

复杂软件系统的不确定性*

檀超^{1,2}, 张静宣^{1,2}, 王铁鑫^{1,2}, 岳涛^{1,2}

¹(南京航空航天大学 计算机科学与技术学院, 江苏 南京 211106)

²(高安全系统的软件开发与验证技术工业和信息化部重点实验室(南京航空航天大学), 江苏 南京 211106)

通信作者: 张静宣, E-mail: jxzhang@nuaa.edu.cn



摘要: 复杂软件系统(如信息物理系统 CPS、物联网 IoT 以及自适应软件系统等)在其开发和运行过程中会遇到各种类型的不确定性问题. 针对这些不确定性问题, 研究人员开展了大量的研究工作, 提出了一系列的方法, 取得了诸多成果. 然而, 由于此类系统本身固有的复杂性和其内在与外在不确定性的共同作用, 截止目前研究人员针对该研究领域仍然缺乏系统性和整体性的理解和分析. 为了深入探究该领域的研究现状, 采用系统研究的方法(systematic mapping study)识别出 142 篇相关文献, 并基于这些文献研究信息物理系统和物联网等系统生命周期中各个阶段和系统开发过程中产生的各种制品的不确定性及其处理方法. 通过对相关文献进行分析发现针对复杂系统的不确定性研究主要集中在其生命周期的设计定义、系统分析和运行等阶段. 此外, 首先将文献划分为 3 种不确定性类型, 包括外部不确定性、内部不确定性和传感器不确定性, 并将 142 篇相关论文关注的不确定性进行了分类. 为了深入探究不确定性, 将外部不确定性细分为环境不确定性、基础设施不确定性、用户行为不确定性以及经济属性不确定性, 将内部不确定性细分为系统结构不确定性、内部交互不确定性、支持系统运行的技术不确定性以及处理系统运行技术的不确定性. 针对复杂系统中的开发制品, 提出了对应的不确定性类型, 包括模型的不确定性、数据的不确定性和参数的不确定性等 8 类. 针对复杂系统的不确定性问题, 研究人员主要采用不确定性下的决策、不确定性推理和不确定性规约/建模等方法进行不确定性分析和处理. 基于文献分析结果, 进一步探讨和展望了该领域未来的研究趋势.

关键词: 不确定性; 系统研究; 信息物理系统; 物联网

中图法分类号: TP311

中文引用格式: 檀超, 张静宣, 王铁鑫, 岳涛. 复杂软件系统的不确定性. 软件学报, 2021, 32(7): 1926–1956. <http://www.jos.org.cn/1000-9825/6267.htm>

英文引用格式: Tan C, Zhang JX, Wang TX, Yue T. Uncertainty-wise software engineering of complex systems: A systematic mapping study. Ruan Jian Xue Bao/Journal of Software, 2021, 32(7): 1926–1956 (in Chinese). <http://www.jos.org.cn/1000-9825/6267.htm>

Uncertainty-wise Software Engineering of Complex Systems: A Systematic Mapping Study

TAN Chao^{1,2}, ZHANG Jing-Xuan^{1,2}, WANG Tie-Xin^{1,2}, YUE Tao^{1,2}

¹(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

²(Key Laboratory of Safety-Critical Software, Ministry of Industry and Information Technology (Nanjing University of Aeronautics and Astronautics), Nanjing 211106, China)

Abstract: Complex software systems (e.g., cyber-physical systems, Internet of Things, and adaptive software system) encounter various types of uncertainties in their different phases of development and operation. To handle these uncertainties, researchers have carried out a

* 基金项目: 国家自然科学基金(61872182)

Foundation item: National Natural Science Foundation of China (61872182)

本文由“面向非确定性的软件质量保障方法与技术”专题特约编辑陈俊洁副教授、汤恩义副教授、何啸副教授以及马晓星教授推荐.

收稿时间: 2020-09-15; 修改时间: 2020-10-26; 采用时间: 2020-12-14; jos 在线出版时间: 2021-01-22

lot of research work, proposed a series of methods, and achieved considerable results. However, there is still a lack of systematic understanding of the current state-of-the-art approaches. Motivated by this observation, this paper reports a systematic mapping study of 142 primary studies collected by following a rigorous literature review methodology. The scope of the study is about investigating on how the literature deals with uncertainties appearing in various phases or artifacts produced during a development lifecycle of cyber-physical systems and Internet of Things. Results show that uncertainties mainly appear in the phases of design definition, system analysis, and operation. Based on the 142 primary studies, uncertainties are first defined and classified into external uncertainty, internal uncertainty, and sensor uncertainty, and descriptive statistics are reported in terms of this classification. In order to explore the uncertainty in depth, external uncertainty is subdivided into environmental uncertainty, infrastructure uncertainty, user behavior uncertainty, and economic attribute uncertainty, and internal uncertainty is subdivided into uncertainty in system structure, internal interaction uncertainty, uncertainty in the technology supporting system operation, and uncertainty in the technology dealing with system operation. Furthermore, another classification is presented and descriptive statistics for those primary studies where uncertainties in eight different types of artifacts are discussed, including model uncertainty, data uncertainty, and parametric uncertainty. Results also show that researchers mainly focused on decision-making under uncertainty, uncertainty reasoning, and uncertainty specification/modeling when dealing with uncertainties. Based on the results, the future research trend is commented on in this area.

Key words: uncertainty; systematic mapping study; cyber-physical systems; Internet of Things

1 引言

随着信息化技术的不断发展,信息物理系统 CPS^[1]、物联网 IoT^[2]、智能系统^[3]等复杂软件系统在人们的生产和生活中扮演了越来越重要的角色.为了使复杂系统按照人们的预期进行工作,研究人员和实践者们通过为系统建模并构建预定义的规则来保证系统在给定输入情况下的正确输出.因此,复杂系统通常善于处理确定性的事务.然而,在复杂系统开发和运行过程中会遇到各种各样的不确定性(uncertainty).复杂系统的不确定性涉及其开发生命周期的各个阶段,涵盖大量的开发制品,从不同方面影响了系统的正常功能.由于不确定性具有未知性和不可预测性,通常会增大复杂系统的开发难度.

正确地理解和应对复杂系统开发中的不确定性是非常困难的,主要是因为此类系统本身具有的跨学科性质.其开发需要综合考虑 4 个方面的因素:控制(control)、平台(platform)、网络(internet)以及人(human).不确定性不仅存在于软件中,而且还存在于硬件、通信、人以及它们之间的交互中.此外,在 CPS 和 IoT 系统开发和运行过程中需要考虑到两个世界,即物理世界和逻辑世界,在工业应用中经常遇到软件(逻辑)和物理世界之间的方法和技术的 mismatch 问题.比如,物理系统中度量不确定性(measurement uncertainty)的方法非常成熟,然而度量计算不确定性的方法以及整合物理和逻辑度量不确定性的方法目前还处在研究阶段^[4].

为了解决复杂系统面临的不确定性问题,研究人员针对其具体的开发阶段和开发制品提出了相应的解决办法.近年来,随着该领域相关文献数量的逐年递增,研究人员分析和整理这些文献变得越来越困难.涉及到复杂系统不确定性的诸多问题尚未经过有效探索和系统性的研究,比如复杂系统在其不同的开发阶段会遇到哪些种类的不确定性因素?这些不确定性来自哪些方面?有哪些开发制品会含有不确定性信息?研究人员提出了哪些关于不确定性处理的方法?针对上述问题,本文围绕复杂系统面临的不确定性展开研究,旨在帮助研究人员快速了解该领域的研究现状,为他们进一步提出新的研究内容和技术方法提供一定的参考.同时,本文还进一步讨论了该领域未来的发展趋势,并系统且全面地对该领域进行了分析,希望能够吸引更多的研究者投入到这一重要的研究领域中来.

为了调研上述重要的研究问题,本文使用系统研究的方法从不同角度对复杂系统面临的不确定性进行分析.系统研究框架包括 6 个模块,即研究问题定义、相关文献搜索、文献筛选、文献质量评估与保障、属性抽取与细化、研究问题调研.在研究问题定义模块,我们首先定义了围绕复杂系统不确定性的 8 个研究问题,从不同的角度对不确定性进行调研.接着,在相关文献搜索模块,我们以回答这些研究问题为目标,定义了一系列的搜索字符串,并在常用的文献数据库(包括 IEEE Xplore、ACM Digital Library 等)进行搜索,共搜索到 6 831 篇文献.针对搜索到的文献集,我们在文献筛选模块定义了筛选标准,对文献集进行筛选,以获取相关文献.接着,我们定义了质量评估指标,对筛选后的文献进行质量评估,并通过文献质量评估进一步保留高质量的相关文献,最终

筛选出 142 篇高质量的相关文献.然后,为了更好地回答预定义的研究问题,我们提出并定义了关于不确定性的多个层面的分类标准(包括不确定性类型、开发制品的类型和不确定性处理方法的类型等).基于这些分类标准,我们从这些相关文献中抽取出一系列的文献属性,并对文献属性的一致性进行检查和校正.最后,以这些相关文献及其不同的属性为数据基础,调研并回答复杂系统不确定性相关的研究问题.

基于对 142 篇相关文献的系统性分析,我们观察到:(1) 复杂系统不确定性相关文献在 2015 年之后的发表数量呈不断上升的趋势,说明该领域逐渐成为一个研究热点;(2) 复杂系统不确定性相关的文献广泛分布在大量不同的期刊和会议上,反映了这个研究领域的多学科交叉的特点;(3) 在不确定性相关文献使用的关键词方面,使用频次最高的关键词为 *uncertainty*、*cyber-physical systems* 和 *Internet of Things*,这和我们设定的搜索字符串高度吻合;(4) 在不确定性出现的复杂系统开发生命周期的不同阶段,针对复杂系统的不确定性研究主要集中在其开发周期的设计定义、系统分析和运行等阶段;(5) 在不确定性应用的领域方面,研究人员主要关注信息物理系统、物联网和自适应信息物理系统等;(6) 在不确定性分类方面,研究人员主要关注复杂系统的外部不确定性,开发制品的不确定性类型主要集中在模型不确定性、数据不确定性和参数不确定性等上面;(7) 在应对不确定性方法方面,研究人员较多关注的是不确定条件下的决策、不确定性推理和不确定性规约/建模等;(8) 在自动化和软件工程领域关注的不确定性不同点方面,针对不确定性应用领域,软件工程主要关注信息物理系统、物联网和自适应信息物理系统等,自动化主要关注信息物理系统和物联网;针对不确定性类型,软件工程和自动化主要关注外部不确定性和内部不确定性;针对系统开发制品的不确定性类型,软件工程主要关注模型不确定性和数据不确定性等,自动化主要关注模型不确定性和参数不确定性等;针对不确定性处理方法,软件工程主要关注不确定性规约/建模、不确定性推理、不确定性条件下的决策和不确定性传播,自动化主要关注不确定性推理和不确定性条件下的决策等.

本文第 2 节介绍复杂系统不确定性的背景知识.第 3 节详细描述本文采用的系统分析方法.第 4 节重点调研预定义的研究问题.第 5 节讨论复杂系统的不确定性未来的研究趋势.第 6 节对全文进行总结.

2 背景知识和相关工作

不确定性通常被定义为“缺乏确定性(*lack of certainty*)”.针对不确定性的研究可以追溯到希腊哲学家亚里士多德在哲学范畴内对知识(*knowledge*)的确定性(*certainty*)的研究^[5].自此之后,不确定性在哲学、经济学等不同领域被赋予了不同的含义^[6].一种较常见且通用的不确定性分类是偶然(*aleatory*)和认识(*epistemic*)的不确定性.前者通常专注于评估事件发生的倾向性,通常通过频率来度量并且归因于随机行为;后者主要专注于一个事件的未来以及对错,通常通过对知识的信心(*confidence*)来度量并且归因于缺少信息或专业知识^[7,8].在过去的 10 多年中,不确定性,尤其是针对自适应系统的运行环境中的外部不确定性的研究引起了人们的关注^[9,10].比如,世界知名的信息物理系统领域专家 Manfred Broy 在与 Geisberger 共同编辑的 *agendaCPS* 中指出,复杂系统开发面临的一个重要挑战就是保证 CPS 在日渐不确定、不可预测、开放、互联的环境中运行的鲁棒性^[11].

近年来,工业复杂系统在不同工业领域广泛应用,复杂系统及相应的复杂性科学问题变得日益突出,这给软件工程学科带来了前所未有的挑战,如高复杂度、高动态、大规模、持续集成与部署等.随着快速发展的人工智能技术在越来越多的工业产品中得到应用与集成,现代及未来工业复杂系统的自身及其运行的环境会出现更多的无法避免的“客观不确定性”^[12].同时,由于系统的高复杂性及开发人员自身知识的局限性,各种假设被引入到系统的开发过程中,由此带来了“主观不确定性”^[13].由于工业复杂系统具有持续集成与持续部署等显著特征,在系统设计与开发阶段,开发人员无法完全预测未来的可能部署,从而引入了另外一个层面的主观与客观不确定性.此外,复杂系统的不确定性又经常被分为系统“内部不确定性”和“外部不确定性”(包括运行环境不确定性)^[14].其中,较为常见的外部不确定性是“传感器不确定性”.在系统开发过程中会产生很多的制品(如需求、系统架构模型、测试用例),这些制品同样可能包含不确定信息,尤其是人工开发的制品,如需求.由此可以看出,针对复杂系统的不确定性分类和特征化将是多角度和多维度的.鉴于此,本综述的主要目的是系统地了解针对复杂系统的不确定性研究现状,从而讨论和展望未来的研究趋势,为面向不确定性的复杂系统设计、开发、运行和

维护奠定基础。

下面,我们简要介绍针对不确定性进行系统文献综述的相关工作。Marinho 等人收集了 1994 年~2013 年之间发表的相关文献,并以此为依据针对软件项目管理中的不确定性进行了系统的文献调研^[15]。他们的研究目标是调研在软件项目中的不确定性的来源、等级和管理技术等。研究结果表明,采用新方法的项目和具有高风险的项目更有可能在其开发过程中遇到不确定性问题,大部分的不确定性出现在没有足够多的开发经验可以参考和创新的创新项目中。同时,他们还发现不确定性的来源包括市场的不确定性(比如不确定的客户、供应商、合作伙伴和当前市场状况)、技术的不确定性(比如现有技术的可用性和成熟程度)、内部和外部组织环境的不确定性(比如人口变化)、项目团队的认知复杂性和不确定性,以及他们与其他组织机构的关系。

Asan 和 Soyer 对不确定性下多准则风险优先级划分方法展开了文献调研,主要目标是识别和确定潜在的失效模式及其在不确定性下影响的优先级^[16]。这里的不确定性包括不完整、模糊和/或模棱两可的信息以及由于缺乏知识、模糊的评估和分级标准、碎片化的判断导致专家的主观判断缺少足够的信心。在该工作中,两位作者还确定并罗列了应对不确定性的方法(例如证据分析、普通模糊集)。他们的研究结果表明,不同的方法倾向于处理不同类型的不确定性,并且没有一种研究方法能够很好地处理所有类型的不确定性。比如,基于粗糙集理论的方法主要用于处理专家的主观判断中的不确定性。

Shevtson 等人针对控制理论软件的适应性进行了系统的文献调研,控制理论软件的适应性主要是为了处理不确定的操作条件,例如动态的服务可用性和软件目标的变化^[17]。他们收集了 42 篇相关文献,并展开调研。调研结果表明,尽管软件通常被认为是高度非线性的,但大多数研究都在使用线性模型。同时,他们也鼓励在控制理论的软件适应性方面进行跨学科研究(比如同时考虑软件工程和控制论)。但是,他们在文献[17]中并没有将不确定性作为主要的研究问题。

Salih 等人通过系统的文献综述将处理需求工程中不确定性的方法进行分类^[18]。他们收集了 167 篇相关文献,并以此为基础开展调研。调研结果表明,研究人员针对需求审查和分析中出现的确定性展开的研究最多,其次是面向目标的建模。在这项工作中,虽然在搜索字符串中明确定义了不确定性,但是 Salih 等人并没有提出针对不确定性的研究问题。

