

# 面向电网调度自动化系统鲁棒性的 验证用例生成方法研究

路 辉<sup>1</sup>, 王诗琪<sup>1</sup>, 万 雄<sup>2</sup>, 张若柳<sup>1</sup>, 裘雨音<sup>3</sup>, 申建强<sup>4</sup>

(1. 北京航空航天大学 电子信息工程学院, 北京 100191;

2. 国家电网有限公司国家调度控制中心, 北京 100031;

3. 国网浙江省电力有限公司, 杭州 310007;

4. 国网湖州供电公司, 浙江 湖州 313000)

**摘要:** 电网调度自动化系统的稳定运行是保障电力系统安全的核心; 对于电网调度自动化系统进行鲁棒性测试, 是保障系统面对异常、极端输入下性能良好的前提; 对面向电网调度自动化系统鲁棒性的验证用例生成方法进行了综合分析; 对电网调度自动化系统的特点与测试难点、现存电网调度自动化系统验证用例生成方法进行了分析, 发现了现存方法在电网调度自动化系统鲁棒性测试方面的不适用性; 研究了面向一般复杂系统鲁棒性测试的验证用例生成的方法框架; 对于框架中涉及到的代理模型在不同维度下的拟合精度与速度进行了实验分析, 经实验测试发现高斯过程分类模型适合于低维问题; 支持向量机模型适合于中高维问题; 对于超过 20 维的问题, 人工神经网络的拟合速度达到支持向量机模型的 20 倍, 并且能达到令人满意的精度; 详细分析了将鲁棒性测试验证用例生成框架用于电网调度自动化系统亟待解决的问题和可能存在的技术途径。

**关键词:** 电网调度自动化系统; 电力系统; 鲁棒性测试; 验证用例生成; 自适应采样

## Research on the Verification Case Generation Method for Robustness of Power Dispatching Automation Systems

LU Hui<sup>1</sup>, WANG Shiqi<sup>1</sup>, WAN Xiong<sup>2</sup>, ZHANG Ruoliu<sup>1</sup>, QIU Yuyin<sup>3</sup>, SHEN Jianqiang<sup>4</sup>

(1. School of Electronic and Information Engineering, Beihang University, Beijing 100191, China;

2. National Electric Power Dispatch and Control Center of State Grid Corporation of China, Beijing 100031, China;

3. State Grid Zhejiang Electric Power Limited Company, Hangzhou 310007, China;

4. State Grid Huzhou Electric Power Supply Company, Huzhou 313000, China)

**Abstract:** The stable operation of the power dispatching automation system is essential for ensuring power system security. Conducting robustness testing on such systems is a prerequisite for guaranteeing their performance under abnormal and extreme conditions. A comprehensive analysis was carried out on methods for generating verification cases aimed at evaluating the robustness of power dispatching automation systems. This study examined the characteristics and testing challenges of power dispatching automation systems, as well as existing methods for generating verification cases. The inadequacy of current approaches in robustness testing for power dispatching automation systems was identified. A methodological framework for generating test cases for testing other complex system robustness was investigated. The analysis focusing on fitting accuracy and speed of variant surrogate models was conducted. Experimental tests showed that the Gaussian process classification model suits low-dimensional problems. The support vector machine model fits medium to high-dimensional problems. For problems exceeding 20 dimensions, the artificial neural network's fitting speed is 20 times faster than the support vector machine model. It also achieves satisfactory accuracy. The unresolved issues and potential technical pathways for applying this

收稿日期: 2025-09-04; 修回日期: 2025-10-10。

基金项目: 国家电网公司科学技术项目(5100-202455380A-3-2-ZX); 北京航空航天大学博士研究生卓越学术基金。

作者简介: 路 辉(1977-), 女, 工学博士。

引用格式: 路 辉, 王诗琪, 万 雄, 等. 面向电网调度自动化系统鲁棒性的验证用例生成方法研究[J]. 计算机测量与控制, 2026, 34(1): 259-268.

framework to power dispatching automation systems was further analyzed.

