

DOI: 10.19595/j.cnki.1000-6753.tces.200257

能源互联网数据交易：架构与关键技术

郭庆来^{1,2} 王博弘^{1,2} 田年丰^{1,2} 孙宏斌^{1,2} 温柏坚³

(1. 清华大学电机工程与应用电子技术系 北京 100084

2. 电力系统及发电设备控制和仿真国家重点实验室(清华大学) 北京 100084

3. 广东电网有限责任公司信息中心 广州 510630)

摘要 随着大数据技术的发展,数据成为一种愈发重要的特殊资产。然而,许多能源互联网中的利益主体对于隐私的保护使得能源互联网中的数据相互割裂、无法流通,严重限制了数据转换为实际经济价值的潜力。在能源互联网逐渐成熟的背景下,能源和数据信息的高度融合将极大地促进二者的协同发展。因此,构建完整的数据交易体系架构并研究相关的技术实现对于能源互联网的发展而言具有重大的战略意义。该文对能源互联网中的数据交易架构进行了展望,针对数据交易中亟需解决的数据价值评估、数据定价、数据确权和数据隐私保护等诸多问题进行了梳理和探讨,阐明这些问题之间的关联关系,同时基于信息论、博弈论、区块链技术等现有理论或技术给出初步解决方案,为能源互联网中数据交易技术的形成和推广提供借鉴。

关键词: 能源互联网 数据交易 数据价值评估 数据定价 数据确权 数据隐私保护

中图分类号: TM73

Data Transactions in Energy Internet: Architecture and Key Technologies

Guo Qinglai^{1,2} Wang Bohong^{1,2} Tian Nianfeng^{1,2} Sun Hongbin^{1,2} Wen Bojian³

(1. Department of Electrical Engineering Tsinghua University Beijing 100084 China

2. State Key Laboratory of Control and Simulation of Power Systems and Generation Equipments
Tsinghua University Beijing 100084 China

3. Information Center Guangdong Electric Power Company Guangzhou 510630 China)

Abstract With the development of big data technology, data have become an increasingly important special asset. However, the protection of privacy by many members in the energy Internet makes the data in the energy Internet cut off from each other and unable to circulate, which seriously limits the potential of data to be converted into actual economic value. Under the background of the gradual maturity of energy Internet, the high integration of energy and data information will greatly promote the coordinated development of them. Therefore, it is of great strategic significance for the development of energy Internet to build a complete data transaction system architecture and study the related technology implementation. This paper summarizes the data transaction architecture in the energy Internet, and comprehensively surveys and discusses many problems in data transaction, such as data value evaluation, data pricing, data right confirmation, data privacy protection, etc. Besides, this paper explains the relationships between these problems and gets their preliminary solutions based on information theory, game theory, block chain technology, and other existing theories or technologies, which will guide the formation and promotion of data transaction technology

in the energy Internet.

Keywords: Energy internet, data transactions, data value evaluation, data pricing, data right confirmation, data privacy protection

0 引言

随着互联网和计算机技术的发展,人们收集、加工和分享数据的能力得到了极大的提升,大数据等数据科学技术^[1]可以对海量数据进行分析、挖掘和利用,进而获取相应的价值。因此,数据逐渐成为促进经济增长和社会发展的主要生产要素^[2]。

能源互联网作为能源技术革命的代表,核心是基于互联网的理念对能源网络进行重构,促进能源的市场化、高效化和绿色化,是能源系统与信息系统高度融合的结果^[3-4]。数据资产的交易与共享作为能源互联网中各个利益主体之间信息交互的有效手段,日益得到重视。

信息论的奠基人香农(C. E. Shannon)认为信息最主要的特征是可以消除不确定性^[5]。这一结论尽管源自于香农在信息“语法”层面的研究成果,但是扩展到信息“语义”和“语用”层面也仍然适用^[6],即可以认为信息的传递与流通对于不同利益主体的价值也体现在信息所消除的不确定性中。

在能源互联网中,随着分布式可再生能源渗透率的提高,源侧和荷侧的强不确定性给能源系统带来巨大的挑战^[7-10]。从广义上看,各种运行与规划场景绝大部分都是在不确定性下如何做出风险决策的问题。如果通过数据的交易、共享或者挖掘获取信息,使得能源互联网中的利益主体在决策中所面临的不确定性被减少或消除,相应决策策略的潜在风险成本也将减少,由此带来相应的经济效益,而这就是数据或信息所带来的效益增量。

近年来,大数据交易发展迅速、前景广阔。根据贵阳大数据交易所发布的《2016年中国大数据交易白皮书》显示,2015年全球大数据产业规模达到了1403亿美元,2020年预计可以达到10270亿美元^[11]。根据中国信息通信研究院发布的《大数据白皮书(2018年)》显示,我国大数据产业规模自2015年以来连续多年维持约30%的同比增长,在2020年预计可以达到10100亿元人民币^[12]。因此,数据交易的相关研究具有重要的战略意义。

目前,数据交易的模型在公共管理领域已有部分研究^[13-14],数据交易的机制设计在经济学领域^[15-16]和法学领域^[17-18]的论述也较多。除此之外,数据确权问题由于区块链技术的兴起日益成为各学科研究

的焦点^[15,19-20]。但数据交易在电力系统和能源互联网学术领域的探讨^[21]不多,大部分停留在系统架构的设计和和应用场景的展望。

如今,以电力交易为代表的能量交易市场日趋成熟,在能量交易市场中通过对能源及其金融衍生品的交易可以在能源系统中实现资源的优化配置。参考能量交易市场的思路,能否构建电力/能源系统的数据交易市场?如果可行,那么对于打造能源互联网的新业态,打破不同利益主体之间的信息壁垒,提高数据交易共享的效率,充分发挥数据的内在价值将具有重要意义。可以预想,能量交易和数据交易之间存在相辅相成、协同发展的关系,能量交易中通过消除信息不确定性所带来的收益前景推动了数据交易的发展,而数据交易带来的价值又服务于能量交易,促进整体市场交易成本下降,提升了能源系统运行效率,形成良性循环。二者之间的关系如图1所示。