张等人针对机器学习赋能的软件系统自适应性开展了系统的文献调研工作^[19]。他们将自适应环境中的不确定性明确归类为触发自适应的原因之一,还提出了面向自适应软件系统设计和运行中的不确定性的全面分类,并以此为目标开展研究。基于收集到的 78 篇相关文献,他们针对软件系统自适应性的 3 个关注点(即场景不确定性、环境动态变化、效果不确定)进行了分析,并将这 3 个关注点与自适应策略生成与演化、参数评估和环境预测等自适应系统各项功能进行了关联,同时将这些自适应系统的各项功能进一步与不同的机器学习方法相关联。此外,他们还在文献[19]中详细介绍了目前该领域研究存在的问题,并对未来工作加以展望。

3 研究框架

由于目前复杂系统不确定性涉及的领域众多,且文献数量庞大,难以进行全面的搜索分析,因而难以采用系统文献综述(systematic literature review,简称 SLR)方法^[20]。本文调研的研究问题较为宽泛,更适合采用系统的调研(systematic mapping study)方法。系统调研常用于某一类研究搜索范围确定的领域,以得到某一主题的当前技术发展水平或实践水平的概括。因此,在本文中,我们使用系统的调研方法来调研复杂系统及其开发和运行中的不确定性。系统调研是一种通过分析文献来识别、评估和解释与某个研究兴趣相关内容的方法。Kitchenham 等人提出了被研究人员广泛使用的系统调研的具体准则^[21],我们按照该准则进行实验。根据这一准则的要求,系统调研包括 3 个阶段,分别为调研的规划(planning)、调研的实施(conducting)、调研的分析与展示(reporting)。以这 3 个阶段为根据,我们详细设计了本文的研究方法。如图 1 所示,本文的研究方法分成 6 个模块,分别为研究问题定义、相关文献搜索、文献筛选、文献质量评估与保障、属性抽取与细化以及研究问题调研。

具体来说,根据本文的研究目标,我们首先定义了一系列的研究问题,从不同角度调研针对复杂系统的不确定性问题。接着,以这些定义好的研究问题为依据,设定了一系列的搜索关键词及其同义词,并将这些搜索关键

词在主流的在线文献数据库中进行搜索.然后,针对这些在线文献数据库搜索到的文献,设置删除和保留标准,用来筛选与本文的研究目标相关的文献.接着,为了提高分析结果的有效性,我们为文献设定了一系列的质量评估标准,并以此为依据为筛选后的文献进行质量评估,以选取文献质量较高的文献作为最终的参考文献集.然后,我们定义了与研究问题相关的分类标准并抽取和清洗对应的文献属性.最后,结合我们获取的相关文献和每个文献对应的属性,调研各个研究问题的结果,并分析这些结果对复杂系统的不确定性问题的影响.

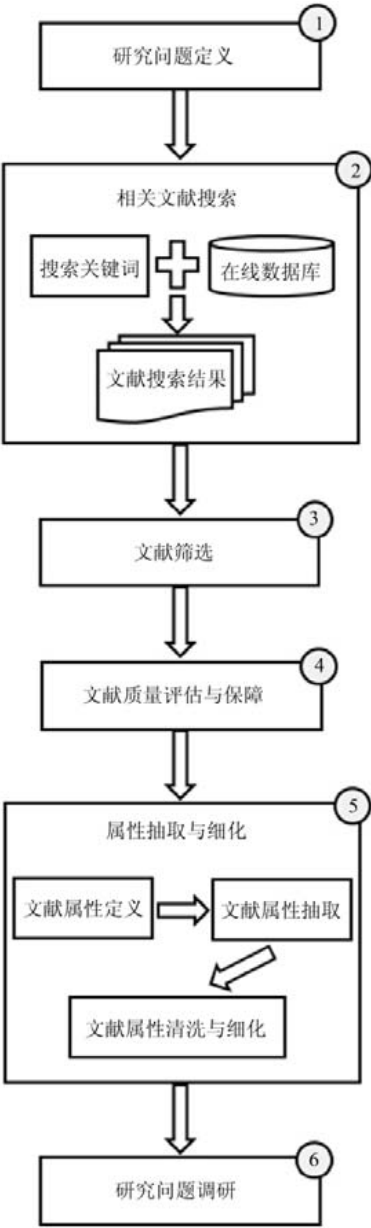


Fig.1 The research method applied for conducting the systematic mapping study
图 1 系统调研研究方法

3.1 研究问题

本文的总体研究目标是调研和分析复杂软件系统不确定性研究的各个方面现有的研究成果.基于此目标,

在本文中,我们对以下 8 个问题展开研究.

(1) 复杂系统的不确定性研究的相关文献随时间推移的发表情况是怎样的?通过研究此问题,我们可以了解该领域相关文献的发表数量随着时间的演变情况,以便掌握该领域的整体研究趋势.

(2) 复杂系统的不确定性研究的相关文献发表在哪些刊物上?文献发表数量排名靠前的期刊和会议都有哪些?通过研究此问题,我们可以了解该领域文献发表的相关期刊或者会议,为后续的研究人员的投稿刊物的选择提供参考.

(3) 复杂系统的不确定性研究具有哪些高频关键词以及它们之间的关系是怎样的?研究人员通常会在文献中定义一些关键词来反映该文献的主题概念.通过研究此问题,我们可以了解复杂系统不确定性研究的主题概念主要集中在哪些关键词,在此基础上进一步了解关键词之间的共现关系.

(4) 复杂系统的不确定性相关文献关注于系统开发生命周期的哪些阶段?通过此研究问题,我们可以进一步明确复杂系统在开发过程中的哪些阶段会遇到不确定性问题.

(5) 不确定性的研究主要是针对哪些复杂系统?通过研究此问题,我们可以进一步了解研究人员关注和解决哪些复杂系统的不确定性.

(6) 复杂系统不确定性可分成哪些类型?通过此研究问题,我们可以发现复杂系统不确定性具有哪些类型,以供研究人员参考.

(7) 针对复杂系统的不确定性,研究人员提出了哪些不确定性的处理方法?通过此研究问题,我们可以总结出处理不确定性的常用方法,并为研究人员提出新的方法提供借鉴.

(8) 针对复杂系统的不确定性,自动化和计算机领域(尤其是软件工程)的研究,各有哪些关注点?是否有交叉?通过此研究问题,我们可以了解以上两个领域关于不确定性研究关注点的差异,为后续这两个领域研究人员进行学科交叉研究提供参考依据.

3.2 相关文献搜索

我们遵循了 Kitchenham 和 Charters 等人提出的 PICO(Population, Intervention, Comparison and Outcomes) 结构^[21],并根据研究问题确定搜索关键词.同时,我们在确定搜索关键词时考虑了关键词的适用范围,以尽可能地将所有相关的文献都检索到,避免遗漏.我们设定的关键词包括以下 3 个方面:(1) 搜索关键词须为不确定性的描述词汇,包括 uncertainty、unpredictability、indeterminacy 以及 non-determinism;(2) 搜索关键词需要覆盖复杂软件系统生命周期的各个阶段,包括 requirement、modeling、analysis、architecture、design、testing、development、specification、operation、verification 以及 validation;(3) 搜索关键词包括的应用领域为信息物理系统(cyber-physical systems)和物联网(Internet of Things).同时,我们在搜索相关文献时发现有些文献并没有直接提及上述两种复杂系统,而是以智能系统(smart system)作为统称进行研究.因此,为了尽可能地搜索到所有相关文献,我们将 cyber-physical systems、Internet of Things 以及 smart system 作为搜索关键词.

通过调研在软件工程领域的系统研究中经常使用的文献数据库^[22-25],我们确定使用如下 4 个文献数据库作为文献数据来源:IEEE Xplore、ACM Digital Library、ScienceDirect 和 Springer.由于不同的文献数据库有不同的搜索方式,我们针对每一个文献数据库的搜索方式,将 3 个方面的搜索关键词进行不同方式的组合,以形成最终的搜索字符串.具体的文献数据库及其对应的搜索字符串详见表 1.由于 ScienceDirect 文献数据库搜索引擎对搜索字符串的数量有一定的限制,我们将搜索字符串分成 3 个子集进行搜索.对于 Springer,我们将搜索范围限制为计算机科学和自动控制学,在计算机科学侧重于两个子学科,即软件工程和人工智能(包括机器人学).

我们利用制定好的搜索字符串分别在这 4 个文献数据库进行搜索.在对应的文献数据库搜索相应的字符串时,限定文献的发表时间是在最近 10 年,即 2011 年 1 月~2020 年 3 月之间,以保证我们得到的文献分析结果是基于最新的前沿研究.我们在 IEEE Xplore 文献数据库进行搜索获得文献 192 篇,在 ACM Digital Library 文献数据库搜索获得 399 篇文献,在 ScienceDirect 文献数据库搜索获得 4 497 篇文献,在 Springer 文献数据库搜索获得 1 743 篇文献,共计 6 831 篇文献.我们以这 6 831 篇文献作为原始依据进行分析.

Table 1 The search strings used in the four digital libraries
表 1 本文使用的文献数据库及其对应的搜索字符串

文献数据库	搜索字符串
IEEE Xplore	(("Abstract":uncertainty 或 "Abstract":unpredictability 或 "Abstract":indeterminacy 或 "Abstract":non-determinism) 和 ("Abstract":requirement 或 "Abstract":modeling 或 "Abstract":analysis 或 "Abstract":architecture 或 "Abstract":design 或 "Abstract":testing 或 "Abstract":development 或 "Abstract":specification 或 "Abstract":operation 或 "Abstract":verification 或 "Abstract":validation) 和 ("Abstract": "cyber-physical systems"或"Abstract": "internet of things"或"Abstract": "smart system"))
ACM Digital Library	recordAbstract:(("uncertainty 或 unpredictability 或 indeterminacy 或 non-determinism)和(requirement 或 modeling 或 analysis 或 architecture 或 design 或 testing 或 development 或 specification 或 operation 或 verification 或 validation)和("cyber-physical systems"或"internet of things"或"smart system"))
ScienceDirect	(uncertainty 或 unpredictability 或 indeterminacy 或 non-determinism)和(modeling 或 analysis 或 architecture) 和 ("cyber-physical systems" 或 "internet of things" 或 "smart system")(uncertainty 或 unpredictability 或 indeterminacy 或 non-determinism) 和 (design 或 testing 或 development 或 specification) 和 ("cyber-physical systems" 或 "internet of things" 或 "smart system")(uncertainty 或 unpredictability 或 indeterminacy 或 non-determinism)和(operation 或 verification 或 validation)和("cyber-physical systems"或"internet of things"或"smart system")
Springer	(uncertainty 或 unpredictability 或 indeterminacy 或 non-determinism)和(requirement 或 modeling 或 analysis 或 architecture 或 design 或 testing 或 development 或 specification 或 operation 或 verification 或 validation)和("cyber-physical systems"或"internet of things"或"smart system")

3.3 文献筛选

我们针对这些搜索到的文献,设计了一系列的筛选标准,用来删除不符合预定义标准的文献,保留相关文献.经过对文献进行筛选,可以保证筛选后的文献与本文的研究主题相关,有助于回答设定好的研究问题(详见第 3.1 节).我们首先设计了一系列能够自动判断的删除标准,分别为:

- (1.1) 该文献是否是英文文献?如果不是,则删除;
- (1.2) 该文献是否存在重复的文献?如果有,则删除重复;
- (1.3) 该文献是否可以全文下载?如果不可以,则删除;
- (1.4) 该文献的长度是否不少于 4 页?如果不是,则删除.

同时,为了保证文献确实与本文的研究目标相关,我们还设计了一个需要人工判断的文献筛选标准.

- (2.1) 该文献是否研究针对复杂系统的不确定性问题?如果是,则保留.

针对检索到的原始文献数据,我们首先依次通过工具应用筛选标准 1.1~1.4,再应用人工筛选标准 2.1.如果某个文献同时满足上述的所有标准,则保留.反之,则删除.由于人工选择标准的判定比较主观,为了保证文献筛选的质量和有效性,我们按照如下步骤进行文献筛选.

- (1) 同一篇文献由本文的两位不同作者同时进行判定,两位作者不知道对方的判定结果.
- (2) 如果他们得到了一致的判定结果,则该文献的最终结果为两位本文作者的一致结果.
- (3) 如果他们得到了不同的判定结果,则两位本文的作者需要进行协商以得到一致的结果.

3.4 文献质量评估与保障

针对筛选后的文献,我们从质量报告、严谨性、可信性和相关性 4 个维度对其进行质量评估^[26].通过详细调研相关工作,我们设计了 12 个文献质量指标,见表 2,针对每一篇文献从以下 4 个维度对其质量进行评估.

- (1) 质量报告,主要涉及到研究的相关原理、目标和背景的报告质量,涉及到的文献质量指标为 Q1~Q4.
- (2) 严谨性,主要涉及到研究方法、研究设计、分析方法的严谨性、相应的工具支持以及收集足够的项目数据集和案例,涉及到的文献质量指标为 Q5~Q9.
- (3) 可信性,该维度主要为了确保文献相关结论的有效性以及具有一定的意义,涉及到的文献质量指标为 Q10 和 Q11.
- (4) 相关性,该维度为了探讨文献是否对工业界和学术界有价值,涉及到的文献质量指标为 Q12.

针对每一个质量指标,我们制定了对应的评分标准.其中 3 个文献质量指标(Q2、Q8 和 Q9)采用基于二元度

量的质量得分值.如果回答“是(Y)”,那么回答这个问题的得分为 1 分;如果回答“否(N)”,那么回答这个问题的得分是 0 分.其中 8 个文献质量指标(Q1、Q3、Q4、Q5、Q6、Q10、Q11 和 Q12)如果不能充分且自信地回答“是”或者“否”,那么回答这个问题的得分是 0.5 分.其中一个文献质量指标(Q7)中度量文献利用数据集和案例进行验证的数量.考虑到当数据集或者案例大于 1 时,更有说服力,所以制定:在该质量指标中,当只有一个数据集或者案例时,该问题的评分为 1 分;当有多个数据集或者案例时,该问题的评分为 2 分;当没有数据集或者案例时,该问题的评分为 0 分.

当获得每个文献的所有质量指标得分之后,我们计算这 12 个问题的回答得分的总和来评估该文献的质量.在进行文献质量评估过程中,与文献筛选的过程类似,每篇文献均由本文的两位不同的作者独立进行评估,当未达到共识时,需要对该文献进行重新质量评估和充分讨论,最终得到一致的结果.

根据文献[22,27]中给出的实验结果,上述文献质量指标可以有效地度量文献的质量.为了保障文献分析的质量和有效性,与文献[24,27,28]类似,我们设定了最小文献质量阈值为 8 分.如果文献的整体质量得分少于 8 分,则进一步将这些文献删除.这样,我们只保留质量得分不小于 8 分的文献作为最终的文献数据集.最终,通过对文献的质量进行评估,筛选得到了 142 篇相关文献(<https://chaotan201.github.io/ComplexSystemUncertaintyPapers.html>),并将这些相关文献作为最终的文献数据集进行调研.

Table 2 The metrics for measuring the quality of the primary studies and scores

表 2 文献质量度量指标及对应的得分值

序号	维度	质量指标	质量得分值
Q1	质量报告	该文献是否有清晰明确的目标? ^[27]	Y=1,N=0,P=0.5
Q2		该文献是否开展了一系列的相关实验或是报告了经验? ^[22]	Y=1,N=0
Q3		该文献有没有对研究上下文进行充分的介绍(如某个行业、实验环境等)? ^[22,27]	Y=1,N=0,P=0.5
Q4		该文献是否明确说明了研究目标? ^[22]	Y=1,N=0,P=0.5
Q5	严谨性	该文献是否清晰地描述了所使用的技术? ^[24]	Y=1,N=0,P=0.5
Q6		该文献是否详细介绍了研究设计的理由? ^[27]	Y=1,N=0,P=0.5
Q7		该文献中的实验是否收集了足够多的项目数据集或案例研究? ^[24]	Y=1 or 2,N=0
Q8		该文献是否有工具支持? ^[28]	Y=1,N=0
Q9		该文献是否经过实证评估? ^[28]	Y=1,N=0
Q10	可信性	该文献是否明确讨论了这项研究的局限性? ^[25]	Y=1,N=0,P=0.5
Q11		该文献是否有关于研究结果的讨论? ^[28]	Y=1,N=0,P=0.5
Q12	相关性	该文献是否对工业界或学术界有价值? ^[22,24]	Y=1,N=0,P=0.5

3.5 属性抽取与细化

本模块包括 3 个子模块,分别为:文献属性定义、文献属性抽取、文献属性清洗和细化.