**Keywords:** power dispatching automation systems; power system; robustness testing; verification case generation; adaptive sampling

## 0 引言

电网调度自动化系统是电力系统的核心组成部分,是综合利用计算机、远动和远程通信技术,对电网一次设备进行运行监视、控制和调度的自动化系统<sup>[1]</sup>,承担着实时监控、调度控制的核心功能<sup>[2]</sup>。电网调度自动化系统集成数据采集、控制、通信和分析决策功能于一身,为各类应用提供统一的模型、数据、网络通信、人机界面、系统管理以及分析计算等服务<sup>[3]</sup>。电网调度自动化系统的稳定、可靠、安全运行是保证电力系统安全以及经济社会稳定发展的基础。然而,电网调度自动化系统面临着大规模、具有较高不确定性的输入以及多种外部环境的影响,由于输入异常、人为不当操作以及不完善的防误机制造成的电网事故时有发生,对于电网安全运行造成重要影响,对于电网调度自动化系统的测试验证是保障其稳定可靠、降低运行风险的重要环节。

2014年11月,国家电网公司正式发布执行 Q/GDW 11205-2014《电网调度自动化系统软件通用测试规范》<sup>[1]</sup>,该标准详细规定了电网调度自动化系统软件测试的测试对象、测试环境以及测试方法,将对于电网调度自动化系统软件的测试按照不同目的分为:源代码测试、功能测试、性能测试、标准符合测试、兼容性测试、安全性测试以及稳定性测试。其中稳定性测试既包括可靠性测试(重点关注系统无错误运行时间),也包括鲁棒性测试(重点关注系统对于异常数据的处理能力)。已有一些学者对于电网调度自动化系统的测试验证展开了研究<sup>[4-6]</sup>,但他们大多集中于对于电网调度自动化系统的基本功能进行验证,或对于软件代码进行测试,测试方法也通常局限于人工设置验证用例或代码走查等面向传统软件的测试方法,对于电网调度自动化系统鲁棒性的测试验证的研究较少。但是,基本功能验证以及软件代码测试只能对于设计者已经考虑到的软件逻辑以及防误措施进行测试验证,无法覆盖所有系统实际运行时可能面临的极端或异常输入,无法发现系统设计时未考虑到的未知错误。通过基本功能以及软件代码测试无误的系统,在实际运行面临多变、不确定的输入时,依然可能会出现未知的错误。为了全面衡量电网调度自动化系统在面对未知、可能存在扰动的输入下的稳定性以及面临极端情况下的正确性,对于系统进行鲁棒性测试非常重要。

IEEE组织对于一般软件的鲁棒性的定义为:在压力条件以及扰动下系统保持正确的能力<sup>[7]</sup>。由于鲁棒性重点衡量系统在异常输入下保持正确的能力,对于复杂

系统的鲁棒性的测试与验证通常基于验证用例生成的方法进行,通过生成边界验证用例或失效验证用例等对于系统具有挑战性的输入,来对于系统进行鲁棒性测试。目前已有大量的学者对于包括自动驾驶汽车<sup>[8]</sup>、船舶避碰系统<sup>[9]</sup>在内的多种复杂系统展开了鲁棒性验证用例生成方法的研究,然而,面向电网调度自动化系统鲁棒性的验证用例生成的研究还较少。一方面,由于电网调度自动化系统具有高并发性和高实时性的特点,面临的输入数据量巨大且为时间序列,为验证用例生成带来很大的挑战;另一方面,电网调度自动化系统性能的影响因素众多,存在较大不确定性,进一步加剧了验证用例生成的难度。随着信息技术以及新能源产业的不断发展,电力系统逐步趋向智能化,电网调度自动化系统引入了新型自动控制软件和智能化调度算法<sup>[2]</sup>,其所面临的新能源场站的数据具有更大的不确定性,这为电网调度自动化系统的鲁棒性验证用例生成带来了更大的挑战。