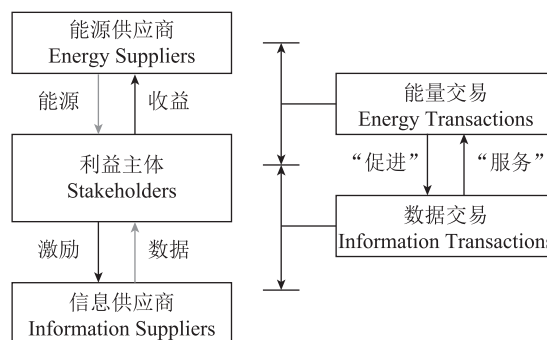


图1 能量交易与数据交易之间的关系

Fig.1 Relationships between energy transactions and data transactions

能源互联网数据交易涵盖诸多关键技术问题,包括数据价值评估、数据定价、数据确权、数据隐私保护等。其中,数据价值评估主要由参与数据交易的利益主体完成,而数据定价、数据确权和数据隐私保护则主要与数据交易市场运行机制的设计相关。这些问题尽管侧重于数据交易的各个方面,但仍然存在着紧密的联系。数据价值评估是数据定价的基础,依据数据价值评估结果进行市场定价将有助于动用市场力保证数据交易市场的良性竞争;数据确权是数据定价的前提,一旦数据的权属被冒用或盗用,将不仅直接影响相关权利人的经济利益,而且还间接影响利益主体共享交易数据的意愿,违

背了数据交易市场建立的初衷；数据确权是数据隐私保护的条件，数据确权将确认数据隐私保护的对象，从而应用相关的数据安全技术保护利益主体的数据隐私，建立良好和谐的数据生态。这些关键技术问题是数据交易架构形成的基础，也构成了数据交易机制设计的核心问题。四项问题之间的联系如图 2 所示。

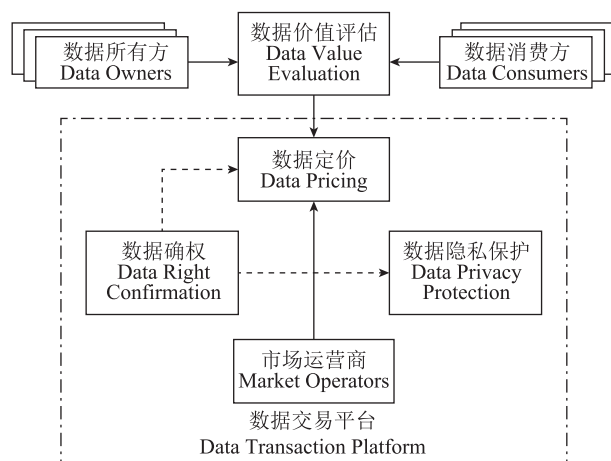


图 2 数据交易关键技术问题之间的联系

Fig.2 Relationships among key technology problems of data transactions

本文将重点对上述问题进行讨论，提出相应的初步解决方案与技术展望，抛砖引玉，期待与更多专家学者共同研讨和畅想电力与能源领域数据交易的未来愿景。

1 数据交易架构

数据交易架构是数据交易的实现基础，包括数据交易技术体系和数据交易组织形式两个部分。数据交易技术体系是数据相关技术的运用方式，数据交易组织形式是利益主体之间数据的交互形式。

1.1 数据交易技术体系

从数据交易的具体形式看，一般分成两种。一是直接交易原始数据，即将原始数据（可脱敏）整体从一个交易主体转移给另外主体，本质上交易的是数据的所有权和使用权，比如将某个片区的所有负荷或者光伏历史数据去除隐私信息后售出，核心技术挑战在于数据确权问题；二是交易原始数据衍生品（或称之为“数据产品”），即交易/交换原始数据处理后的关键特征与二次信息，交易过程中数据不可逆，无法还原原始数据本身，这本质上是一种数据服务的交易。换言之，仅交易数据的使用权，不交易所有权，比如从数据中提取关键的负荷分类信息、典型负荷曲线、可再生能源出力概率分布、

用户用电行为、企业履约情况等，再将这些二次“加工”后的信息进行交易，核心技术挑战除了确权问题外，还必须在保证原始数据安全的前提下解决大规模、异地、异构数据在动态环境下进行高性能、联合复杂分析的问题。

数据交易技术体系可借鉴云服务的体系架构、综合虚拟数据库等技术进行建构^[22]。其中，虚拟数据库可以对多方数据进行整合，并对数据进行配给；基于虚拟数据库的云服务从原始数据出发，由数据应用服务开发平台生成，为数据消费方提供数据产品和服务；数据应用服务开发平台将以数据为素材，依赖深度学习、图分析和统计分析等算法或技术对数据进行深层次的分析和挖掘。

在能源互联网中，存在以能源互联网调度控制中心、电力交易中心等为代表的数字所有者和以售电公司等为代表的数字消费方，一些高级的数据服务需要跨利益主体进行数据检索和计算，此时数据所有方的数据可以整合形成一个非物理集合形式的虚拟数据库，将原始数据保留在原地，实现隐私保护前提下的虚拟计算，具体技术可参见 5.2 节。数据交易技术体系如图 3 所示。

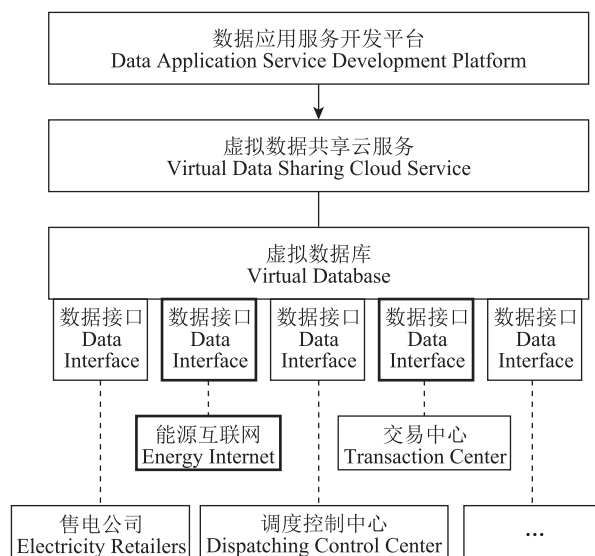


图 3 数据交易技术体系

Fig.3 Technology system of data transactions

1.2 数据交易组织形式

数据交易的组织形式包括集中式数据交易市场和分布式数据交易市场。

1.2.1 集中式数据交易市场

基于第三方交易平台的集中式交易是当前电子商务的主要形式之一，用于数据交易也较为合适。在能源互联网的背景下，数据交易将较大可能在主体之间以 B2B 模式开展，数据所有者和消费方将在