在文献属性定义子模块,我们定义了 12 个文献属性,见表 3.前 4 个属性为全局属性,各个研究问题都需要使用这 4 个属性.属性 5~属性 12 分别对应我们已给出定义的 8 个研究问题(详见第 3.1 节).

在文献属性抽取子模块,基于上述已定义的属性,我们对所有文献进行各个属性值的抽取.从表格中可以看出,部分文献属性(属性 1~属性 7)为客观属性,可以很容易地通过工具自动获取.还有部分属性为主观属性(属性 8~属性 12),需要人工对文献进行阅读和分析,从而获得该属性的正确值.为了保证主观属性值抽取的正确性,我们采取了以下措施.

- (1) 同一篇文献由本文的两位不同作者同时进行属性值的抽取,两位作者不知道对方的属性值抽取结果.
- (2) 如果他们的属性值抽取的结果一致,该文献属性值的最终结果即为两位作者的一致结果.
- (3) 如果他们抽取到了不同的属性值,则两位作者需要进行协商以得到一致的结果.

在文献属性清洗和细化子模块,针对客观属性(属性 1~属性 7),我们对文献的部分属性进行了清洗和细化,使这些属性的值在各个文献的书写形式上保持统一,比如作者姓名、作者机构、出版物类型以及关键词.为了

使这些属性值在不同文献的形式统一,我们采取了以下措施.

(1) 针对作者名,考虑到 IEEE Xplore、ACM Digital Library、ScienceDirect 和 Springer 数据库中提供的作者名的形式有所不同,我们从文献及其对应的发表网站上爬取了所有文献中所有作者的全名,以作者全名代替所有的缩写名以及不同的书写形式.

(2) 针对作者机构,考虑到有些作者可能有多个所属的科研机构,我们以该作者所提及的第 1 单位作为其所在的单位.同时,不同的文献数据库对不同机构的书写形式也不一样.同样,我们爬取了所有机构的全称,以全称代替所有的缩写形式.

(3) 针对出版物类型,我们同样从文献相关的网站上爬取了出版物的全称,并以出版物的全称代替所有的其他书写形式.

(4) 针对关键词,考虑到不同的作者有着不同的写作习惯和表达方式,他们会使用不同的关键词来表示相同的主题,比如 uncertainty、unpredictability、indeterminacy 等词经常替代使用.因此,我们建立了一个同义词表,将不同的词替换为同一个词,比如将 unpredictability、indeterminacy 替换为 uncertainty,以更好地发现文献的相关主题及其共现关系.同时,有些关键词的全称有其简写形式,比如 Unified Modeling Language 的简写形式为 UML.因此,我们还建立了一个关键词简称和全称对照表,并以此为根据将关键词的简写替换为其对应的全称.

Table 3 The attribute details of the primary studies

表 3 文献属性的详细信息

序号	属性名	属性描述	相关的研究问题
1	编号	文献唯一标识	全部
2	标题	文献的标题	全部
3	作者姓名	文献全部作者的姓名	全部
4	作者机构	作者所在的科研机构	全部
5	出版时间	文献的出版时间	RQ1
6	出版物类型	文献出版类型,如期刊、会议等	RQ2
7	关键词	文献中涉及到的关键词	RQ3
8	生命周期阶段	不确定性研究涉及到的系统生命周期的阶段	RQ4
9	运用领域	不确定性研究涉及到的领域,如信息物理系统、物联网等	RQ5
10	不确定性类型	不确定性研究涉及到的不确定性类型	RQ6
11	不确定性处理方法	不确定性研究涉及到的不确定性处理方法	RQ7
12	学科领域	不确定性研究涉及到的学科领域,如自动化、计算机(软件工程)等	RQ8

3.6 研究问题调研

基于相关文献及其对应的属性,我们构建了文献-属性表格,并根据此表格调研和回答各个研究问题.需要强调的是,这个过程是数据收集(尤其是针对文献属性 8~属性 12)中一个至关重要的环节.所以整个数据收集过程,由本文的两位作者进行数据收集和反复的协商直至得到一致的结果,主要目的是完善不确定性分类以及不确定性处理方法.

4 实验结果

本节基于收集到的数据,针对每个研究问题,展开详细讨论.

4.1 研究问题1:复杂系统的不确定性研究的相关文献随时间的发表情况是怎样的?

为了回答这个研究问题,我们统计了从 2011 年~2020 年(截止到 3 月)每年文献的发表数量.图 2 展示了相关文献发表年份的分布情况.从图 2 中可以发现,在 2015 年及其之前复杂系统的不确定性领域文献发表的数量增幅相对平缓,发表文献较多的是 2013 年和 2015 年,文献发表数量分别为 9 篇和 10 篇.在 2015 年之后,文献发表的数量呈逐年递增的趋势,到 2018 年达到了 37 篇(37/142).这表明,研究人员对复杂系统的不确定性研究的关注

度越来越高,越来越多的研究人员投入到该领域的研究中,该领域逐渐成为热点研究领域.

调研结果表明:从总体趋势来看,过去 10 年中,关注复杂系统不确定性的研究呈现出上升趋势,这说明不确定性研究越来越受研究人员的关注.

4.2 研究问题2:复杂系统的不确定性研究的相关文献发表在哪些刊物上?文献发表数量排名靠前的期刊和会议都有哪些?

在本研究问题中,我们对于复杂系统的不确定性相关文献发表的刊物进行了分析,发现相关文献发表在期刊(journal)、会议(conference)、研讨会(workshop)以及书章节(book section)这 4 种类型的文献上.图 3 展示了相关文献在期刊、会议、研讨会和书章节的发表数量的分布情况.与复杂系统不确定性相关的 59 篇文献发表在 39 个不同的期刊上、69 篇文献发表在 60 个不同的会议上、9 篇文献发表在 8 个不同的研讨会上、5 篇文献发表在书的章节中.这说明,复杂系统不确定性相关的文献广泛分布在大量的期刊和会议中.从图 3 可以看到,发表在期刊和会议上的文献数量占据很大的比例,所占比例超过了 90%.所以我们很容易看出针对复杂系统,如 CPS 和 IoT 的研究的跨学科特性.这一特征也符合其他研究发现,如文献[17]总结到自适应软件研发的跨学科特性.

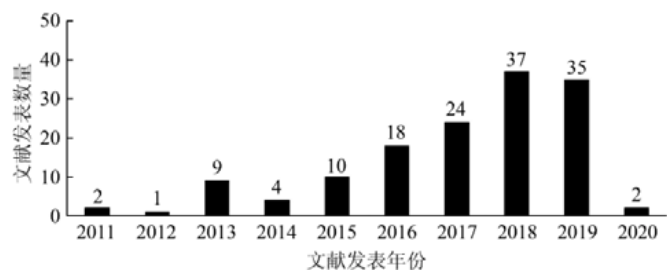


Fig.2 The descriptive statistics on the number of the primary studies published each year
图 2 每年发表的文献数量统计情况

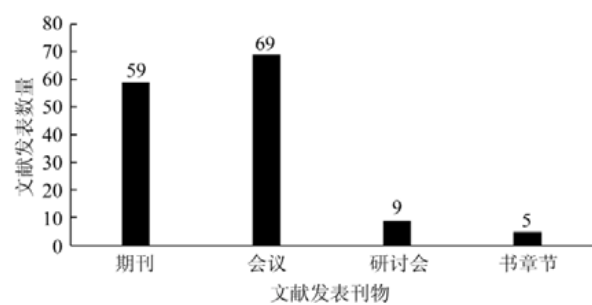


Fig.3 The distribution of the primary studies published in the four different venues
图 3 相关文献在不同刊物的发表数量的分布情况

我们进一步分析发表文献数量排名靠前的期刊和会议.表 4 是文献发表数量不少于 2 篇的期刊,其中,IEEE Access、Automatica、Journal of Systems and Software(JSS)和 Software & Systems Modeling(SoSyM)这 4 个期刊发表的论文数量是 4 篇及 4 篇以上.值得一提的是,Automatica 是控制领域的知名期刊,JSS 和 SoSyM 是软件工程领域的知名期刊.

表 5 是文献发表数量不少于 2 篇的会议,其中,ACM/IEEE Int'l Conf. on Cyber-Physical Systems 和 Doctoral Conf. on Computing,Electrical and Industrial Systems 这两个会议受研究人员关注较多.需要再次强调的是,复杂系统不确定性方面的研究发表的期刊和会议刊物非常分散,142 篇文章发表在 112 个不同的刊物,说明这个研究领域涉及的范围广泛.

Table 4 The top 9 journals having at least two primary studies published
表 4 发表文献数量不少于 2 篇的前 9 种期刊

排名	期刊	发表数量
1	IEEE Access	7
2	Automatica	4
3	Journal of Systems and Software	4
4	Software & Systems Modeling	4
5	ACM Transactions on Embedded Computing Systems	2
6	Future Generation Computer Systems	2
7	IFAC-PapersOnLine	2
8	Information and Software Technology	2
9	Journal of the Franklin Institute	2

Table 5 The top 6 conferences having at least two primary studies published
表 5 发表文献数量不少于 2 篇的前 6 种会议

排名	会议名	发表数量
1	ACM/IEEE International Conference on Cyber-Physical Systems	4
2	Doctoral Conference on Computing, Electrical and Industrial Systems	3
3	European Conference on Modelling Foundations and Applications	2
4	European Conference on Software Architecture	2
5	NASA Formal Methods Symposium	2
6	Software Engineering for Self-Adaptive Systems III. Assurances	2

调研结果表明:复杂系统不确定性的文献分布在多个领域的多种刊物上,反映了多学科交叉的特性.

4.3 研究问题3:复杂系统的不确定性研究具有哪些高频关键词以及它们之间的关系是怎样的?

文献中的关键词可以反映文献的主题概念,我们抽取了相关文献中研究人员定义的关键词,并统计了各个关键词出现的频次.我们按照出现的频次对关键词进行排序,并在表 6 中列出了出现频次不少于 4 次的关键词,包括 uncertainty(30 次)、cyber-physical systems(26 次)、Internet of Things(25 次)、decision making(6 次)、self-adaptation(6 次)等.从这些高频关键词中我们可以发现以下 3 个现象.一是高频关键词与我们设定的搜索字符串高度相关.比如我们设定的搜索字符串包括待分析的目标系统为信息物理系统和物联网,而这两个关键词也是出现频次比较高的两个关键词.二是这些高频关键词关注不确定性的不同方面,比如自适应性(self-adaptation)和复杂事件过程(complex event process)等.三是这些高频关键词和处理不确定性问题的技术相关,比如基于模型的测试(model-based testing)、决策制定(decision making)和概率模型检测(probabilistic model checking)等.

每篇文献的关键词可以反映出其研究主题或使用的主要技术.两个关键词在不同的文献中共同出现的次数越多,代表这两个关键词所描述的主题越密切.为了更好地对复杂系统的不确定性领域的热点研究主题进行分析,我们利用关键词共现关系构建关键词共现网络,并将关联密切的关键词聚类成关键词簇.我们利用荷兰莱顿大学开发的针对文献知识进行分析的可视化工具 VOSviewer^[29,30]来抽取和构建关键词共现矩阵,进而形成关键词共现关系网络和聚类结果.为了更好地展示关键词的共现关系和聚类效果,我们选取了共现次数大于 1 的关键词.

图 4 展示了复杂系统的不确定性相关文献的关键词的共现关系和聚类结果,其中圆圈和对应标签表示一个关键词,圆圈之间的连线表示两个关键词共现超过 1 次,圆圈的大小表示该关键词与其他关键词共现的次数,圆圈越大,表示该关键词与其他关键词共现的次数越多.通过关键词的共现关系可以看出 uncertainty 分别与 cyber-physical systems、Internet of Things 共现,通过关键词出现的频次可以看出 uncertainty、cyber-physical systems 以及 Internet of Things 与其他关键词共现频次较高,且 Bayesian network、context awareness 与其他关键词共现频次较低.

Table 6 The top 10 keywords appearing in the provided keywords lists of the primary studies
表 6 不确定性研究出现频次不少于 4 次的前 10 个关键词

排名	关键词	频次
1	uncertainty	30
2	cyber-physical systems	26
3	Internet of Things	25
4	decision making	6
5	self-adaptation	6
6	model-based testing	5
7	industry 4.0	5
8	probabilistic model checking	4
9	cluster	4
10	complex event process	4

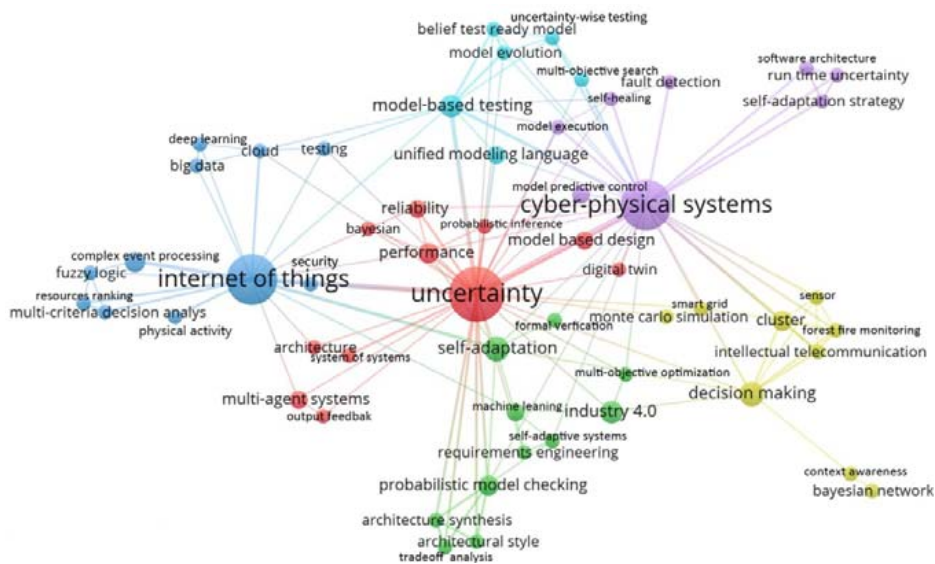


Fig.4 The keywords co-occurrence relations and clustering results in the primary studies
图 4 不确定性相关文献中的关键词共现关系及聚类结果

另外,图 4 中的不同颜色代表不同的聚类结果.从图中可以看出,我们共得到了 6 个关键词的聚类结果.

红色代表的聚类:主要内容为运用相关的技术保证复杂系统的性能和可靠性,如输出反馈(output feedback)和贝叶斯(Bayesian)模型;绿色代表的聚类:主要内容为运用概率模型检查(probabilistic model checking)技术、形式化验证(formal verification)以及机器学习(machine learning)等支持软件自适应;蓝色代表的聚类:主要内容为利用大数据(big data)、深度学习(deep leaning)以及模糊逻辑(fuzzy logic)等技术来保障系统安全;紫色代表的聚类:主要关注不确定性发生在信息物理系统(cyber-physical systems)运行时的不确定性(run time uncertainty)、自适应策略(self-adaptation strategy)、自我修复(self-healing)以及故障检测(fault detection);黄色代表的聚类:主要内容是通过诸如聚类(clusters)、贝叶斯网络(Bayesian network)等技术和方法支持不确定性决策;青色代表的聚类:主要内容是利用相关测试技术对不确定性系统进行建模和测试,如不确定性测试(uncertainty-wise testing)、基于模型的测试(model-based testing)和信念测试就绪模型(belief test ready model).

调研结果表明:通过对关键词词频进行统计,复杂系统不确定性主要针对信息物理系统和物联网等系统,主要关注自适应和复杂事件过程中的不确定性,以及基于模型的测试、决策制定和概率模型检测等处理不确定性的相关技术.

4.4 研究问题4:复杂系统的不确定性相关文献关注在系统开发生命周期的哪些阶段?

