering Village and China Knowledge Resource Integrated Database, and the status quo of big data technique application in civil engineering and the trends are analyzed. This paper facilitates the future research on the application of big data techniques in civil engineering.

Key words: Big Data Technique; Civil Engineering; Big Data Application; Data Processing; Value Mining