第三方的数据交易平台上进行数据投标, 由平台进行监管和集中出清, 从而充分发挥平台的优势, 降低存储成本, 节省协商时间, 有助于实现大规模的数据交易。

集中式数据交易在数据价值评估和数据定价方面更具优势。由于市场条件的限制, 分布式数据交易往往通过执行区块链技术中的智能合约实现, 而这对于数据价值评估和数据定价而言过于抽象和笼统。在集中式数据交易中, 数据价值将依赖于具体应用场景的建模进行准确评估, 数据定价也可以根据供需关系和市场行为的变动及时进行调整。因此, 在集中式数据交易中整体的模型分析和机制设计将更加灵活。

1.2.2 分布式数据交易市场

分布式数据交易相对于集中式数据交易而言尚不够成熟, 其中最大的发展阻碍在于无法很好地应对交易双方“信息不对称”的问题以及由此导致的市场低效和信息安全隐患。而目前的区块链技术有望成为解决这一问题的突破口。尽管区块链技术也带来了信息存储成本高和扩展性差等局限性, 但是随着区块链技术的发展, 分布式数据交易市场的发展前景仍然值得期待。

分布式数据交易在数据确权和数据隐私保护方面更具优势。在集中式数据交易市场中, 第三方交易平台承担交易过程的所有环节, 每次交易的相关内容只有交易双方和第三方平台可知。对于数据交易而言, 信息的不对称和不完备是有潜在风险的, 拥有较多信息的一方将有可能借助信息优势谋取额外收益, 甚至侵犯对方的隐私。因此, 在集中式数据交易中利益主体的数据权利和数据隐私保障是重要挑战。而在分布式数据交易中, 区块链技术使得信息透明公开, 不仅降低了信任成本, 而且对数据确权和数据隐私保护具有重要意义。

2 数据价值评估

2.1 研究意义

数据是一种新兴的无形资产, 因此数据交易与普通商品交易不同, 具有价值的不确定性等多种特征^[17,22]。数据的价值在交易时经常是未知的, 这是因为人们对数据的了解和使用过程相分离, 在对数据进行分析 and 应用之前无法对数据价值进行准确评估, 因此为了获取信息很可能需要付出与信息价值不匹配的成本。

从新制度经济学的角度, 可以得出相同的结论。

英国经济学家科斯(R. H. Coase)全面地介绍了交易成本理论和产权理论^[23-25]。由于产权包括数据的所有权和使用权, 因此人们需要明确界定数据的产权从而通过应用数据获取利润。然而, 由于数据效用的不确定性, 产权问题难以得到明确的界定, 数据交易的成本居高不下, 这无疑严重阻碍了数据的流通和共享。因此, 有必要通过具有一般性的数据价值评估方法作为数据定价的基础, 从根本上规范数据交易市场行为。

需要指出的是, 本文所探讨的数据价值是指数据的使用价值, 数据本身并没有特定价值, 也即数据只有被应用于特定的场景中时才能体现其使用价值。为了使数据价值评估具有可操作性和实际意义, 本文也将具体化数据价值, 将宽泛的“大数据”概念落实到有实际物理意义的“小数据”层面并进行理论分析。

2.2 数据价值评估的方法

在高比例分布式可再生能源渗透的背景下, 由源侧和荷侧注入能源互联网的强不确定性不仅将给能源互联网的规划、调度和运行带来许多挑战, 而且将影响能源互联网中的能量交易。其中, 不确定性对能量交易的影响更易由经济价值进行衡量, 因此, 能源互联网中的能量交易将是本文数据价值评估的应用场景。

下面以虚拟电厂作为实例对数据价值评估的方法进行说明。在虚拟电厂中, 可再生能源发电、需求侧响应资源^[26]均存在固有的不确定性, 这将使得虚拟电厂在进行电力调度或电力市场投标等决策的过程中面临较大的风险, 并可能直接导致虚拟电厂的经济损失。因此, 虚拟电厂有意愿通过数据交易获取关于可再生能源出力和负荷变化的更多微观信息以消除不确定性并辅助决策。

考虑到数据在消除不确定性方面的“语用”, 数据价值被定义为数据消除不确定性的作用对数据消费方经济利益的影响。因此, 能源互联网中的数据价值评估可以分为两个阶段进行, 如图4所示, 分别是: 阶段1: 基于数据交易, 定量分析数据对不确定性消除的作用; 阶段2: 基于能量交易, 定量分析不确定性的消除对利益主体经济利益的影响。

阶段1对不确定性的定量分析依据对不确定性的度量方式和数据与不确定性之间的关系。这一阶段不依赖于特定应用场景, 不确定性的消除是数据的直接“语用”。其中, 常用的不确定性表示方式包括区间数^[26-27]、模糊数^[28-29]、随机变量^[30-31]、不确

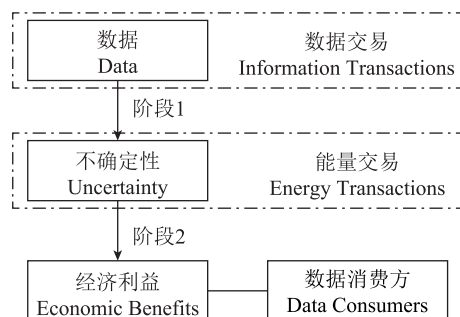


图 4 数据价值评估流程分解图

Fig.4 Process decomposition of data value evaluation

定性集^[32]等，相应的不确定性度量方式包括区间的长度、模糊数的隶属度、随机变量概率分布的标准差或熵、不确定性集的范围等。仍以虚拟电厂为例，首先需要确定所获取的信息映射到相应度量模型后对不确定性的影响，例如虚拟电厂获取额外的天气预报信息使得风电出力预测概率分布的分散度降低，或者获取个人用户行为信息使得负荷预测的置信区间长度减小。

阶段 2 是对经济利益的定量分析，需要构建不同数据消费方的收益模型，以及不确定性对收益影响的数学模型。这一阶段依赖能源互联网的应用场景，经济效益的提升是所交易数据的间接“语用”，也是能源互联网中数据价值评估结果。例如，对虚拟电厂而言，以考虑不确定性的风险优化模型为基础进行决策，那么就可以计算重要变量的不确定性程度变换对最终收益目标的影响；对于可控负荷、风电、光伏、储能装置等资源的调度模型，可以通过目标函数对风电或光伏出力预测方差的灵敏度分析，得到相应不确定性的减小为虚拟电厂带来的收益大小。根据不同的决策模型可以实现不同应用场景下的数据价值评估，具有良好的扩展性。