为了调研复杂系统不确定性相关文献关注的系统开发生命周期的阶段,我们首先根据 ISO/IEC/IEEE^[31]的标准将复杂系统的生命周期分为 14 个阶段:业务和任务分析(business or mission analysis)、利益相关者的要求和需求定义(stakeholder needs and requirements definition)、系统需求定义(system requirements definition)、架构定义(architecture definition)、设计定义(design definition)、系统分析(system analysis)、实现(implementation)、集成(integration)、验证(verification)、迁移(transition)、确认(validation)、运行(operation)、维护(maintenance)和处置(disposal).接着,我们对收集的相关文献按上述阶段进行分析.需要强调的是,一篇文献可能关注生命周期的多个不同阶段.

从图 5 中可以看出,研究人员在研究复杂系统本身及其开发和运行中的不确定性时,重点关注系统开发过程中的设计定义(87 次)、系统分析(25 次)、运行(21 次)、验证(17 次)、架构定义(17 次)、确认(13 次)等阶段,而业务和任务分析、迁移、处置等阶段到目前为止没有被研究人员所关注.图 6 展示了研究人员关注的系统生命周期不同阶段随时间的变化情况,横坐标表示系统生命周期的各个阶段,纵坐标表示不同的年份.横坐标和纵坐标的交叉位置有对应的圆形图案,圆的大小表示系统开发生命周期不同阶段在对应的年份发表文献的数量.从图中可以看出,对于设计定义阶段,从 2013 年开始,此研究领域开始吸引较多的研究人员关注,2016 年及以后一直是研究人员关注的热点.这个增长趋势与复杂系统,如信息物理系统,与物联网在工业应用的增长趋势相一致.工业 4.0 作为一项德国战略倡议于 2013 年正式提出^[32],此后,各种复杂系统开始在工业广泛应用,并出现了一系列的不确定性问题,进而推动了该研究领域的发展.据我们推测,该阶段这一趋势与自适应系统有关^[33],其中 2013 年有 3 位研究者关注自适应系统的设计定义.对于架构定义阶段,2013 年就有研究人员开始关注,但是一直处于波动的状态,直到 2018 年才趋于平缓.而系统分析和运行阶段整体上从 2013 年开始引起研究人员的关注,直到 2017 年才被研究人员重点关注.对于验证和确认阶段,研究人员前期并未重点关注,分别到 2015 年、2017 年才引起研究人员的关注,并一直成为研究人员重点关注的阶段.我们认为这一趋势与复杂系统,诸如信息物理系统和物联网近年来广泛在人们生产生活和工业中的应用,导致对验证和确认(如测试)需求的增加有关.

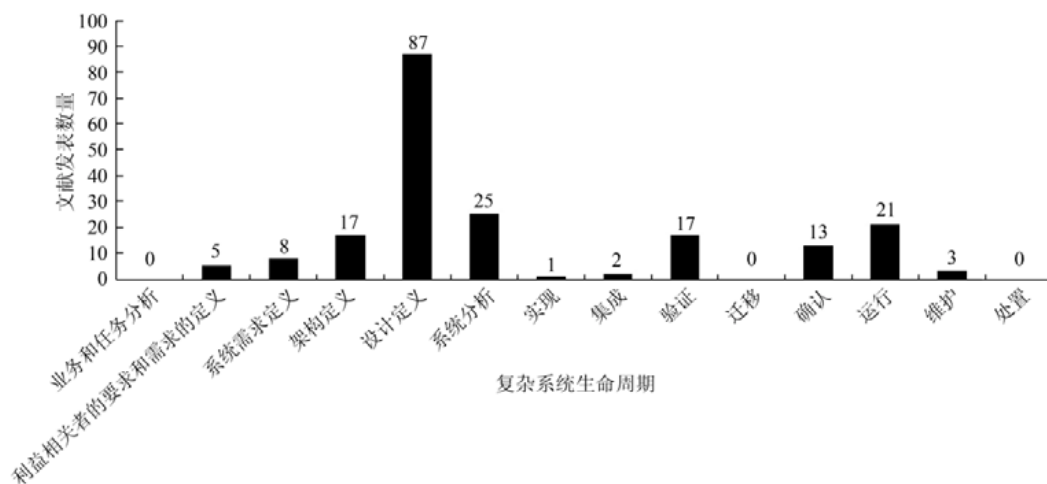


Fig.5 The distribution of the primary studies according to uncertainty along with the various stages of the system development life cycle

图 5 相关论文关注的复杂系统生命周期的不同阶段的分布情况

调研结果表明:设计定义阶段是研究人员关注最多的阶段,自 2013 年以后,复杂系统设计定义阶段的不确定性研究出现了明显的增长趋势.对于架构定义阶段,2013 年就有研究人员开始关注,但是一直处于波动状态,直到 2018 年才趋于平缓.而系统分析和运行阶段整体上从 2013 年起引起研究人员的关注,直到 2017 年才被研究

人员重点关注.对于验证和确认阶段,研究人员前期并未重点关注,分别到 2015 年、2017 年才引起研究人员的关注,并一直成为研究人员重点关注的阶段.

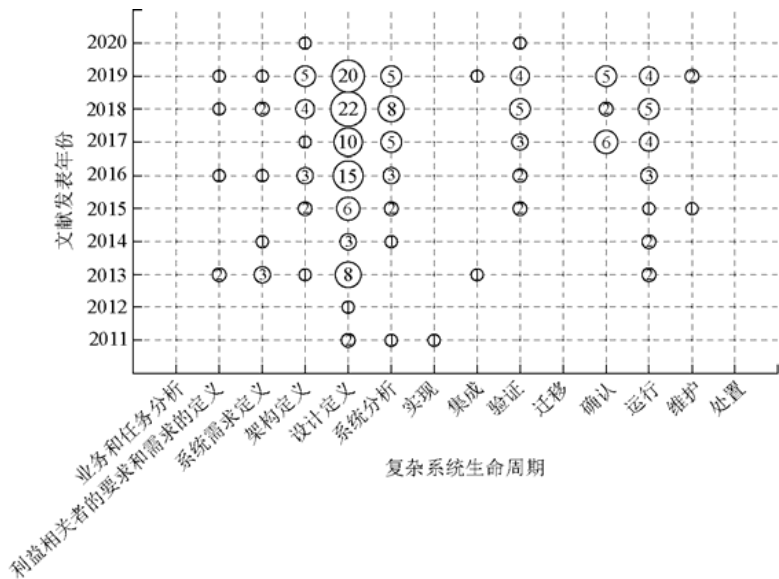


Fig.6 The distribution of the primary studies published in various years along with the various stages of the system development life cycle
图 6 复杂系统生命周期的不同阶段及不同时间段的分布情况

4.5 研究问题5:不确定性出现在哪些复杂系统中?

由于在检索相关文献时设定的不确定性的应用领域为信息物理系统(CPS)和物联网(IoT),我们统计了不确定性在这两个系统应用的分布情况.如图 7 所示,其中,文献数表示涉及到纵轴对应的复杂系统类型的文献数量,总体来说,有 70 篇文献(49.3%)的研究对象仅限于信息物理系统,45 篇文献(31.7%)的研究对象仅限于物联网.此外,8 篇文献的研究对象同时包含信息物理系统和物联网.由于不确定性是自适应系统需要考虑的关键因素^[34,35],因此我们进一步将研究对象是自适应信息物理系统和自适应物联网的相关文献作为单独的类别进行深入分析.统计结果显示,针对自适应信息物理系统不确定性的研究有 16 篇文献,针对自适应物联网的研究文献有 2 篇,而针对自适应信息物理系统和自适应物联网的研究文献仅有 1 篇.从这个统计结果可以看出,在过去 10 年中,针对不确定性的研究对象主要是信息物理系统.然而,学术界和工业界对信息物理系统和物联网的定义以及两者之间的界限一直存在争议.正如 Greer 等人所指出的^[4],随着技术的发展,信息物理系统和物联网的定义将趋于融合,共同强调通过集成物理和逻辑功能设计的混合系统.

图 7 展示了不确定性在不同复杂系统生命周期不同阶段的分布情况.其中,图中总计表示涉及到纵轴对应的复杂系统类型与横轴对应该系统开发生命周期文献数量的总和,需要注意的是,在 1 篇相关文献中可能涉及到生命周期的多个阶段,所以图 7 中总计和文献数对应的数量不一致.从图 7 可以看出,研究人员研究不确定性时更关注设计定义阶段的不确定性并应用到信息物理系统和物联网.同时,研究人员在研究信息物理系统和物联网中的不确定性时关注设计定义、系统分析以及运行阶段.同时,关注信息物理系统和物联网的文献以及关注自适应信息物理系统的文献大部分涉及系统生命周期的设计定义阶段.

调研结果表明:复杂系统不确定性主要关注于信息物理系统和物联网.而针对自适应系统,自适应信息物理系统是研究关注的重点.针对不确定性在生命周期以及运用领域的调研,研究人员研究不确定性更关注设

计定义阶段的不确定性并将其运用到信息物理系统和物联网.而自适应信息物理系统更多地关注设计定义阶段和运行阶段.

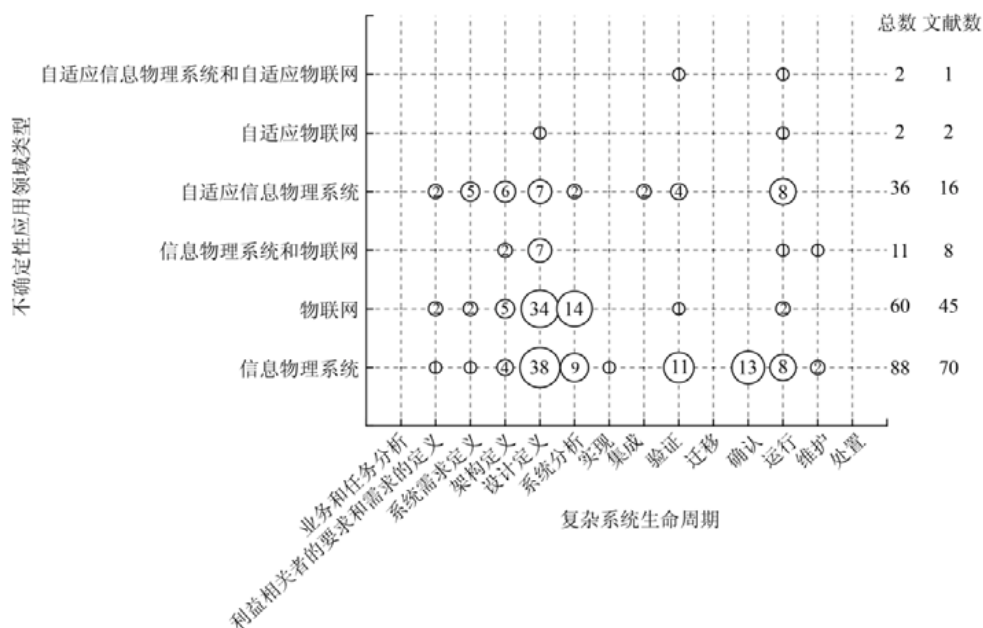


Fig.7 The distribution of the primary studies concerning different types of systems along the various stages of the system development life cycle

图7 不确定性在系统生命周期及运用领域的分布情况

4.6 研究问题6:复杂系统的不确定性可分成哪些类型?

4.6.1 外部、内部和传感器不确定性

针对复杂系统的不确定性类型,我们通过对文献进行深入的分析,提出了3种不同类型的不确定性:外部不确定性(external uncertainty)、内部不确定性(internal uncertainty)和传感器不确定性(sensor uncertainty).外部不确定性着重考虑系统边界外的不确定,内部不确定性指的是系统本身的不确定性.传感器为联系系统外部和内部的纽带,很多文献都将传感器不确定性从外部不确定性这个类别里独立出来成为一个新的分类.因此,在本文中,我们也将传感器不确定性作为一个独立的类别.基于上述定义,我们统计了研究人员处理不同类型的不确定性文献统计的数量.如图8所示,研究人员更多地关注外部不确定性(64次),内部不确定性(54次)次之,而传感器不确定性(11次)较少.其中,33篇文献同时关注内部和外部两种不确定性.另外,在调研的142篇文献中,46篇文献没有讨论外部、内部以及传感器不确定性.

为了更好地展示不确定性不同层次的分类情况,我们进一步将3种不确定性类型进行了细粒度的分类.外部不确定性是一个非常受研究人员关注的不确定性类型,系统在其开发和运行过程中,时刻受到多方面外部因素的影响.本文将外部不确定性分成环境不确定性、基础设施不确定性、用户行为不确定性以及经济属性不确定性,并对文献进行标注并统计,统计结果如图9所示,4种外部不确定性的相关文献分别为44篇、15篇、8篇以及6篇.进一步分析后得到了4种外部不确定性出现的原因.首先,系统在运行过程中直接与环境进行交互,而环境是时刻变化的,所以研究人员更多地关注环境的不确定性(如文献[36-38]).其次,随着复杂系统在各行业的广泛应用,复杂系统的规模越来越庞大.系统运行需要更多的基础设施,从而引入了更多的不确定性,比如网络的不确定性(如文献[39,40]).此外,系统在与用户交互时,用户行为的不确定性也是导致系统行为变化或不确定的主要因素之一^[13,41].最后,由于一些信息物理系统(如智能电网 Smart Grid)的运行依赖于一些动态变化的经济

属性(如电的浮动市场价格),这些经济属性直接影响到信息物理系统为了实现成本优化和平衡做出的智能决策和控制(如文献[42]).

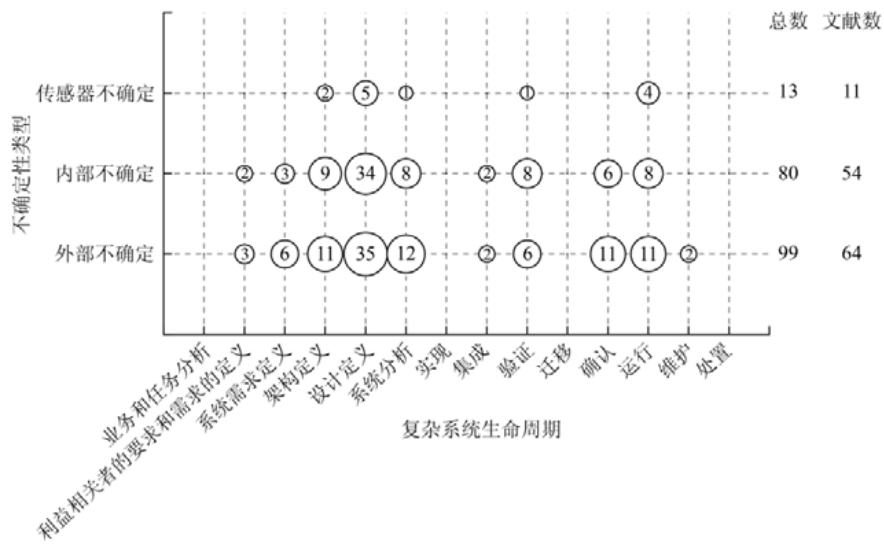


Fig.8 The distribution of the primary studies according to the internal/external/sensor uncertainty classification along with the various stages of the system development life cycle
图 8 不确定性分类在系统生命周期的不同阶段的分布情况

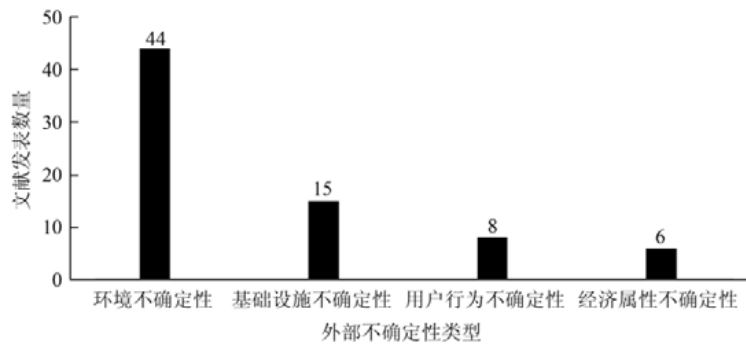


Fig.9 The distribution of the primary studies according to the four different types of external uncertainties
图 9 相关文献在外部不确定性不同类型的发表数量的分布情况

系统在运行过程中也会出现其本身多个方面的内部不确定性.一是系统内部的系统结构具有不确定性,比如系统构件不确定性(如文献[43])和系统架构演化的不确定性(如文献[44]);二是系统内部交互时会出现不确定性,比如数据不确定性(如文献[45]);三是支持系统运行的相关技术会出现不确定性,比如深度学习中深度神经网络本身固有的不确定性(如文献[46]);四是处理系统运行的技术的不确定性,比如利用机器学习技术来评估深度神经网络的预测不确定性或预测错误(prediction uncertainty or errors)(如文献[47]).本文通过对相关文献进行分析,提出了 4 种内部不确定性,包括系统结构不确定性、内部交互不确定性、支持系统运行的技术的不确定性和处理系统运行技术的不确定性.4 种内部不确定性类型的数量统计结果如图 10 所示,研究人员更多关注的是系统结构不确定性(18 次)和内部交互不确定性(26 次),而支持系统运行的技术不确定性(5 次)以及处理系统运行技术的不确定性(9 次)关注得较少.