综合考虑对于电网调度自动化系统进行鲁棒性测试的需求以及验证用例生成所面临的巨大挑战,本文对于面向电网调度自动化系统鲁棒性的验证用例生成方法进行了研究,详细论述了电网调度自动化系统的特点以及测试难点,分析了现有电网自动化系统测试验证方法,介绍了一般复杂系统鲁棒性验证用例的生成框架及鲁棒性评估指标,对于框架中涉及到的不同代理模型对于维度的适用性进行了实验分析,并结合电网调度自动化系统的特点以及验证用例生成框架阐明了面向电网调度自动化系统鲁棒性的验证用例生成的关键挑战以及未来研究方向。

## 1 电网调度自动化系统的特点与鲁棒性测试的难点

### 1.1 电网调度自动化系统的特点

电网调度自动化系统所面临的海量时序输入数据、控制策略与防误机制的复杂逻辑、高实时性与高并发处理要求,是其区别于其他系统的主要特征。本文将电网调度自动化系统的特点总结如下。

#### 1) 高并发输入与高实时性响应:

不同于一般系统的数据输入,电网调度自动化系统会面临上千个控制区域和控制对象的实时数据、市场数据、计划数据等数据同时作为输入,数据量十分庞大,且各控制对象数据间并无耦合关系,每个对象的数据都可能同时会出现异常。另外,系统必须对于数据进行实时响应和处理,若电网发生故障,短时间内就会积攒大量数据。例如,对于自动发电控制软件(AGC, automa-

ted generation control), 要求数据读取周期不大于5秒。

### 2) 系统逻辑复杂:

电网调度自动化系统通常具有多种调度模式和控制逻辑,除了这些复杂的调度控制逻辑外,还具有一系列数据质量校验环节和防误措施,这些环节的错误都将导致系统最终输出出错。若将电网调度自动化系统整体视为黑盒,其内部的复杂功能和逻辑在测试时需要着重考虑。

### 3) 影响因素众多:

电网调度自动化系统输出正确与否,不仅与输入数据有关,还与人为操作、软件参数设置以及外部环境有关,人工操作错误以及设备故障等都有可能造成调度结果出错。而人工操作错误与设备故障具有很强的不确定性和偶发性,难以通过仿真手段复现及遍历。

## 1.2 电网调度自动化系统鲁棒性测试难点

电网调度自动化系统的鲁棒性测试需求为确保在各种正常/异常输入组合下,系统的控制逻辑正确且防误机制有效拦截错误数据/指令。考虑到鲁棒性测试的重点是考察异常或极端输入对于系统的影响,测试时只需考虑系统对于这些输入响应,因此,鲁棒性测试一般基于黑盒测试的方法进行,通过产生对系统具有挑战性的验证用例作为输入,分析系统对于这些验证用例的响应,从而完成对于系统的鲁棒性测试。为了对于系统进行摸底探底,需要充分探索所有可能导致电网调度自动化系统出现错误的输入,从而设计新的防误措施,保证系统稳定正确运行。电网调度自动化系统复杂输入、逻辑与输出的特点为面向其鲁棒性测试的验证用例生成带来了状态空间爆炸、故障模式隐蔽、极端输入占比小和高测试效率需求等难点,具体分析如下。

### 1) 状态空间维数爆炸:

状态空间由所有会对系统输出产生影响的因素构成,验证用例搜索将在此空间内进行,每一个验证用例对应一组影响因素的取值组合。对于电网调度自动化系统,一方面,系统可能面临成千上万条时序数据作为输入,这些输入可能会同时发生异常,对于系统的防误机制带来很大的挑战,验证用例生成需要同时考虑这些输入;另一方面,人为操作、主备机切换、系统运行环境等多种外部因素也会对系统性能产生影响。多种因素的耦合作用下,导致状态空间维数爆炸,失效/边界验证用例的搜索难度和数目也随着状态空间维数的增加指数级增加。

### 2) 故障模式隐蔽:

电网调度自动化系统鲁棒性测试的重点在于发现设计时控制逻辑以及防误措施忽略的故障模式,而“发现未知故障”的目的为电网调度自动化系统的鲁棒性测试进一步带来了挑战。首先,面对“发现未知故障”的目