3 数据定价

3.1 研究意义

数据价值评估是数据交易的基础，但真正建立完善的数据交易架构不仅需要评估数据的价值，还需要通过机制设计约束数据交易市场行为，降低数据交易成本，激励能源互联网中的利益主体主动参与数据交易，这涉及在价值评估基础上的定价理论。

数据定价机制的不合理将导致部分数据所有方在数据交易市场中采取虚假报价等策略性行为为自身谋取利益，由于市场力的存在，这种行为将严重影响数据交易市场效率。因此，设计合理、有效的

数据定价机制对于数据交易市场的良好运行具有重要意义。

数据交易和普通商品交易的一大区别在于普通商品作为物理实体，存在特定的度量单位，在交易时可以叠加，而且互相之间不存在相关性；然而对于数据而言，没有统一的单位度量，不同数据的效用可能存在重复，无法直接叠加，而且不同数据之间存在相关关系，一种数据的部分价值可以由其他数据所替代。因此，在设计数据定价机制的过程中，除了借鉴能量定价机制之外，还需要考虑这些区别带来的影响。

3.2 数据定价的机制设计

3.2.1 数据交易市场运行模式

由于分布式数据交易往往通过区块链技术中的智能合约进行出清，较少涉及数据交易市场中不同主体之间的竞争，因此在数据定价的问题中，本文将以集中式数据交易为背景进行说明。

在完全竞争的数据交易市场中，能源互联网中的利益主体（数据消费方）按照能量交易的模型得到风险期望成本与不确定性程度的关系曲线，并根据模型生成数据需求曲线；信息供应商（数据所有方）则以数据的获取成本和技术成本等为依据提供数据供给能力。数据交易平台将根据数据消费方提供的数据需求曲线和数据所有方提供的数据供给能力，以社会福利最大化为目标进行市场出清。在市场出清后，数据交易平台将向中标的数据所有方支付相应的数据收益。另外，在各个利益主体的数据信息由相互割裂走向共享交易的过程中，不可避免地会出现少数数据所有方对部分数据存在垄断的局面。因此，在完全垄断或寡头垄断的市场环境下的数据定价机制也是值得考虑的。与此同时，健全的数据交易监管体系也应逐步建立，以维护良好的数据交易环境。

3.2.2 数据定价机制设计

借鉴电力交易市场的机制，在完全竞争的数据交易市场中，主要可以采用拍卖定价或以 VCG(Vickrey-Clarke-Groves)机制为基础的数据定价机制。

拍卖定价是一种较为普遍的定价机制，同样可以用于数据定价。由于海量的非结构化数据无法划分至特定的应用场景中，因此数据价值评估难以直接实现，利益主体需要通过集中竞价的方式，以自身对于数据价值的衡量参与数据交易。拍卖定价的优点在于适用范围广，与现行的能量交易市场的定

价机制相似;缺点是难以激励利益主体进行真实报价,容易导致部分主体通过虚假报价使市场出清结果偏离真实情况,妨碍市场竞争。

针对拍卖定价中市场机制难以实现激励相容的困境,可以设计以 VCG 机制为基础的数据定价机制激励利益主体提供真实数据报价。

VCG 机制是一种满足优势策略激励相容的机制^[33],在能源互联网中风电、储能、电网与气网耦合、需求侧资源管理等多个方向的机制设计层面得到了广泛应用^[34-37]。在以 VCG 机制为基础的数据定价机制中,市场将数据所有方提供的数据为数据消费方带来的经济效益的提升换算为报酬,支付给数据所有方。该机制可以满足激励相容、个体理性和社会福利最大化,充分激励数据所有方提交真实数据成本。然而,该机制的局限性在于需要依赖准确的数据价值评估结果,往往适用于在特定应用场景中进行的数据交易,适用范围较小。

因此,在数据价值明确的场景中,以 VCG 机制为基础的数据定价机制具有明显的优越性;而在数据价值无法得到精确评估或数据供求关系不清晰的场景中,拍卖定价将是一种更为灵活、适用性更好的数据定价机制^[13]。

另外,在完全垄断或寡头垄断的数据交易市场中,数据所有方可以采用价格歧视等多种定价手段,对不同的数据消费方或在不同时间段内设置不同的数据价格,从而获取额外的收益。

3.2.3 数据定价机制的特点

数据定价机制与能量交易机制相似,但是由于数据本身的特性仍然存在较为明显的区别。

(1) 由于数据的不可叠加性,本文认为数据所有方针对每次数据交易,将提供数据供给能力点的一维供给信息,而无法得到数据供给曲线的二维供给信息。

(2) 由于数据的相关性,本文认为数据消费方获取了部分信息后可能依据所获取信息估计得到更多信息,因此若数据消费方所具有的数据信息不同,其在参与数据交易时面临的数据供给能力也可能存在不同。

(3) 由于数据难以划分单位、实现度量,仅通过信息熵或其他表征不确定性的参数对数据所蕴含的信息量的描述难以适应所有数据消费方的需求,因此本文认为应对数据交易市场划分应用场景,使得数据交易主要在相似的应用场景条件下进行,跨场景的数据交易则需要在定价机制中考虑场景转换带来的附加价值。

4 数据确权

4.1 研究意义

数据不同于普通商品,可以被广泛地复制和共享^[6],具有绝对的非排他性^[16]。因此,数据交易和普通商品交易的另一个区别是:普通商品交易使得商品的所有权由所有方转移到消费方,而数据交易则将数据的所有权由所有方复制到消费方,数据所有方并不会失去数据所有权。

然而,数据交易中数据所有权的可复制性使数据的权属不清晰。在数据交易发生后,数据消费方能够以数据所有方的身份与其他数据消费方进行更多的数据交易,从而可能侵犯了原数据所有方的权益,对于原数据所有方而言无异于数据被盗用,从而极易导致数据交易市场失调。因此,在数据交易的过程中,数据确权问题需要得到妥善处理。

数据确权是数据交易机制建立与发展的根本问题,其核心是确定数据相关权利人,即明确数据所有权、使用权等权利的归属,从而保护相应权利人的合法权益^[38],保障数据交易程序的公平。

4.2 区块链技术与数字水印技术的应用

区块链技术起源于比特币,是比特币作为虚拟货币发行和流通的基础。近年来,区块链技术本质上作为一种基于密码学方法的去中心化分布式记账技术,被应用于更广阔的领域。数字水印技术是一种非密码机制信息隐藏技术,通过将相关标识嵌入数据中以保障信息安全。本文将应用区块链技术和数字水印技术解决数据确权问题,其中区块链技术用于对数据进行定位和追溯,数字水印技术用于确认数据的具体权属。