在复杂系统与环境交互时,传感器扮演着对物理过程进行测量和对物理世界进行感知的重要角色.传感器不确定性一方面关注传感器度量的不确定性,比如传感器度量的精度和准确性(如文献[48]),另一方面关注系统外部(比如环境)数据的不确定性(如文献[49–52]).在我们调研的文献中,绝大多数(10 篇/11 篇)关注的是传感器采集的数据不确定性.比如,对于面向服务的体系结构,当传感器在采集数据过程中遇到数据的不确定性时,如何利用模糊逻辑去降低数据的不确定性(如文献[51]).

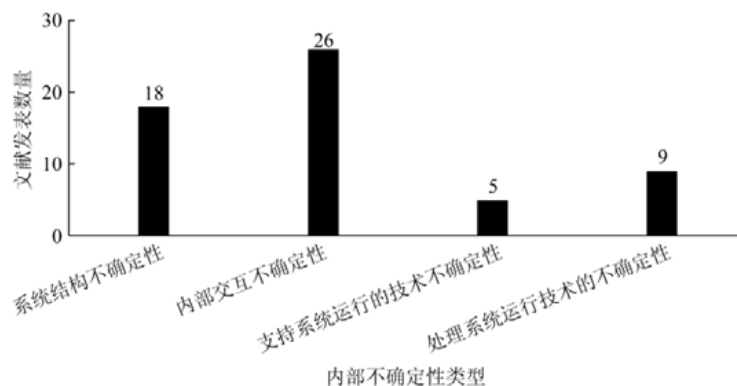


Fig.10 The distribution of the primary studies according to the four different types of internal uncertainties

图 10 相关文献在内部不确定性不同类型的发表数量的分布情况

复杂系统及其开发和运行中的不确定性种类多种多样,其分类标准很多情况下与信息物理系统和物联网的具体设计和应用领域有关,目前还缺少统一的分类标准.基于目前收集到的数据,我们提出了 3 种不确定性类型并进行了数据统计.我们未来的一项工作是针对信息物理系统和物联网开展全面且深入的分析,以提出专门针对这两种复杂系统的不确定性分类.

图 8 也展示了不同类型的不确定性在系统生命周期不同阶段的分布情况.我们可以发现,外部不确定性和内部不确定性都主要出现在设计定义、系统分析、确认、架构定义和运行阶段,尤其是架构定义、设计定义和系统分析阶段的文献较多(占总论文数的 42.3%).研究人员在系统架构和设计方面投入了很多精力,以移除或降低内外部不确定性对系统行为及其属性(如安全性)的负面影响.

而针对传感器不确定性,研究人员关注的阶段主要包括设计定义和运行.由于传感器不确定性相关的文献数量较少(只有 11 篇),而且这些文献涉及到系统生命周期的 5 个不同的阶段,尚不能有效地分析出传感器不确定性分布在这些阶段的原因.

调研结果表明:复杂系统的不确定性主要关注于外部不确定性和内部不确定性.而外部不确定性更多地关注于环境不确定性,内部不确定性更多地关注于内部交互不确定性.针对不确定性分类在系统生命周期的不同阶段的调研,外部不确定性和内部不确定性,主要关注于设计定义阶段.

4.6.2 系统开发过程中各制品的不确定性

不同的研究人员关注的不确定性的角度以及研究的粒度不同,为了更好地调研不确定性分类的研究现状,我们根据另外一种分类标准对不确定性进行分类.在我们收集的部分文献中明确地讨论了系统开发过程中的一些制品(例如需求、数据、模型、约束)的不确定性,这些制品的不确定性是在系统开发生命周期中产生的.这种分类与根据内部不确定性和外部不确定性的分类不同,因此我们进一步根据系统开发过程中产生的制品对不确定性进行分类.

我们对文献进行深入的分析,提出了 8 种制品的不确定性.

- 需求不确定性(requirements uncertainty):是指需求不准确、不可靠或不足,需求不断变化^[53–55].
- 模型不确定性(model uncertainty):是指模型结构和模型输出中的不确定性,也称为模型可变性(variability)^[48,56,57].

- 参数不确定性(parametric uncertainty):是指数学模型参数的不确定性,也称为参数可变性^[58-60].
- 数据不确定性(data uncertainty):是指(大)时空数据的不确定性^[61-63].
- 演化不确定性(evolution uncertainty):是指系统开发过程或开发过程中创造的人工制品(artifacts)演化中的不确定性,比如系统架构的变化^[44].
- 约束不确定性(constraint uncertainty):是指系统属性(如安全性)或变量之间依赖关系的约束不确定性^[64,65].
- 信息不确定性(information uncertainty):是指不完整、模糊和不确定的信息^[66].
- 时间不确定性(time uncertainty):是指时间的不确定性,比如时间延迟、时间约束的不确定性^[67,68].

图 11 展示了基于系统制品的不确定性分类在生命周期不同阶段的分布情况,需要指出的是,在收集的 142 篇相关文献中,只有 65 篇文献明确涉及到了不同制品的不确定性,同时,一篇相关文献可能涉及到多个不同制品的不确定性,所以明确涉及到不同制品的不确定性的文献数小于图中显示的文献数之和.从图 11 可以看出,研究人员更多地关注模型不确定性(23 次)、数据不确定性(28 次)以及参数不确定性(6 次),而针对需求不确定性、演化不确定性、约束不确定性、信息不确定性以及时间不确定性只有少量的研究人员关注.据分析是因为建模与数据采集和处理是信息物理系统和物联网开发过程中至关重要的两个步骤,所以模型不确定性和数据不确定性成为了研究人员最关注的两个研究点.

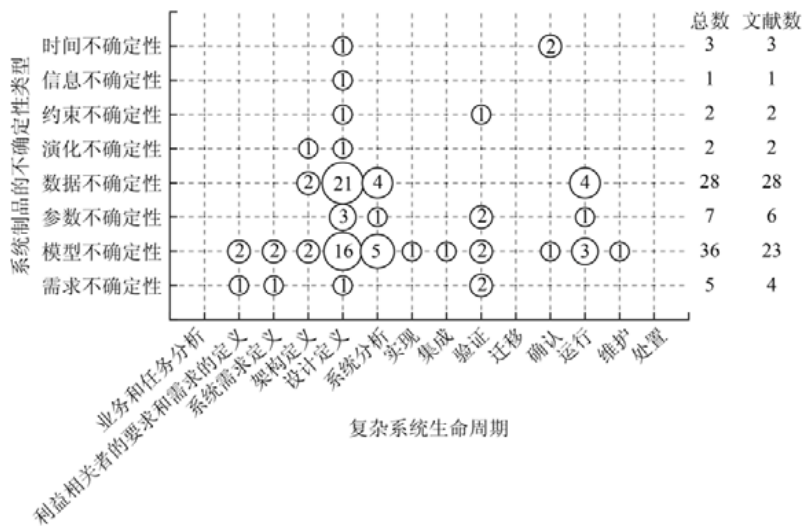


Fig.11 The distribution of the primary studies concerning uncertainties in different system artifacts along with the various stages of the system development life cycle

图 11 基于系统制品的不确定性分类在生命周期不同阶段的分布情况

另外,从图 11 可以看出,在利益相关者的要求和需求定义以及系统需求的定义阶段,研究人员关注的是需求不确定性和模型不确定性,主要原因是在信息物理系统和物联网生命周期的这两个阶段,需求描述和建模是需求定义的主要方法.在架构定义阶段,研究人员关注的是模型不确定性、数据不确定性以及演化不确定性.出现这种现象的主要原因包括:建模是系统架构定义的主要方法;数据采集和处理是架构定义(尤其是物联网架构定义)必须要考虑的内容.文献[44]关注基于学习的系统架构设计的自演化,这是唯一一篇我们收集到的关于架构模型自演化的文献.

在设计定义阶段,研究人员关注的不确定性类型相对于其他阶段具有更强的多样性,涵盖了所有 8 种类型.其中,数据不确定性和模型不确定性受到研究人员的更多关注.我们认为其主要原因是,建模是系统设计的主要方法,并且数据采集和处理是信息物理系统和物联网设计需要考虑的重要设计因素之一.在系统分析阶段,研究

人员关注的是模型不确定性、参数不确定性和数据不确定性.基于模型的系统属性(如安全性、可靠性)分析是信息物理系统和物联网开发过程的常用方法.此外,在系统设计和分析阶段,数学模型中的参数不确定性受到研究人员的关注,主要原因是数学模型通过调节和优化模型参数取值以期达到系统设计优化的目标.

在验证阶段,研究人员关注的是需求不确定性、模型不确定性、参数不确定性和约束不确定性.这一结果说明,需求、模型以及模型参数是重要的验证对象.在确认阶段,研究人员关注的是模型不确定性和信息不确定性,主要原因是信息物理系统和物联网主要的确认方法是测试.针对不确定性信息建模,并基于构建的模型和模型转换技术实现测试用例自动生成、优化和执行是目前比较常见的测试方法(如文献[69–71]).在运行阶段,研究人员主要关注的是模型不确定性以及数据不确定性.在实现、集成和维护阶段,研究人员关注的是模型不确定性.但是这些阶段的文献数量都很少,很难分析出具体的原因.

调研结果表明:针对系统开发过程中制品的不确定性,研究人员更多地关注模型不确定性、数据不确定性以及参数不确定性.针对系统制品的不确定性分类在生命周期不同阶段的调研,模型不确定性、数据不确定性以及参数不确定性主要分布在设计定义阶段.

4.7 研究问题7:针对复杂系统的不确定性,研究人员提出了哪些不确定性的处理方法?

对于不同的不确定性类型和所关注的研究重点,研究人员采取了不同的方法对不确定性进行处理,解决复杂系统开发和运行的不同阶段所面临的挑战.我们通过对相关文献进行分析,发现研究人员处理不确定性的方法可以分为以下6类.

- 不确定性度量(uncertainty quantification):使用不同的度量方法(如概率和粗糙集)定量表征不确定性^[8,72].
- 不确定性管理(uncertainty management):在需要更多知识时,寻求信息以纠正知识水平下降的过程^[73,74].其中减少不确定性(uncertainty reduction)是不确定性管理中的一个重要分支^[75].
- 不确定性规约/建模(uncertainty specification/modeling):在诸如需求、模型等人工制品的基础上对不确定性信息进行规约和建模^[76].
- 不确定性推理(uncertainty reasoning):旨在用于知识表示和推理的方法,例如,贝叶斯推理和模糊逻辑^[6].
- 不确定条件下的决策(decision making under uncertainty):支持系统行为优化和适应性测试等考虑不确定性的系统决策^[77].
- 不确定性传播(uncertainty propagation):不确定性来源(如输入参数数据不确定性)对系统行为(如自适应系统的自适应行为)不确定性的影响性分析^[50].

图12展示了不确定性处理方法的统计分布情况.如图所示,研究人员更关注的是利用不确定条件下的决策(58次)、不确定性推理(44次)以及不确定性规约/建模(29次)方法来应对复杂系统开发和运行中所遇到的各种不确定性.需要注意的是,一个不确定性处理方法可能出现在多个系统生命周期阶段.为了更深层次地了解对于复杂系统生命周期的不同阶段,研究人员更倾向于利用哪些不确定性处理方法来处理该不确定性,我们也对不确定性处理方法在开发周期不同阶段的分布情况进行了统计.如图12所示,我们发现对于设计定义阶段,研究人员使用了各种不确定性处理方法,其中不确定性条件下的决策和不确定性推理方法使用得较多,分别达到了34次和22次,而不确定性度量方法使用得较少,仅9次.在设计阶段,不确定性条件下的决策(如预测前面车辆的车速、到下一个交通道路交叉口的距离以及路段中的车道数量^[40])是不确定性处理的主要方法.

不确定性会极大地增加系统及其环境行为和属性推理的复杂度,而且简单的逻辑并不能充分地表达不确定信息,所以研究人员尝试了不同的不确定性推理方法,比如概率^[65]、模糊集^[53]以及贝叶斯概率^[78]等.面向复杂系统的不确定性推理的主要目标包括推理系统或其运行环境的状态和可能采取的决策行为等.从我们收集的数据来看,与其他方法相比,在架构设计、系统分析、验证、确认以及运行阶段不确定性推理是比较常用的不确定性处理方法.尤其是在验证阶段,利用不确定性推理方法的文献达到了13篇,用于验证系统属性(如可靠性)的正确性^[79]、系统运行时的需求满足性^[53]以及系统脆弱度(fragility)^[80]等.此外,在运行阶段,不确定性条件下的决策和不确定性推理是研究人员关注的主要不确定性处理方法,这主要是因为大部分的系统是自适应系统,如文献[81,82].

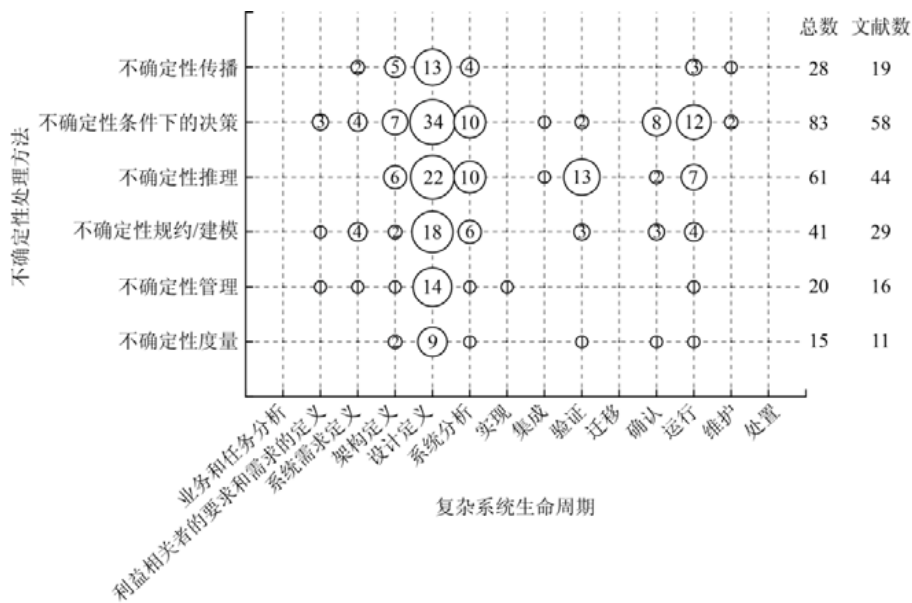


Fig.12 The distribution of the primary studies concerning methods dealing with uncertainty at the various stages of the system development life cycle
图 12 不确定性处理方法在系统生命周期不同阶段的分布情况

在本文中,我们没有从外部/内部/传感器不确定性和不确定性处理方法的相关性分析中观察到更多的信息,得到有效的结论.我们认为这主要是因为不确定性处理方法(如不确定性传播)都是建立在通用理论(如概率论、贝叶斯统计)之上,而且处理方法也是可以通用到外部、内部以及传感器等不同类别的不确定性.此外,我们未从不同的人工制品的不确定性(如数据不确定性、需求不确定性)与不确定性处理方法的分析中观察到它们之间的相关性.比如,基于概率的不确定性度量方法可以被用来度量模型、数据以及其他人工制品.另外,不同的不确定性规约/建模方法可以用来表述/建模诸如模型及参数、数据、约束的各种人工制品的不确定信息.然而,我们收集的数据并没有细分到不同的不确定性规约/建模方法,所以目前无法建立更细级别的关联性.同理,我们也认为,更细级别的分类也同样会有助于将具体的不确定性条件下的决策等各种处理方法与具体的不确定类型联系起来.未来,我们将着重开展这方面的工作.