的,代码走查、性能测试等传统软件测试方法失效,因为传统软件测试方法只能发现已完成设计和编写部分的错误和漏洞;其次,由于故障模式未知且隐蔽,缺乏专家经验对于验证用例搜索过程的指引,先验信息的缺乏进一步加剧了验证用例生成的难度;最后,由于故障模式未知,对于每一个验证用例,需要人工判断电网调度自动化系统的输出是否正确,既额外增加了测试成本,也给自动化测试带来了困难。

### 3) 异常/极端用例占比小但风险高:

对于一个设计完备的电网调度自动化系统,状态空间中的异常/极端区域必然较小,但一旦出现,也会对电网安全造成严重影响。鲁棒性测试必须在这些异常、故障、极端区域生成多剖面组合验证用例,超高维状态空间下的微小异常/极端区域使得生成失效/边界验证用例更加困难。

### 4) 高测试效率需求:

为了加快测试进程,减少测试时间,控制测试成本,使得电网调度自动化系统尽快投入使用,需要测试方法具有较高的测试效率。一方面,虽然面临的状态空间维数巨大,但鲁棒性测试所需验证用例应尽可能少,以使用最少的验证用例覆盖系统的异常/极端情况;另一方面,生成验证用例需要尽可能快。需要充分利用电网调度自动化系统各级仿真系统,减少访问真实电网调度自动化系统的次数,减少测试成本,加快测试效率。

## 2 电网调度自动化系统的测试验证方法

《电网调度自动化系统软件通用测试规范》(下文简称《规范》)中将对于电网调度自动化系统的测试方法分为静态测试与动态测试两种,其中静态测试用于对于文档和代码的检查与测试,动态测试则用于其他测试类别,采用黑盒测试和白盒测试方法进行<sup>[1]</sup>。《规范》中针对每一个测试类别分别给出了适用的测试方法,其中,功能测试、性能测试、标准符合测试、兼容性测试、安全性测试以及稳定性测试都采用黑盒测试方法。但《规范》中表示:应选择适用的测试工具进行测试,或由人工进行手工测试,并未给出具体的测试步骤以及验证用例的设计方法。

目前已有部分学者对于电网调度自动化系统展开测试验证。可以按照不同测试方法、不同验证用例生成手段对于现存的电网调度自动化系统的测试方法进行分类。

按照不同测试方法,可以将目前面向电网调度自动化系统的测试方法分为白盒测试方法、灰盒测试方法以及黑盒测试方法。白盒测试方法需要知道系统内部逻辑或代码,一般针对软件代码进行测试与验证;黑盒测试方法只需获得系统对于输入的响应,一般针对系统性能测试、功能测试等其他测试;灰盒测试方法是两者的结

合。对于白盒测试方法，杨清波等人提出面向智能电网控制系统的试验验证体系及平台，其中包括源代码验证步骤，并通过该步骤发现 4 106 处代码缺陷或隐患<sup>[10-11]</sup>。对于黑盒测试方法，大量学者提出使用验证用例生成的方法对于电网调度自动化系统进行测试，但都未给出具体的验证用例生成方法。例如，杨清波等人在所提出的测试平台中设计了调度自动化主站系统控制类软件的闭环测试方法和雪崩测试方法，但也并未涉及验证用例的具体设计方法<sup>[10-11]</sup>。郑光明等人提出对于基于 iES500 Extreme 的调度自动化软件进行可靠性测试，分为基本功能测试和高级应用部分的测试，均采用验证用例生成的方法<sup>[6]</sup>。夏文岳等人提出基于多维场景组合验证的电网自动控制系统测试方法<sup>[2]</sup>，通过结合电网仿真模型，充分考虑电网运行方式、人为操作、环节状态等因素，设计软件运行各类正常场景、异常场景、故障场景和极端场景等多维场景，对于系统进行性能、功能、可靠性等方面的测试。但是他们重点在于介绍测试流程，验证用例的生成规则只提到了在数据上添加随机扰动以及依据特定条件生成特定扰动两种，生成的验证用例对于状态空间覆盖性较低，无法发现所有极端/异常情况。李宇佳等人提出一种调度自动化主站系统软件测试方法及其平台<sup>[12]</sup>，为调度自动化软件提供包括基于网页的自动化测试、基于虚拟技术的环境搭建测试和基于 WEB 访问的测试 3 种测试模式，实现对其进行静态测试及功能、性能及稳定性、标准符合性、兼容性等动态测试类别的测试。但他们的重点在于论述平台实现流程，对于具体如何设计验证用例以完成稳定性测试没有具体的介绍。