4.2.1 数据交易与区块链技术的结合

区块链技术的核心是通过去中心化建立所有交易主体之间对于交易内容的共识,打破不同交易主体之间的信任壁垒,保障交易内容的真实可信,从而解决部分信息不对称的问题,降低主体之间的信任成本,提高数据交易效率。

区块链中的关键技术包括分布式网络、共识机制、加密算法和智能合约等。在将区块链技术应用用于数据交易的过程中,首先在参与数据交易的数据所有方和数据消费方之间形成分布式网络连接,保证所有网络节点的地位均等。在数据交易的过程中,基于区块链的智能合约可以保障基于预设条件自然形成买卖协议,从而实现数据交易和结算的自动化。市场出清后,为了防止数据被泄露,数据所有方和消费方之间往往对数据采用加密算法处理后再进行传

输。一段时间内的数据交易结果经过共识机制的验证和确认后将记录在区块中，各个区块按照时间顺序逐个连接成不可篡改的区块链。最终，区块链将以副本的形式在分布式网络的各个节点上分别储存^[39]。

4.2.2 数字水印技术用于数据确权

目前，较为常用的数据确权方式是专家评审^[40]，但是这种方式在部分情况下难以保证公平性。因此基于计算机算法的数字水印技术^[40]是一种较好的数据确权手段。在数据交易前，数据所有方可以通过计算机算法生成水印，并将水印嵌入待交易的数据模块中，即对数据资产完成了标识。

数字水印往往具有较好的隐蔽性、鲁棒性，不仅难以被感知或篡改，而且可以抵御多种形式的数字攻击^[41]。因此，在出现数据权属的纠纷时，对数字水印的提取和检测可以快速确定数据的权属。

4.2.3 数据确权的实现

本文认为在能源互联网中的数据交易背景下，尽管数据所有方和数据消费方之间通过数据交易实现数据的传输和流通，但是数据的所有权不应该在数据交易的主体间进行转移，如图 5 所示。

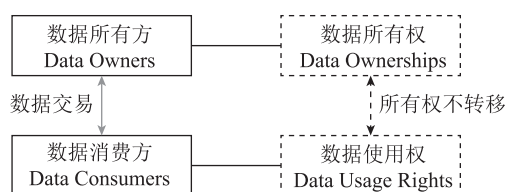


图 5 数据交易中的权属

Fig.5 Right ownership in data transactions

数据消费方在数据交易完成后将获得数据的使用权，也可以通过对数据进行分析和挖掘生产更具价值的数据产品。然而，为了保障数据的所有权及相应的收益权，数据消费方不得冒充数据所有方，将购得的数据直接用于另外的交易。由于数字水印技术对数据资产进行了标识、区块链技术对数据交易的历史结果进行了登记和储存而且内容不可篡改，因此数据资产的来源和具体流通路径是可追溯的，从而充分保障了数据所有方对数据的所有权且不被侵犯。

5 数据隐私保护

5.1 研究意义

数据隐私保护问题属于信息安全问题，研究数据安全共享与交易的机制和技术，旨在解决数据安全、个人隐私、企业保密数据等敏感问题。如何在泄露用户隐私的前提下，提高大数据的利用率，挖掘大数据的价值，是目前大数据研究领域的关键

挑战。目前能源互联网中各个主体之间的数据之所以互相割裂、无法流通，其重要原因是数据隐私保护问题无法得到解决，数据交易带来的收益无法弥补隐私暴露导致的严重后果。因此，数据隐私保护是能源互联网数据交易的必要环节。

5.2 数据隐私保护技术的应用

5.2.1 通用数据安全技术

数据安全的主要内容包括数据保密性与安全共享、身份认证以及访问控制等^[42]。现有的数据保密性和安全共享方案通常采用基于非语义的通用加密技术来实现。应用于数据交易，其常规流程是由数据所有方预先对外包数据进行加密处理和上传操作，并在需要时由数据使用者解密^[43]。身份认证的主要研究内容包括单一域内身份认证、跨域认证和切换认证，主要用于解决每个实体的身份分配问题。在数据交易过程中，各实体首先要通过授权中心的安全认证才能够获取存储和计算等服务^[44]。设计具有隐私保护特性的身份认证协议也是数据交易安全的关键技术和方法，包括基于属性^[45]和基于角色^[46]的访问控制。

从对数据隐私保护的视角看，现有的数据交易架构可分为三类：“移动计算、不移动数据”、“不移动计算、移动数据”和“移动计算、移动数据”。

5.2.2 移动计算，不移动数据

移动计算、不移动数据的交易架构适用于存在多个数据所有方的数据交易。这种数据交易存在明显的网络异构化，对数据的安全性、横向拓展性和动态性有较高要求，对数据的分析性能也有一定要求。由于各数据所有方的数据管理需要，不宜进行数据集中，往往采用虚拟数据中心的形式实现数据信息的集散，因此需要移动计算。基于虚拟任务整合的数据任务机制如图 6 所示。

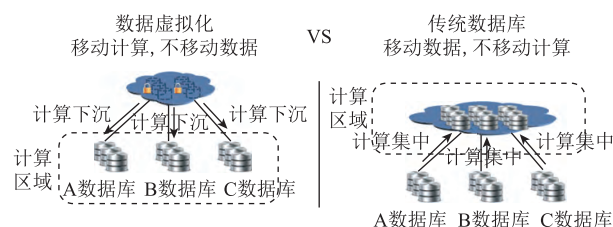


图 6 基于虚拟数据整合的计算任务机制^[48]

Fig.6 The computing mechanism based on virtual data integration^[48]

移动计算、不移动数据技术体系架构的技术实现主要依赖边缘计算、数据虚拟化、区块链等技术。

通过数据虚拟化技术,实现大规模异地、异构数据源动态环境下虚拟整合,以“虚拟数据库^[47]”形态进行呈现。通过边缘计算、GPU 加速等技术赋能,支持包括 SQL JOIN、机器学习在内的复杂数据分析,获得与数据物理集中类似的效果^[48]。在这种技术体系下,最终数据消费方的数据结构对数据所有方不可见,无法知悉细节,但是可以通过平台层数据二次加工获得数据使用权。这其中保证计算任务的可分解性是关键。