调研结果表明:针对复杂系统不确定性处理方法,研究人员更关注于利用不确定性条件下的决策、不确定性推理以及不确定性规约/建模.针对不确定性处理方法在系统生命周期不同阶段的调研,在设计阶段,研究者关注于不确定性条件下的决策、不确定性推理.在验证阶段,研究者关注于不确定性推理.在运行阶段,研究者关注于不确定性条件下的决策和不确定性推理.

4.8 研究问题8:针对复杂系统的不确定性,自动化和计算机领域(尤其是软件工程)研究的关注点有什么差异?

复杂系统不确定性研究具有多学科交叉特性.对于不确定性的研究,不同领域的研究人员关注的研究主题有所差异.为了探究自动化和计算机领域研究人员对于不确定复杂系统关注点的差异,我们将相关文献划分为自动化领域和计算机领域.通过从不确定性出现在哪些复杂系统、复杂系统的不确定性有哪些分类(外部、内部和传感器不确定性以及系统开发过程中各制品的不确定性)以及针对复杂系统的不确定性、研究人员提出哪些不确定性的处理方法这 3 个方面出发,分析两个领域关注点的差异.

本文将 142 篇文献按照领域分成 4 种类型,计算机、自动化、计算机/自动化以及其他,具体如图 13 所示.从图中可以看出,计算机相关文献达到 103 篇(72.5%),自动化相关文献达到 22 篇(15.5%),计算机/自动化相关文献达到 10 篇,其他相关文献达到 7 篇.其中,计算机和自动化的文献数量相差较大,若直接对这两种领域相关文

献进行研究分析,容易对分析的结果造成偏差.为了更好地进行分析,本文按照中国计算机学会的分类标准(计算机体系结构/并行与分布式计算/存储系统、计算机网络、网络与信息安全、软件工程/系统软件/程序设计语言、数据库/数据库挖掘/内容检索、计算机科学理论、人工智能、人机交互与普适计算、跨学科交叉以及计算机图形学与多媒体)进一步对计算机领域进行分类,具体结果如图 14 所示.从图中可以看出,软件工程文献达到 62 篇,而其他类型文献所占比例较少,如数据库/数据挖掘/内容检索(2 篇),计算机科学理论(1 篇),跨学科交叉(1 篇),计算机图形学与多媒体(0 篇).为了更好地进行对比,本文着重对计算机领域中的软件工程(62 篇)和自动化学科(22 篇)的相关文献进行分析.

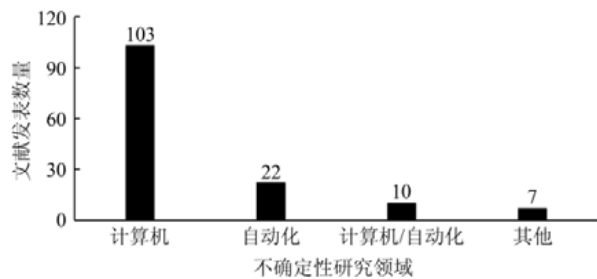


Fig.13 The distribution of the primary studies published in various research fields
图 13 相关文献在不同研究领域发表数量的分布情况

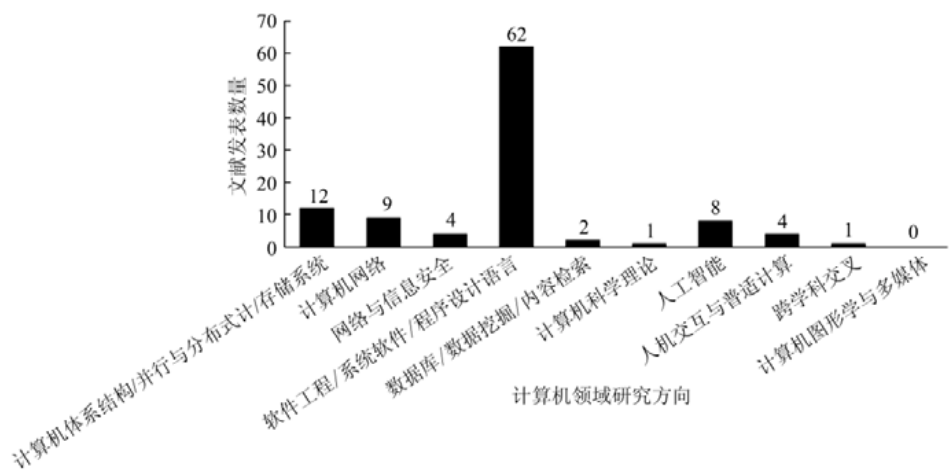


Fig.14 The distribution of the primary studies published in various research direction in computer science
图 14 相关文献针对不同的计算机研究方向发表数量的分布情况

图 15 展示了软件工程和自动化相关领域的文献在不确定性应用领域发表数量的分布情况.软件工程领域研究者主要关注于信息物理系统、物联网以及自适应信息物理系统,文献数量分别达到 32 篇(51.6%)、11 篇(17.7%)以及 13 篇(21.0%),自动化领域研究者主要关注信息物理系统,文献数量达到 18 篇(81.8%).这一现象与自动化学科的特性有关,信息物理系统中的控制(control)是自动化关注的重点.但是,针对自适应信息物理系统,自动化领域并没有涉及,却是软件工程领域研究者关注的重点.

图 16 展示了软件工程和自动化相关领域的文献在不确定类型中发表数量的分布情况.从图中可以看出,软件工程领域研究复杂系统的不确定性涉及到了外部不确定性、内部不确定性以及传感器不确定性,文献数量分别达到了 32 篇(51.6%)、25 篇(40.3%)以及 7 篇(11.3%).值得注意的是,一篇文献可能关注多个不确定性类型,所以软件工程相关文献在不确定性类型文献统计之和大于 62.自动化领域研究者只关注于外部不确定性和内

部不确定性,文献数量分别达到 10 篇(62.5%)和 6 篇(37.5%),没有涉及到传感器不确定性.

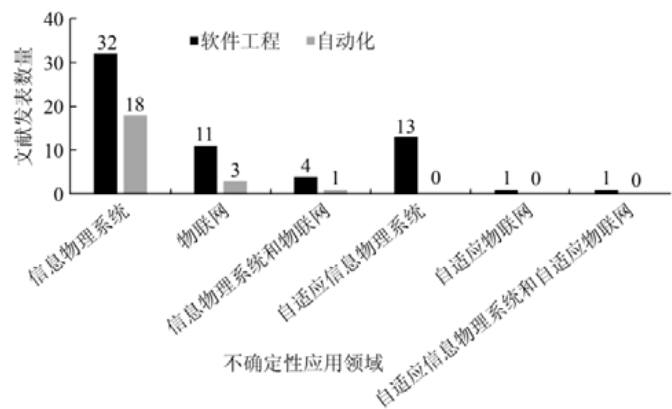


Fig.15 The distribution of the primary studies along with various types of systems concerned (software engineering vs. automation)

图 15 软件工程和自动化领域相关文献在不确定性应用领域发表数量的分布情况

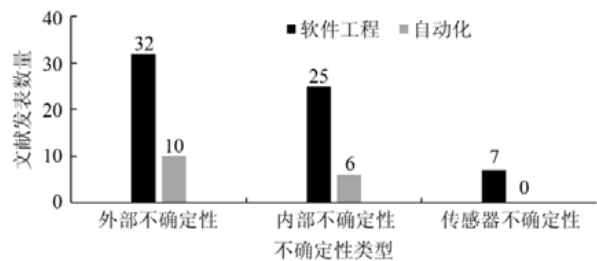


Fig.16 The distribution of the primary studies along with the internal/external/sensor uncertainty classification (software engineering vs. automation)

图 16 软件工程和自动化领域相关文献在不确定性类型中发表数量的分布情况

图 17 展示了软件工程和自动化相关领域的文献在系统制品的不确定性类型中发表数量的分布情况.从图中可以看出,软件工程领域研究者主要关注模型不确定性和数据不确定性,文献数量分别达到 8 篇(30.8%)和 12 篇(46.2%).自动化领域研究者主要关注模型不确定性和参数不确定性,文献数量分别达到 7 篇(58.3%)和 3 篇(25.0%).自动化领域关注于模型不确定性、参数不确定性、数据不确定性以及时间不确定性.这主要是因为其他类型的不确定性和这个领域的相关度较小(如需求不确定性).而针对时间的不确定性,软件工程和自动化都只有少量关注(各 1 篇),其中,软件工程关注于测试用例执行时间的不确定性^[67],自动化关注于转换系统轨迹分类的鲁棒时序逻辑推理的理论框架的时空不确定^[57].综上,软件工程和自动化领域关于不确定性的研究交叉集中在模型不确定性、参数不确定性以及数据不确定性.

图 18 展示了软件工程和自动化相关领域的文献在不确定性处理方法方面发表数量的分布情况.从图中可以看出,软件工程领域研究者主要关注不确定性规约/建模、不确定性推理、不确定性条件下的决策以及不确定性传播的不确定性处理方法,文献数量分别达到 17 篇(21.3%)、21 篇(26.3%)、24 篇(30.0%)以及 10 篇(12.5%).自动化领域研究者主要关注不确定性推理和不确定性条件下的决策,文献数量分别达到 4 篇(17.4%)和 13 篇(56.5%).可以看出,软件工程和自动化两个领域的关注点存在着很大的差异,自动化主要关注于不确定性条件下的决策,软件工程则关注于不确定性规约/建模、不确定性推理、不确定性条件下的决策.综上,针对不确定性条件下的决策,我们强烈推荐:除了软件工程领域的文献,也需要紧密关注自动化领域的文献,借助双方理论方

法的优势,以更好地处理复杂系统的不确定性.

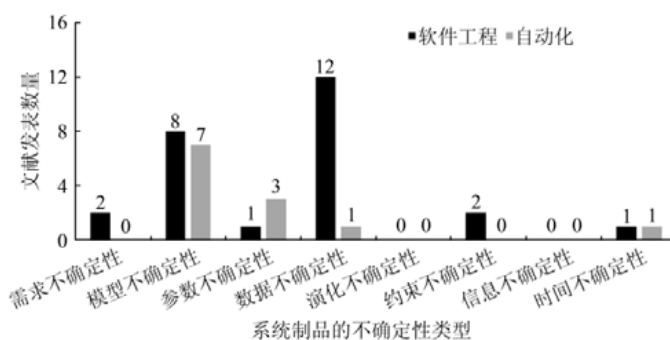


Fig.17 The distribution of the primary studies along with the various types of uncertainty-related systems artifacts (software engineering vs. automation)

图 17 软件工程和自动化领域相关文献在系统制品的不确定性中发表数量的分布情况

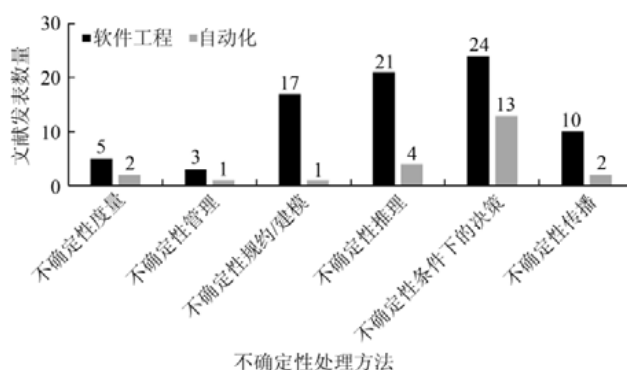


Fig.18 The distribution of the primary studies along with the various types of uncertainty handling methods (software engineering vs. automation)

图 18 软件工程和自动化领域相关文献在不确定性处理方法类型中发表数量的分布情况

关于软件工程和自动化领域在复杂系统不确定性的研究,首先针对复杂系统不确定性的应用领域,软件工程领域在自适应信息物理系统的不确定性有较多的研究,但是自动化领域却没有这方面的研究,可以考虑将软件工程针对自适应信息物理系统研究的理论和方法运用到自动化领域对这方面不确定性的研究.其次针对复杂系统不确定性的类型,软件工程领域研究者对传感器不确定性类型有颇多的关注,但是自动化领域没有这方面的研究,可以考虑扩展自动化领域对传感器不确定性的研究,同时将软件工程针对传感器不确定性的理论和方法运用到自动化领域对这方面不确定性的研究.再次针对复杂系统的系统制品的不确定性类型,软件工程领域研究者针对数据不确定性有颇多的研究,但是自动化领域却很少涉及到这方面的研究,可以考虑将软件工程针对数据不确定性研究的理论和方法运用到自动化领域对这方面不确定性的研究.同时可以看出,软件工程和自动化领域对于模型不确定性都有较多的关注和研究,软件工程领域更多关注的是抽象的软件需求、设计以及开发层面的模型,而自动化领域更多关注描述精准的数学模型.可以考虑将软件工程和自动化领域关于这方面的研究加以结合,借助双方的理论方法优势,进行更深层次的研究.最后针对复杂系统的不确定性处理方法,其中不确定性规约/建模、不确定性推理以及不确定性传播的不确定性处理方法,软件工程领域有较多的研究,但是自动化领域却很少涉及到这方面的研究,可以考虑将软件工程针对不确定性规约/建模、不确定性推理以及不确定性传播研究的理论和方法运用到自动化领域对这方面不确定性的研究中.

调研结果表明:针对复杂系统不确定性应用领域,软件工程的关注较为广泛,包含信息物理系统、物联网以及自适应信息物理系统,自动化领域主要关注于信息物理系统.针对复杂系统不确定性类型,软件工程和自动化都主要关注于外部以及内部不确定性,而传感器不确定性,只有软件工程领域的少量文献关注.针对复杂系统制品的不确定性,软件工程和自动化领域关于不确定性研究的交叉集中在模型不确定性、参数不确定性和数据不确定性上.针对复杂系统不确定性的处理方法,软件工程主要关注于不确定性规约/建模、不确定性推理、不确定性条件下的决策以及不确定性传播.自动化领域的研究者主要关注不确定性条件下的决策.

5 总结和发展趋势

在不确定性文献发表的数量和刊物方面,从总体趋势来看,过去 10 年中,关注复杂系统不确定性的研究呈现出上升趋势.这一发现说明不确定性研究越来越受到研究人员的关注.此外,我们调研的 142 篇文献发表在 112 个不同的期刊、会议、研讨会以及书的章节中;通过分析发表文献数量排名靠前的期刊和会议,我们发现《Automatica》是控制领域的知名期刊,《JSS》和《SoSyM》是软件工程领域的知名期刊,说明本研究领域具有多学科交叉的特点.该特点为本领域的研究带来了一定的挑战,同时也提供了多学科交叉研究的机遇.需要指出的是,这一多学科交叉研究的特点,主要的来源是此次文献调研中关注的信息物理系统和物联网自身的多学科性,如软件工程和自动化等.这一发现给研究人员提供了一些研究启示:在进行该领域的调研时,需要进行跨领域搜索.在跨学科特点的基础上,不确定性本身在理论上(如概率论、粗糙集)和方法上(如蒙特卡洛方法)的复杂性进一步加深了复杂系统不确定性研究的困难程度.

在不确定性文献涉及的关键词方面,针对复杂系统不确定性的热门研究主题,我们采用词频统计和关键词共现的聚类方法进行挖掘.通过对关键词词频进行统计,复杂系统不确定性主要针对信息物理系统和物联网等系统,主要关注自适应和复杂事件过程中的不确定性,以及关注处理不确定性相关技术包括基于模型的测试、决策制定和概率模型检测等.通过对关键词进行聚类得到 6 个热门研究主题:不确定性处理技术保证复杂系统的服务质量、不确定性处理技术保证复杂系统的性能和可靠性、不确定性处理技术支持系统进行决策分析和测试、不确定性处理技术支持不确定性决策、复杂系统开发和运行过程中的不确定性以及复杂系统的不确定性建模和测试.这一研究结果可以使研究人员更好地了解该领域的研究热点主题以及它们之间的关系.

在不确定性文献涉及的生命周期方面,我们针对 142 篇相关文献关注的系统开发生命周期的 14 个阶段进行分类,结果显示:设计定义阶段是研究人员关注最多的阶段;自 2013 年以后,复杂系统设计定义阶段的不确定性研究出现了明显的增长趋势.这个增长趋势与信息物理系统和物联网工业应用的增长趋势是一致的^[4].此外,我们也观察到,针对自适应信息物理系统的不确定性研究也受到了较多研究人员的关注,这主要是因为不确定性是自适应系统设计和运行需要考虑的关键因素^[34,83].然而,面向信息物理系统和物联网的不确定性研究依然处于初始阶段,这主要是因为:(1) 这方面的研究随着物联网和信息物理系统的兴起逐渐受到了关注;(2) 由于多学科交叉导致的系统本身的复杂性、其运行环境的复杂性以及不确定性的类型和级别的多样性(如软件决策的不确定性、硬件设备的不确定性、网络通信中的不确定性、特征交互导致的不确定性)导致此方向的研究难度较大^[4].我们同时也观察到,在软件工程领域,没有系统地支持复杂系统全生命周期不确定性研究的方法、工具平台、数据集、基准以及标准.未来的研究应该更多地着眼于尽可能地覆盖复杂系统全生命周期的多个阶段,贯穿不确定性需求、设计、管理与维护等方面的系统研究.

随着诸如机器学习等人工智能技术在信息物理系统和物联网的进一步应用,系统内部行为的不确定性显著增强,处理复杂系统内部不确定性将是未来研究的热点.据本文统计,11 篇文献关注了机器学习等人工智能技术在复杂系统中的应用导致的系统内容的不确定性.比如文献[84]从 3 个角度讨论这方面的不确定性:(1) 模型拟合相关的不确定性;(2) 通过传感器和人工输入等方式收集到的数据质量的问题导致的不确定性;(3) 模型应用超出其基于上下文决定的应用范围导致的模型输出不确定性.这类在人工智能领域的不确定性研究工作可以有机地与软件工程学科不确定性研究相结合,比如通过软件工程的方法提高数据质量从而在一定程度上保证训练模型的质量.目前这方面的研究大多集中在人工智能领域,比如 AI 系统的模型和(输入)数据的不确定性,

分别被认为是广泛意义下的认知(epistemic)不确定性和偶然(aleatory)不确定性;模型不确定性可以通过提供用户足够多的训练数据来消除,诸如测量误差等不可消除的不确定性.不确定性度量也是该研究领域比较关注的一个研究方向,比如剑桥大学的 Gal 和 Ghahramani^[85]提出的 Monte Carlo Dropout 方法可以在不改变神经网络结构或优化技术的情况下近似模型不确定性,谷歌的 Lakshminarayanan 等人提出的 Deep Ensemble^[86]是一种易于实现和并行化,需要很少的超参数调整的不确定性度量方法.我们认为,这些近期在人工智能领域的研究结果可以极大地促进软件工程领域针对复杂系统的不确定性的研究.此外,机器学习等技术也被广泛应用于处理复杂系统的不确定性,据本文统计,有 38 篇(38/142)文献利用机器学习等人工智能技术(如马尔可夫决策过程和贝叶斯网络等)处理复杂系统的不确定性.如文献[87]利用参数化的马尔可夫决策过程处理自适应系统运行时定量验证过程中出现的不确定性.

在不确定性不同层次的分类方面,讨论复杂系统外部、内部和传感器不确定性的文献中有 64 篇(49.6%)关注外部不确定性(包括传感器不确定性).这一观察结果符合信息物理系统和物联网直接与环境交互的特性以及运行环境具有实时变化的特点.这方面的研究涉及到的不确定性多种多样,包括用户行为的不确定性、环境的不确定性、系统的经济属性以及基础涉及的不确定性.在我们调研的文献中,针对内部不确定性的研究主要集中在系统结构不确定性、系统内部交互以及支持和处理系统运行的相关技术(如深度神经网络本身固有的不确定性)的不确定性上.由此可以看出,针对这些不同种类的不确定性进行细粒度和多维度的分类和特征化是非常有必要的.研究人员在这方面进行了一些工作,比如张等人提出的 U-Model^[13]是一个比较通用的描述不确定性及其特征的概念模型.Moreno 等人在文献[88]中详细讨论了如何通过减少不确定性来增强自适应系统的适应性决策应对各种不确定性的能力.基于以上方法,正在 OMG 对象管理组织制定的 Precise Semantics for Uncertainty Modeling(PSUM)从不确定性建模的角度对不确定性进行了分类.但是,目前仍缺乏一个比较完善的不确定性分类.我们希望这样的不确定性分类应考虑到内外部不确定性、主客观不确定性、偶然和认知不确定性、模型/参数和数据不确定性、人机交互不确定性、各制品不确定性等.

在不确定性涉及的开发制品的分类方面,对收集到的 65 篇明确涉及到制品不确定性的文献的调研结果表明,信息物理系统和物联网系统开发中关注的制品种类繁多(如需求、模型及其参数、数据、制品演化、约束等).此外,模型不确定性和数据不确定性是研究人员关注的重点,主要是因为需求建模、系统架构和设计建模是信息物理系统和物联网系统开发中使用的主要方法,数据采集和处理是此类系统需要考虑的重要设计因素之一.另外,我们也观察到,学术界目前并没有系统地考虑不同制品不确定性信息的相关性.比如,从模型不确定性到需求不确定性的可追溯性(traceability)并没有受到足够的关注.整体来说,这种欠缺在某种程度上是对产品质量及其运行和维护的一个不可忽视的威胁,但其同时也是未来可能的发展趋势.此外,如 Salay 等人在文献[89]中提到,制品之间的追溯链接本身也会有不确定性,所以在管理追溯链接的同时也需要系统地管理不确定性.

在不确定性的处理方法方面,研究人员处理不确定性的方法多种多样,如不确定性传播、不确定性下的决策、不确定性推理、不确定性规约/建模、不确定性管理以及不确定性度量.其中,不确定性下的决策、不确定性推理以及不确定性规约/建模是研究人员关注的重点.更细粒度的不确定性分类有助于将具体的不确定性条件下的决策处理方法与具体的不确定性类型以及相关的理论和技术(如贝叶斯概率和模糊集)联系起来.未来,我们将着重开展这方面的工作.

在不确定性研究的学科领域方面,为了防止自动化领域和计算机领域的文献数据量差异带来的分析偏差,本文重点针对自动化领域和计算机的子领域(软件工程领域)关于复杂系统不确定性在不确定性应用领域、不确定性类型、系统制品的不确定性类型以及不确定性处理方法这 4 个方面展开研究.(1) 关于不确定性应用领域,软件工程领域更多地关注信息物理系统、物联网以及自适应信息物理系统,而自动化领域更多地关注信息物理系统.(2) 关于不确定性类型,软件工程领域和自动化领域对外部不确定性和内部不确定性有较多的研究,而对于传感器不确定性,软件工程领域研究者关注得较少,自动化领域的研究者没有涉及到这方面的研究.(3) 关于系统制品的不确定性类型,软件工程和自动化领域关于不确定性的研究交叉集中在模型不确定性、参数不确定性和数据不确定性上.这一现象是软件工程和自动化学科特性所导致的,如需求不确定性等是软件工

程所特有的。(4) 关于不确定性处理方法,软件工程和自动化都主要关注于不确定性条件下的决策,在探究该方法处理复杂系统开发和运行过程中的不确定性时,可以借助这两个领域的优势,以便有效地处理不确定性。通过对自动化领域和软件工程领域在以上 4 个方面的探索,有助于复杂系统不确定性在这两个领域的交叉研究,借助两个领域理论方法的优势互补(如通过软件工程结构模型和自动化的数学模型两个维度表征不确定性)进行更深层次的研究。

6 结束语

不确定性问题对于复杂系统的开发和运行会产生重要的影响,引起了研究人员的广泛关注。因此,针对复杂系统的不确定性研究是当前的热点。本文通过对该领域的相关文献进行系统性的分析和研究,总结了复杂系统不确定性涉及的多个方面,重点包括复杂系统的不确定性的分类、不同系统开发阶段的分布情况以及研究人员提出的处理不确定性的主要方法。在此基础上,本文还进一步总结和归纳了该领域未来的发展趋势,以帮助研究人员从整体上快速把握该领域的研究现状,吸引更多的研究人员投入到这个热点研究领域。我们未来的研究工作侧重在如下 5 个方面。

- **构建更加通用、全面、可扩展的针对信息物理系统和物联网等复杂系统的不确定性分类标准。**复杂系统及其开发和运行的不确定性种类多种多样,且多数情况下与信息物理系统和物联网的具体设计 and 应用领域有关,目前还缺少统一的分类标准,我们将针对信息物理系统和物联网等复杂系统的不确定性展开全面且深入的研究,构建一个通用的、全面的、针对不同应用领域特征的可扩展的分类标准。

- **深入研究复杂系统的不确定性处理方法、相关理论和技术的应用,且系统、全面地建立不确定性处理方法与不同类别不确定性之间的关系。**为了处理复杂系统开发和运行过程中出现的各种不确定性,研究人员采取了不同的方法对不确定性进行处理:经过对 142 篇文献的整理和分析,本文总结出六大类不确定性处理方法(如不确定性传播)。由于种类较多且每种不确定处理方法所应用的技术复杂多样(如不确定性推理方法可以基于贝叶斯推理和模糊逻辑等),本文无法展开详细讨论。所以,基于本文现有的研究,我们将对每种不确定性处理方法、相关理论和技术应用进行详细的分类和特征提炼,且将不确定性处理方法和不确定性类型关联起来。

- **系统且深入地研究复杂系统开发过程中各种不确定性间的相关性。**为了更好地应对系统开发过程中的各种不确定性,未来需要对软件开发过程中各种不确定性的本质进行更详细的分类和特性提炼,进而更好地定位不确定性的来源以及对应的解决方法。比如,对于需求阶段,需要从需求获取、需求分析、需求规约、需求验证以及需求管理等方面进行更细粒度的不确定性分析。在此基础上,探究需求阶段和系统设计、运行等不同系统开发阶段与不确定性的相关性,并探讨沿着系统全生命周期如何系统地降低各种不确定性,如何分析不同开发阶段之间的不确定性影响,以有效地保证所开发的复杂系统的质量以及提升其在运行阶段应对不确定性的能力。

- **深入探究机器学习等人工智能技术对复杂系统不确定性的影响和机遇。**随着诸如机器学习等人工智能技术的快速发展,越来越多的研究者将机器学习等人工智能技术引入到信息物理系统和物联网等复杂系统不确定性的处理中来。机器学习等人工智能技术的应用给复杂系统不确定性的处理带来了便利(如利用机器学习等人工智能技术更加有效地进行不确定性决策),同时也引入了诸多的副作用(如深度学习中深度神经网络本身固有的不确定性以及机器学习技术预测不确定性或预测错误增加复杂系统内部的不确定)。目前研究工作的视角多以解决问题为主,而缺乏不确定性和机器学习算法的相关性。为了更好地利用机器学习技术去处理复杂系统的不确定性,需要对复杂系统的不确定性本质进行更加准确的分类和特征提炼,进而更好地定位不确定性亟待解决的具体问题,从而更精准地将不确定性问题映射为现有的机器学习方法能解决的问题。此外,我们也需要建立机器学习等人工智能技术的固有不确定性与复杂系统的其他内外部不确定性之间的关系,从而为系统地分析复杂系统开发“端对端”全生命周期的不确定性提供基础。

- **深入探究自动化和计算机领域(尤其是软件工程)不确定性研究的相关性。**复杂系统不确定性研究具有多学科交叉特性,对于不确定性的研究,自动化和计算机领域(软件工程)研究人员关注的研究主题有所差异。由于

本文针对不确定性处理方法的调研未能深层次挖掘不确定性处理方法的差异,对于不确定性的分类粒度较大,对制品不确定性的分类更多地是基于软件工程的角度的,所以未能更好地探究自动化领域的不确定性类型和特征,未能更深层地探究两个领域的研究差异.未来我们将综合考虑自动化和计算机领域(尤其是软件工程)的差异,科学地分析两者研究不确定性的类型以及不确定性方法,从而进行两个领域的优势互补,为其他研究者提供一些参考.

References:

- [1] Lee EA. Cyber physical systems: Design challenges. In: Proc. of the 11th IEEE Int'l Symp. on Object and Component-oriented Real-time Distributed Computing (ISORC). IEEE, 2008. 363–369.
- [2] Uckelmann D, Harrison M, Michahelles F. Architecting the Internet of Things. Springer Science & Business Media, 2011.
- [3] Salton G. The Smart System. In: Retrieval Results and Future Plans. 1971.
- [4] Greer C, Burns M, Wollman D, Griffor E. Cyber-physical systems and Internet of Things. NIST Special Publication, 2019(202): 52,1900.
- [5] Lindsay J. Aristotle and the criterion of truth. The Monist, 1921,31(3):470–475.
- [6] Halpern JY. Reasoning about Uncertainty. MIT Press, 2017.
- [7] Der Kiureghian A, Ditlevsen O. Aleatory or epistemic? does it matter? Structural Safety, 2009,31(2):105–112.
- [8] Matthies HG. Quantifying uncertainty: Modern computational representation of probability and applications. In: Proc. of the Extreme Man-made and Natural Hazards in Dynamics of Structures, Springer-Verlag, 2007. 105–135.
- [9] Gerostathopoulos I, Skoda D, Plasil F, Bures T, Knauss A. Architectural homeostasis in self-adaptive software-intensive cyber-physical systems. In: Proc. of the European Conf. on Software Architecture. Springer-Verlag, 2016. 113–128.
- [10] Zavala E, Franch X, Marco J, Knauss A, Damian D. Sacre: Supporting contextual requirements' adaptation in modern self-adaptive systems in the presence of uncertainty at runtime. Expert Systems with Applications, 2018,98:166–188.
- [11] Geisberger E, Broy M. AgendaCPS. Berlin, Heidelberg: Springer-Verlag, 2012.
- [12] Gu X, Easwaran A. Towards safe machine learning for CPS: Infer uncertainty from training data. In: Proc. of the 10th ACM/IEEE Int'l Conf. on Cyber-physical Systems. 2019. 249–258.
- [13] Zhang M, Bran S, Shaukat A, Yue T, Okariz O, Norgren R. Understanding uncertainty in cyber-physical systems: A conceptual model. In: Proc. of the European Conf. on Modelling Foundations and Applications. Springer-Verlag, 2016. 247–264.
- [14] Esfahani N, Kouroshfar E, Malek S. Taming uncertainty in self-adaptive software. In: Proc. of the 19th ACM SIGSOFT Symp. and the 13th European Conf. on Foundations of Software Engineering. 2011. 234–244.
- [15] Marinho M, Sampaio S, Lima T, Moura H. A systematic review of uncertainties in software project management. arXiv Preprint arXiv: 1412.3690, 2014.
- [16] Asan U, Soyer A. Failure mode and effects analysis under uncertainty: A literature review and tutorial. In: Intelligent Decision Making in Quality Management. Springer-Verlag, 2016. 265–325.
- [17] Shevtsov S, Berekmeri M, Weyns D, Maggio M. Control-theoretical software adaptation: A systematic literature review. IEEE Trans. on Software Engineering, 2017,44(8):784–810.
- [18] Salih A, Omar M, Yasin A. Understanding uncertainty of software requirements engineering: A systematic literature review protocol. In: Proc. of the Asia Pacific Requirements Engineering Conf. on Requirements Engineering for Internet of Things (APRES 2017). 2018. 164–171.
- [19] Zhang MY, Jin Z, Zhao HY, Luo YX. Survey of machine learning enabled software self-adaptation. Ruan Jian Xue Bao/Journal of Software, 2020,31(8):2404–2431 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6076.htm> [doi: 10.13328/j.cnki.jos.006076]
- [20] Kitchenham B, Brereton OP, Budgen D, Turner M, Bailey J, Linkman S. Systematic literature reviews in software engineering—a systematic literature review. Information and Software Technology, 2009,51(1):7–15.
- [21] Kitchenham B, Charters S. Guidelines for performing systematic literature reviews in software engineering. Technical Report, 2007.
- [22] Dybå T, Dingsøyr T. Empirical studies of agile software development: A systematic review. Information and Software Technology, 2008,50(9/10):833–859.