基于“移动计算、不移动数据”技术体系架构,研究能源互联网大数据的多源接入、整合、调度和计算技术,有望消除数据所有方关于数据安全的顾虑,实现数据所有方对数据资产的持续变现的意愿与数据消费方对降低获取数据资源成本的诉求,最终推动能源领域数据交易与共享生态的建立。

5.2.3 移动数据,不移动计算

移动数据、不移动计算的数据交易架构适用于单一数据所有方与多个数据消费方的数据交易,典型的实现方式为基于“云计算”的技术实现。这种数据交易存在数据所有权集中的特点,对数据分析性能有较高的要求,然而对数据的安全性、横向拓展性和动态性要求不高。比如企业的私有云,或者将数据所有权规约于一方的集团公司。由于数据所有权的统一和众多数据消费方的不同需求,因此需要移动数据。

云计算是融合分布式计算、并行计算、效用计算等多种计算机技术的计算服务模式^[49]。其中,云服务提供商对于用户数据的存储将是用户数据隐私泄露隐患的主要来源,需要通过法律和技术手段予以保障。另外,云计算的动态运行会使数据分散于多种网络介质中^[49],这将进一步提高数据隐私保护的难度。

在云计算环境中,属性加密技术和可搜索加密技术是两种数据隐私保护的关键技术^[49]。属性加密技术通过对用户的属性进行加密,密钥来源于访问结构;可搜索加密技术则依赖对数据关键词的搜索。由于云计算作为典型的移动数据的计算方式,关于数据的加密技术可以通过密钥使得数据在所有者和使用者之间互通从而避免被云端服务器获取,因此基于云计算的技术成为移动数据、不移动计算技术体系架构中的核心内容。

5.2.4 移动数据,移动计算

典型的实现方式包括多方安全计算框架。多方安全计算理论是在保护隐私信息以及没有可信第三方的前提下,解决一组互不信任的参与方如何进行

协同计算的问题^[50]。多方安全计算的核心特征是通过数学理论保证参与计算的各方成员输入信息不被暴露的同时,确保计算任务结果的正确性。

在数据交易环境下,企业或个人将数据视为一种重要的资产在数据市场上进行交易。在这种多信任域共存的数据交易场景中,通过多方安全计算技术,数据所有方可以规定数据的用途、用量、有效期等使用属性。多方安全计算技术将数据交易的本质由数据所有权转向数据使用权,保障原始数据所有者的权益,有效遏制原始数据泄漏,并降低数据泄漏引起的数据流通风险,有助于促进数据的大规模应用,建立良好的数据交易生态。

移动数据、移动计算的数据交易架构对数据的安全性、横向拓展性和动态性有较高要求,适用于需要多方利益主体进行较为复杂互动与计算的场景,比如联合的协同优化计算等。多方安全计算是通过较为高昂的计算代价换取信息隐私保护的优势,在兼容异构化网络方面尚有挑战。目前还有大量理论工作和工程实践需要突破。

5.2.5 其他技术

其他数据隐私保护技术包括特征提取与数据生成技术和面向信息语义的伪装保护技术等。

不同数据交易任务对数据有不同的处理方式,可能涉及直接获取实际的数据样本或者从数据挖掘的角度获取卖方数据的统计特征等^[51-52],利用特征提取与数据生成技术可以不暴露原始数据而获取其中最主要的特征。以用户用能数据交易为例,原始用电数据可被用于分析用户各类生活习惯及家庭状况。而真实用电负荷是作为用户隐私的直观反映,因此需要对真实的用能曲线加以修饰混淆,或是通过特征提取反向映射后,利用对抗网络等人工智能方法,生成近似信息用于抹除用户个人隐私,防止用电数据直接泄露客户隐私造成风险^[53]。

信息伪装技术是一类数据脱敏技术。通过信息伪装可削弱私有信息的安全属性,在不影响计算任务的前提下将私有信息转换为伪装信息,此时伪装信息具有保密属性,可由第三方间接访问^[54]。面向数据交易与应用研究信息伪装算法可以与当前通用的加密等手段兼容,作为当前信息加密等安全防护手段的有机补充。信息伪装的实现并不依赖额外的安全措施或设备,而是通过对决策过程中相关信息的数学变换,在保证其核心价值不变的基础上,对信息进行语义层面的伪装,从而实现对关键隐私信息的保护^[55]。

6 总结

能源互联网是能源系统与信息系统高度融合的结果，数据交易是大数据等信息技术发展 to 一定程度的产物。而能源互联网中的数据交易将进一步加强能源与信息的联系，一方面提高能源利用效率，另一方面提高信息传输价值，是能源互联网研究中的重要课题。本文对能源互联网中的数据交易体系架构进行了初步建构，对数据交易相关问题进行了探讨，为数据交易技术的实现和发展提供了思路。目前，笔者也已经在广东电网开始了数据交易的有益尝试。

数据交易作为信息领域的新兴方向，还有许多关键问题和技术有待进一步解决和突破。随着能源互联网的不断演化，会出现更多复杂的场景和需求，能源互联网数据交易技术的落地与应用依然任重道远。