- [23] Uzun B, Tekinerdogan B. Model-driven architecture based testing: A systematic literature review. *Formation and Software Technology*, 2018,102:30–48.
- [24] Achimugu P, Selamat A, Ibrahim R, *et al.* A systematic literature review of software requirements prioritization research. *Information and Software Technology*, 2014,56(6):568–585.
- [25] Ding W, Liang P, Tang A, Van Vliet H. Knowledge-based approaches in software documentation: A systematic literature review. *Information and Software Technology*, 2014,56(6):545–567.
- [26] Kitchenham B. Evidence-based software engineering and systematic literature reviews. In: *Proc. of the Int'l Conf. on Product Focused Software Process Improvement*. Springer-Verlag, 2006. 3.
- [27] Mahdavi-Hezavehi S, Galster M, Avgeriou P. Variability in quality attributes of service-based software systems: A systematic literature review. *Information and Software Technology*, 2013;55(2):320–343.
- [28] Dermeval D, Vilela J, Bittencourt II, Castro J, Isotani S, Brito P, Silva A. Applications of ontologies in requirements engineering: A systematic review of the literature. *Requirements Engineering*, 2016,21(4):405–437.
- [29] Van Eck NJ, Waltman L. Text mining and visualization using VoSviewer. *arXiv Preprint arXiv: 1109.2058*, 2011.
- [30] Waltman L, Van Eck NJ, Noyons ECM. A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 2010,4(4):629–635.
- [31] ISO/IEC/IEEE 15288. Systems and software engineering—System life cycle processes. 2015.
- [32] Kagermann H, Wahlster W, Helbig J. Securing the future of german manufacturing industry: Recommendations for implementing the strategic initiative industrie 4.0. *Final Report of the Industrie*, 2013,4.
- [33] De Lemos R, Giese H, Müller HA, Shaw M, Andersson J, Litoiu M, Schmerl B, Tamura G, Villegas NM, Vogel T, *et al.* Software engineering for self-adaptive systems: A second research roadmap. In: *Software Engineering for Self-adaptive Systems II*. Springer-Verlag, 2013. 1–32.
- [34] Esfahani N, Malek S. Uncertainty in self-adaptive software systems. In: *Software Engineering for Self-Adaptive Systems II*. Springer-Verlag, 2013. 214–238.
- [35] Cheng B, Lemos R, Giese H, Inverardi P, Magee J, Andersson J, Becker B, Bencomo N, Brun Y, Cukic B, Serugendo GDM, Dustdar S, Finkelstein A, Gacek C, Geihs K, Grassi V, Karsai G, Kienle H, Kramer J, Whittle J. Software engineering for self-adaptive systems: A research roadmap. In: *Software Engineering for Self-adaptive Systems*. 2009,5525:1–26.
- [36] Ahmad M, Gnaho C, Bruel JM, Laleau R. Towards a requirements engineering approach for capturing uncertainty in cyber-physical systems environment. In: *Proc. of the Int'l Conf. on Model and Data Engineering*, Springer-Verlag, 2018. 115–129.
- [37] Ramparany F, Mondy R, Demazeau Y. A semantic approach for managing trust and uncertainty in distributed systems environments. In: *Proc of the 21st Int'l Conf. on Engineering of Complex Computer Systems (ICECCS)*. 2016. 63–70.
- [38] Yang P, Hanneghan M, Qi J, Deng Z, Dong F, Fan D. Improving the validity of lifelogging physical activity measures in an Internet of Things environment. In: *Proc. of the 2015 IEEE Int'l Conf. on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*. IEEE, 2015. 2309–2314.
- [39] Bordel B, Alcarria R, De Andrés DM, You I. Securing Internet-of-Things systems through implicit and explicit reputation models. *IEEE Access*, 2018,6:47472–47488.
- [40] Banerjee S, Bouzeffrane S, Mühlethaler P. Mobility prediction in vehicular networks: An approach through hybrid neural networks under uncertainty. In: *Proc. of the Int'l Conf. on Mobile, Secure, and Programmable Networking*. Springer-Verlag, 2017. 178–194.
- [41] Francalanza E, Borg J, Constantinescu C. A knowledge-based tool for designing cyber physical production systems. *Computers in Industry*, 2017,84:39–58.
- [42] Hashemi-Dezaki H, Askarian-Abyaneh H, Shams-Ansari A, DehghaniSanij M, Hejazi MA. Direct cyber-power interdependencies-based reliability evaluation of smart grids including wind/solar/diesel distributed generations and plug-in hybrid electrical vehicles. *Int'l Journal of Electrical Power & Energy Systems*, 2017,93:1–14.
- [43] Cámara J, Garlan D, Schmerl B. Synthesis and quantitative verification of tradeoff spaces for families of software systems. In: *Proc. of the European Conf. on Software Architecture*. Springer-Verlag, 2017. 3–21.

- [44] Agarwal S, Pape LE, Dagli CH, Ergin NK, Enke D, Gosavi A, Qin R, Konur D, Wang R, Gottapu RD. Flexible and intelligent learning architectures for SOS (Fila-SOS): Architectural evolution in systems-of-systems. *Procedia Computer Science*, 2015,44(C): 76–85.
- [45] Li HY, Zhang YR, Chen YX. Pstep—a novel probabilistic event processing language for uncertain spatio-temporal event streams of Internet of vehicles. In: *Proc. of the 2015 IEEE Int'l Conf. on Software Quality, Reliability and Security-companion*. IEEE, 2015. 161–168.
- [46] Amir M, Givargis T. Hybrid state machine model for fast model predictive control: Application to path tracking. In: *Proc. of the 2017 IEEE/ACM Int'l Conf. on Computer-aided Design (ICCAD)*. IEEE, 2017. 185–192.
- [47] Bortolussi L, Cairoli F, Paoletti N, Stoller SD. Conformal predictions for hybrid system state classification. In: *From Reactive Systems to Cyber-Physical Systems*. Springer-Verlag, 2019. 225–241.
- [48] Bacci G, Hansen M, Larsen KG. On the verification of weighted kripke structures under uncertainty. In: *Proc. of the Int'l Conf. on Quantitative Evaluation of Systems*. Springer-Verlag, 2018. 71–86.
- [49] Vrbaski M, Bolic M, Majumdar S. Complex event recognition notification methodology for uncertain iot systems based onmicro-service architecture. In: *Proc. of the 6th IEEE Int'l Conf. on Future Internet of Things and Cloud (FiCloud)*. IEEE, 2018. 184–191.
- [50] Behl M, Nghiem TX, Mangharam R. Model-IQ: Uncertainty propagation from sensing to modeling and control in buildings. In: *Proc. of the 2014 ACM/IEEE Int'l Conf. on Cyber-physical Systems (ICCPS)*. IEEE, 2014. 13–24.
- [51] Bhadoria RS, Chaudhari NS, Samanta S. Uncertainty in sensor data acquisition for SOA system. *Neural Computing and Applications*, 2018,30(10):3177–3187.
- [52] Cámara J, Peng WX, Garlan D, Schmerl B. Reasoning about sensing uncertainty in decision-making for self-adaptation. In: *Proc. of the Int'l Conf. on Software Engineering and Formal Methods*. Springer-Verlag, 2017. 523–540.
- [53] Chatzikonstantinou G, Kontogiannis K. Efficient parallel reasoning on fuzzy goal models for run time requirements verification. *Software & Systems Modeling*, 2018,17(4):1339–1364.
- [54] Ghezzi C. Dependability of adaptable and evolvable distributed systems. In: *Proc. of the Int'l School on Formal Methods for the Design of Computer, Communication and Software Systems*. Springer-Verlag, 2016. 36–60.
- [55] Shevtsov S, Weyns D, Maggio M. Simca* a control-theoretic approach to handle uncertainty in self-adaptive systems with guarantees. *ACM Trans. on Autonomous and Adaptive Systems (TAAS)*, 2019,13(4):1–34.
- [56] Ge H, Yue D, Xie XP, Deng S, Dou CX. A unified modeling of muti-sources cyber-attacks with uncertainties for CPS security control. *Journal of the Franklin Institute*, 2019.
- [57] Zhe Xu, Julius AA. Robust temporal logic inference for provably correct fault detection and privacy preservation of switched systems. *IEEE Systems Journal*, 2019,13(3):3010–3021.
- [58] Yang ZF, Wu M, Lin W. An efficient framework for barrier certificate generation of uncertain nonlinear hybrid systems. *Nonlinear Analysis: Hybrid Systems*, 2020,36:100837.
- [59] Chen B, Niu YG, Zou YY. Security control for Markov jump system with adversarial attacks and unknown transition rates viaadaptive sliding mode technique. *Journal of the Franklin Institute*, 2019,356(6):3333–3352.
- [60] Chen C, Chen D, Yan YN, Zhang GF, Zhou QG, Zhou R. Integration of numerical model and cloud computing. *Future Generation Computer Systems*, 2018,79:396–407.
- [61] Song W, Li WF, Fu XW, Cao YL, Yang L. Rfid based real-time manufacturing information perception and processing. In: *Proc. of the Int'l Conf. on Algorithms and Architectures for Parallel Processing*. Springer-Verlag, 2013. 303–310.
- [62] Chen WH, An JY, Li RF, Fu L, Xie GQ, Bhuiyan MZA, Li KQ. A novel fuzzy deep-learning approach to trafficflow prediction with uncertain spatial-temporal data features. *Future Generation Computer Systems*, 2018,89:78–88.
- [63] Jaribion A, Khajavi SH, Hossein Motlagh N, Holmström J. [wip] a novel method for big data analytics and summarization based on fuzzysimilarity measure. In: *Proc. of the 11th IEEE Conf. on Service-oriented Computing and Applications (SOCA)*. 2018. 221–226.
- [64] Jha S, Raman V. On optimal control of stochastic linear hybrid systems. In: *Proc. of the Int'l Conf. on Formal Modeling and Analysis of Timed Systems*. Springer-Verlag, 2016. 69–84.

- [65] Kate S, Ore JP, Zhang XY, Elbaum S, Xu ZG. Phys: Probabilistic physical unit assignment and inconsistency detection. In: Proc. of the 26th ACM Joint Meeting on European Software Engineering Conf. and Symp. on the Foundations of Software Engineering. 2018. 563–573.
- [66] Ravi M, Demazeau Y, Ramparany F. Reasoning with trust and uncertainty illustration in the internet of things. In: Proc. of the 2015 IEEE/WIC/ACM Int'l Conf. on Web Intelligence and Intelligent Agent Technology (WI-IAT). 2015(2):125–128.
- [67] Shin SY, Nejati S, Sabetzadeh M, Briand LC, Zimmer F. Test case prioritization for acceptance testing of cyberphysical systems: A multi-objective search-based approach. In: Proc. of the 27th ACM SIGSOFT Int'l Symp. on Software Testing and Analysis. 2018. 49–60.
- [68] Cimatti A, Micheli A, Roveri M. Timelines with temporal uncertainty. In: proc. of the 27th AAAI Conf. on Artificial Intelligence. 2013.
- [69] Zhang M, Shaukat A, Yue T, *et al.* Uncertainty-wise cyber-physical system test modeling. *Software & Systems Modeling*, 2019, 18(2):1379–1418.
- [70] Zhang M, Shaukat A, Yue T. Uncertainty-wise test case generation and minimization for cyber-physical systems. *Journal of Systems and Software*, 2019, 153:1–21.
- [71] Ma T, Shaukat A, Yue T. Modeling foundations for executable model-based testing of self-healing cyber-physical systems. *Software & Systems Modeling*, 2019, 18(5):2843–2873.
- [72] Potter K, Rosen P, Johnson CR. From quantification to visualization: A taxonomy of uncertainty visualization approaches. In: Proc. of the IFIP Working Conf. on Uncertainty Quantification. Springer-Verlag, 2011. 226–249.
- [73] De Rocquigny E, Devictor N, Tarantola S. *Uncertainty in Industrial Practice: A Guide to Quantitative Uncertainty Management*. John Wiley & Sons, 2008.
- [74] Mathias KE, Nixon MR, Talbot PJ. Uncertainty management in a decision-making system. *US Patent*, 2009, 7:606,784.
- [75] Kellermann K, Reynolds R. When ignorance is Bliss: The role of motivation to reduce uncertainty in uncertainty reduction theory. *Human Communication Research*, 1990, 17(1):5–75.
- [76] Bertoa MF, Burgueño L, Moreno N, *et al.* Incorporating measurement uncertainty into OCL/UML primitive datatypes. *Software & Systems Modeling*, 2019, 1–27.
- [77] Aziz AR, *et al.* *Decision theory with imperfect information*. World Scientific, 2014(10).
- [78] Akbar A, Kousiouris G, Pervaiz H, Sancho J, Ta-Shma P, Carrez F, Moessner K. Real-time probabilistic datafusion for large-scale IoT applications. *IEEE Access*, 2018, 6:10015–10027.
- [79] Cámara J, Garlan D, Schmerl B. Synthesizing tradeoff spaces with quantitative guarantees for families of software systems. *Journal of Systems and Software*, 2019, 152:33–49.
- [80] Chaves LC, Ismail HI, Bessa IV, Cordeiro LC, de Lima Filho EB. Verifying fragility in digital systems with uncertainties using dsverifier v2.0. *Journal of Systems and Software*, 2019, 153:22–43.
- [81] Sharifloo A, Metzger A. MCaaS: Model checking in the cloud for assurances of adaptive systems. In: *Software Engineering for Self-adaptive Systems III*. 2017. 137–153.
- [82] Kasprzyk JR, Nataraj S, Reed PM, Lempert RJ. Many objective robust decision making for complex environment alsystems undergoing change. *Environmental Modelling & Software*, 2013, 42:55–71.
- [83] Islam N, Azim A. Assuring the runtime behavior of self-adaptive cyber-physical systems using feature modeling. In: Proc. of the 28th Annual Int'l Conf. on Computer Science and Software Engineering. 2018. 48–59.
- [84] Kläs M, Vollmer AM. Uncertainty in machine learning applications: A practice-driven classification of uncertainty. In: Proc. of the Int'l Conf. on Computer Safety, Reliability, and Security. Springer-Verlag, 2018. 431–438.
- [85] Gal Y, Ghahramani Z. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In: Proc. of the Int'l Conf. on Machine Learning. 2016. 1050–1059.
- [86] Lakshminarayanan B, Pritzel A, Blundell C. Simple and scalable predictive uncertainty estimation using deep ensembles. In: *Advances in Neural Information Processing Systems*. 2017. 6402–6413.
- [87] Nia MA, Kargahi M, Faghih F. Probabilistic approximation of runtime quantitative verification in self-adaptive systems. *Microprocessors and Microsystems*, 2020, 72:102943.

- [88] Moreno GA, Cámara J, Garlan D, *et al.* Uncertainty reduction in self-adaptive systems. In: Proc. of the 13th Int'l Conf. on Software Engineering for Adaptive and Self-Managing Systems. 2018. 51–57.
- [89] Salay R, Chechik M, Horkoff J, *et al.* Managing requirements uncertainty with partial models. Requirements Engineering, 2013, 18(2):107–128.

附中文参考文献:

- [19] 张明悦,金芝,赵海燕,罗懿行.基于机器学习的软件自适应性.软件学报,2020,31(8):2404–2431. <http://www.jos.org.cn/1000-9825/6076.htm> [doi: 10.13328/j.cnki.jos.006076]



檀超(1991—),男,博士生,CCF 学生会员,主要研究领域为不确定性软件工程,模型驱动工程.



张静宣(1988—),男,博士,讲师,CCF 专业会员,主要研究领域为软件仓库挖掘,智能软件工程.



王铁鑫(1987—),男,博士,讲师,CCF 专业会员,主要研究领域为基于模型的系统工程,模型驱动方法,领域本体与语义检测.



岳涛(1974—),女,博士,博士生导师,CCF 专业会员,主要研究领域为模型驱动工程,需求工程,产品线工程,基于搜索的软件工程,经验软件工程.