参考文献

- [1] Liu Yue, He Jia, Guo Minjie, et al. An overview of big data industry in China[J]. China Communications, 2014, 11(12): 1-10.
- [2] 申建建, 曹瑞, 苏承国, 等. 水火风光多源发电调度系统大数据平台架构及关键技术[J]. 中国电机工程学报, 2019, 39(1): 43-55.
Shen Jianjian, Cao Rui, Su Chengguo, et al. Big data platform architecture and key techniques of power generation scheduling for hydro-thermal-wind-solar hybrid system[J]. Proceedings of the CSEE, 2019, 39(1): 43-55.
- [3] 孙宏斌, 郭庆来, 潘昭光. 能源互联网：理念、架构与前沿展望[J]. 电力系统自动化, 2015, 39(19): 1-8.
Sun Hongbin, Guo Qinglai, Pan Zhaoguang. Energy internet: concept, architecture and frontier outlook[J]. Automation of Electric Power Systems, 2015, 39(19): 1-8.
- [4] 孙宏斌, 郭庆来, 潘昭光, 等. 能源互联网：驱动力、评述与展望[J]. 电网技术, 2015, 39(11): 3005-3013.
Sun Hongbin, Guo Qinglai, Pan Zhaoguang, et al. Energy internet: driving force, review and outlook[J]. Power System Technology, 2015, 39(11): 3005-3013.
- [5] Shannon C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3): 379-423.
- [6] 钟义信. 信息科学原理[M]. 北京：北京邮电大学出版社, 2013.
- [7] 康重庆, 姚良忠. 高比例可再生能源电力系统的关键科学问题与理论研究框架[J]. 电力系统自动化, 2017, 41(9): 2-11.
Kang Chongqing, Yao Liangzhong. Key scientific issues and theoretical research framework for power systems with high proportion of renewable energy[J]. Automation of Electric Power Systems, 2017, 41(9): 2-11.
- [8] 姚良忠, 朱凌志, 周明, 等. 高比例可再生能源电力系统的协同优化运行技术展望[J]. 电力系统自动化, 2017, 41(9): 36-43.
Yao Liangzhong, Zhu Lingzhi, Zhou Ming, et al. Prospects of coordination and optimization for power systems with high proportion of renewable energy[J]. Automation of Electric Power Systems, 2017, 41(9): 36-43.
- [9] 鲁宗相, 黄瀚, 单葆国, 等. 高比例可再生能源电力系统结构形态演化及电力预测展望[J]. 电力系统自动化, 2017, 41(9): 12-18.
Lu Zongxiang, Huang Han, Shan Baoguo, et al. Morphological evolution model and power forecasting prospect of future electric power systems with high proportion of renewable energy[J]. Automation of Electric Power Systems, 2017, 41(9): 12-18.
- [10] 程耀华, 张宁, 王佳明, 等. 面向高比例可再生能源并网的输电网规划方案综合评价[J]. 电力系统自动化, 2019, 43(3): 33-42.
Cheng Yaohua, Zhang Ning, Wang Jiaming, et al. Comprehensive evaluation of transmission network planning for integration of high-penetration renewable energy[J]. Automation of Electric Power Systems, 2019, 43(3): 33-42.
- [11] 贵阳大数据交易所. 2016 年中国大数据交易产业白皮书[R]. 贵阳, 2016.
- [12] 中国信息通信研究院. 大数据白皮书(2018 年)[R]. 北京, 2018.
- [13] 郭春芳. 不确定性分析视角下大数据信息服务定价研究[D]. 北京：北京交通大学, 2019.
- [14] 倪渊, 李子峰, 张健. 基于 AGA-BP 神经网络的网络平台交易环境下数据资源价值评估研究[J]. 情报理论与实践, 2019, 43(1): 135-142.
Ni Yuan, Li Zifeng, Zhang Jian. Research on data resources value assessment model based on AGA-BP neural network in the background of network platform transaction[J]. Information Studies: Theory & Application, 2019, 43(1): 135-142.
- [15] 宋立丰, 宋远方, 国潇丹. 基于数据权的现实与虚

- 拟闲置资产共享——区块链视角下的共享经济发展研究[J]. 经济学家, 2019(8): 39-47.
- Song Lifeng, Song Yuanfang, Guo Xiaodan. Reality and virtual idle asset sharing based on data right-research on the development of shared economy from the perspective of blockchain[J]. Economist, 2019(8): 39-47.
- [16] 吴江. 数据交易机制初探——新制度经济学的视角[J]. 天津商业大学学报, 2015, 35(3): 3-8.
- Wu Jiang. A preliminary study of the mechanism of data transaction—a view of new institutional economics[J]. Journal of Tianjin University of Commerce, 2015, 35(3): 3-8.
- [17] 朱宝丽. 数据产权界定: 多维视角与体系建构[J]. 法学论坛, 2019, 34(5): 78-86.
- Zhu Baoli. Definition of data property rights: multidimensional perspectives and system construction[J]. Legal Forum, 2019, 34(5): 78-86.
- [18] 史宇航. 数据交易法律问题研究[D]. 上海: 上海交通大学, 2017.
- [19] 韩璇, 袁勇, 王飞跃. 区块链安全问题: 研究现状与展望[J]. 自动化学报, 2019, 45(1): 206-225.
- Han Xuan, Yuan Yong, Wang Feiyue. Security problems on blockchain: the state of the art and future trends[J]. Acta Automatica Sinica, 2019, 45(1): 206-225.
- [20] 许重建, 李险峰. 区块链交易数据隐私保护方法[J]. 计算机科学, 2020, 49(3): 281-286.
- Xu Chongjian, Li Xianfeng. Data privacy protection method of block chain transaction[J]. Computer Science, 2020, 49(3): 281-286.
- [21] 王伟, 蒋菱, 王峥, 等. 基于向量评估遗传算法的智能电网大数据交易模型研究[J]. 电网与清洁能源, 2016, 32(10): 1-8.
- Wang Wei, Jiang Ling, Wang Zheng, et al. Trade model of smart grid big data based on vector evaluated genetic algorithm[J]. Power System and Clean Energy, 2016, 32(10): 1-8.
- [22] 奈存剑. 虚拟化数据管理平台研究[D]. 武汉: 华中科技大学, 2013.
- [23] Coase R H. The federal communications commission[J]. The Journal of Law & Economics, 1959, 2: 1-40.
- [24] Coase R H. The nature of the firm[J]. Economica, 1937, 4(16): 386-405.
- [25] Coase R H. The problem of social cost[J]. The Journal of Law & Economics, 1960, 3: 1-44.
- [26] Wang Bohong, Guo Qinglai, Yang Tianyu. From uncertainty elimination to profit enhancement: role of data in demand response[C]//2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), 2019: 2952-2957.
- [27] Chen Runze, Sun Hongbin, Guo Qinglai, et al. A generation-interval-based mechanism for managing the power generation uncertainties of variable generation[J]. IEEE Transactions on Sustainable Energy, 2016, 7(3): 1060-1070.
- [28] Huang Chaoming, Yang Hongtzer, Huan Ching-lieng. Bi-objective power dispatch using fuzzy satisfaction-maximizing decision approach[J]. IEEE Transactions on Power Systems, 1997, 12(4): 1715-1721.
- [29] Miranda V, Hang P S. Economic dispatch model with fuzzy wind constraints and attitudes of dispatchers[J]. IEEE Transactions on Power Systems, 2005, 20(4): 2143-2145.
- [30] Lee D, Shin H, Baldick R. Bivariate probabilistic wind power and real-time price forecasting and their applications to wind power bidding strategy development[J]. IEEE Transactions on Power Systems, 2018, 33(6): 6087-6097.
- [31] Bitar E Y, Rajagopal R, Khargonekar P P, et al. bringing wind energy to market[J]. IEEE Transactions on Power Systems, 2012, 27(3): 1225-1235.
- [32] Dvorkin Y, Lubin M, Backhaus S, et al. Uncertainty sets for wind power generation[J]. IEEE Transactions on Power Systems, 2016, 31(4): 3326-3327.
- [33] Narahari Y. Game theory and mechanism design[M]. Singapore: World Scientific, 2014.
- [34] Samadi P, Mohsenian-Rad H, Schober R, et al. Advanced demand side management for the future smart grid using mechanism design[J]. IEEE Transactions on Smart Grid, 2012, 3(3): 1170-1180.
- [35] Tang W, Jain R. Aggregating correlated wind power with full surplus extraction[J]. IEEE Transactions on Smart Grid, 2018, 9(6): 6030-6038.
- [36] Nekouei E, Alpcan T, Chattopadhyay D. Game-theoretic frameworks for demand response in electricity markets[J]. IEEE Transactions on Smart Grid, 2015, 6(2): 748-758.
- [37] Contreras-Ocaña J E, Ortega-Vazquez M A, Zhang B. Participation of an energy storage aggregator in electricity markets[J]. IEEE Transactions on Smart Grid, 2019, 10(2): 1171-1183.

- [38] 彭云. 大数据环境下数据确权问题研究[J]. 现代电信科技, 2016, 46(5): 17-20.
Peng Yun. Research on authenticating data rights in big data environment[J]. Modern Science & Technology of Telecommunications, 2016, 46(5): 17-20.
- [39] 于戈, 聂铁铮, 李晓华, 等. 区块链系统中的分布式数据管理技术-挑战与展望[J/OL]. 计算机学报, 2019: 1-27[2020-05-14]. <http://kns.cnki.net/kcms/detail/11.1826.tp.20191029.1604.004.html>.
Yu Ge, Nie Tiezheng, Li Xiaohua, et al. The challenge and prospect of distributed data management techniques in blockchain systems[J]. Chinese Journal of Computers, 2019: 1-27[2020-05-14]. <http://kns.cnki.net/kcms/detail/11.1826.tp.20191029.1604.004.html>.
- [40] 王海龙, 田有亮, 尹鑫. 基于区块链的大数据确权方案[J]. 计算机科学, 2018, 45(2): 15-19.
Wang Hailong, Tian Youliang, Yin Xin. Blockchain-based big data right confirmation scheme[J]. Computer Science, 2018, 45(2): 15-19.
- [41] 李兆璨, 王利明, 葛思江, 等. 基于正交编码的大数据纯文本水印方法[J]. 计算机科学, 2019, 46(12): 148-154.
Li Zhaocan, Wang Liming, Ge Sijiang, et al. Big data plain text watermarking based on orthogonal coding[J]. Computer Science, 2019, 46(12): 148-154.
- [42] 张佳乐, 赵彦超, 陈兵, 等. 边缘计算数据安全与隐私保护研究综述[J]. 通信学报, 2018, 39(3): 1-21.
Zhang Jiale, Zhao Yanchao, Chen Bing, et al. Survey on data security and privacy-preserving for the research of edge computing[J]. Journal on Communications, 2018, 39(3): 1-21.
- [43] 宋磊, 罗其亮, 罗毅, 等. 电力系统实时数据通信加密方案[J]. 电力系统自动化, 2004, 28(14): 76-81.
Song Lei, Luo Qiliang, Luo Yi, et al. Encryption on power systems real-time data communication[J]. Automation of Electric Power Systems, 2004, 28(14): 76-81.
- [44] 尹传烨. 电力信息系统安全策略应用研究[D]. 武汉: 华中科技大学, 2006.
- [45] 房梁, 殷丽华, 郭云川, 等. 基于属性的访问控制关键技术研究综述[J]. 计算机学报, 2017, 40(7): 1680-1698.
Fang Liang, Yin Lihua, Guo Yunchuan, et al. A survey of key technologies in attribute-based access control scheme[J]. Chinese Journal of Computers, 2017, 40(7): 1680-1698.
- [46] 王保义, 邱素改, 张少敏. 电力调度自动化系统中基于可信度的访问控制模型[J]. 电力系统自动化, 2012, 36(12): 76-81.
Wang Baoyi, Qiu Sugai, Zhang Shaomin. A credibility-based access control model in dispatching automation system[J]. Automation of Electric Power Systems, 2012, 36(12): 76-81.
- [47] Gupta A, Harinarayan V, Rajaraman A. Virtual database technology[C]//Proceedings 14th International Conference on Data Engineering, 1998: 297-301.
- [48] 移动计算, 不移动数据[Z/OL]. 即云天下(北京)数据科技有限公司. <http://www.bigknow.com.cn/>.
- [49] 霍成义. 云计算数据隐私保护关键技术研究[J]. 信息安全研究, 2019, 5(12): 1106-1109.
Huo Chengyi. Research on key technologies of cloud computing data privacy protection[J]. Journal of Information Security Research, 2019, 5(12): 1106-1109.
- [50] 孙茂华. 安全多方计算及其应用研究[D]. 北京: 北京邮电大学, 2013.
- [51] 罗永龙, 徐致云, 黄刘生. 安全多方的统计分析问题及其应用[J]. 计算机工程与应用, 2005, 41(24): 141-143.
Luo Yonglong, Xu Zhiyun, Huang Liusheng. Secure multi-party statistical analysis problems and their applications[J]. Computer Engineering and Applications, 2005, 41(24): 141-143.
- [52] 李锋. 面向数据挖掘的隐私保护方法研究[D]. 上海: 上海交通大学, 2008.
- [53] Lan J, Guo Q, Sun H. Demand side data generating based on conditional generative adversarial networks[J]. Energy Procedia, 2018, 152: 1188-1193.
- [54] Xin Shujun, Guo Qinglai, Wang Jianhui, et al. Information masking theory for data protection in future cloud-based energy management[J]. IEEE Transactions on Smart Grid, 2018, 9(6): 5664-5676.
- [55] 辛蜀骏. 面向能量管理系统的信息-物理耦合分析理论研究[D]. 北京: 清华大学, 2017.

作者简介

郭庆来 男, 1979 年生, 副教授, 博士生导师, 研究方向为能量管理、电压控制和信息物理系统。

E-mail: guoqinglai@tsinghua.edu.cn (通信作者)

王博弘 男, 1998 年生, 博士研究生, 研究方向为能源互联网中的数据交易与价值评估理论。

E-mail: wbh19@mails.tsinghua.edu.cn

(编辑 郭丽军)