按照不同验证用例产生方法，可以将目前的电网调度自动化系统测试方法分为人工生成验证用例以及自动生成验证用例两种。人工生成验证用例要求测试人员清楚明白测试目的，且适合于只需要生成有限数目验证用例的测试。例如，应伟刚等人在 AGC 入网前对其进行了测试，他们将 AGC 的测试分为了四个方面：软件安全测试、硬件安全测试、指令安全测试以及整组测试，并对不同测试类别，详细给出了验证用例设计方法<sup>[5]</sup>。杨清波等人提出的针对智能电网控制系统的试验验证平台中针对电网控制类软件闭环验证采用人工设计验证用例的方法对于软件功能进行验证<sup>[10]</sup>。由于人工设计验证用例的测试效率较低，且验证用例覆盖性不够，部分学者提出自动验证用例生成方法。刘卫红等人提出了一种面向智能电网调度系统中参数配置测试的具有覆盖性的验证用例生成方法<sup>[13]</sup>，通过求解解空间树，自动生成满足两两组合覆盖标准的验证用例集合。在生成的验证用例集合中，任意两个参数的任意两种取值组合至少在一个验证用例中出现过，由此达到了对于两两组合参数配置的测试。但此方法只能探讨两个参数组合作用对于系统造成的影响，并且只适用于具有离散取值的变量的测试。陈明恩等人针对配电主站的集成测试进行了方案设计，提出了自动化测试过程，包括测试用例生成器、测试用例生成文档、测试用例解释器、测试行为发生器、终端数据模拟发生器和测试日志分析器<sup>[14]</sup>。但他们的的方法偏向于软件正确性测试，无法验证系统在异常/极端输入下的正确性。对于目前针对电网调度自动化系统的测试方法进一步按照测试目标、方法适用性和优缺点进行总结如表 1 所示。

表 1 面向电网调度自动化系统测试方法总结

方法分类	测试目标	方法适用性	方法代表	优点	缺点
源代码验证	发现代码层面的缺陷或安全隐患，确保软件内部逻辑正确。	适合于代码级的检验，保证基本处理逻辑正确。	[10-11]	能够尽早发现代码错误，修复成本低。	需要访问源代码，无法发现设计逻辑以外的未知错误，无法对智能系统进行测试。
功能测试 (验证用例生成)	验证系统在特定场景、输入、参数取值下的功能、性能、可靠性和安全性。	适合于考虑运行状态下的系统集成测试。	[5]	测试目标明确，用例意图清晰。	人工设计验证用例，覆盖性低，成本高。
			[2],[6],[14]	场景覆盖面广，能综合考虑运行方式、人为操作、环节状态等多种因素。	对于状态空间覆盖性低，场景生成依赖随机扰动，难以发现极端异常情况。
			[13]	对于两两参数组合实现覆盖性测试。	无法适用于更多参数组合共同作用产生的缺陷，仅适用于离散参数变量。
自动化测试平台	通过平台化工具(如网页自动化等)实现自动化测试，覆盖功能、性能、稳定性、兼容性等测试。	适用于需要高效回归测试、标准化测试流程的项目。	[12]	提升测试效率：自动化执行，支持多种测试模式，便于回归测试。测试类别全面：可集成静态测试、动态测试等多种测试类型。	对于极端用例的覆盖性不足。
雪崩测试	测试系统在极端高负载或海量数据冲击下的稳定性和处理能力。	适用于评估调度自动化系统等关键系统在面临数据挤压等情况时的抗压能力和崩溃临界点。	[10-11]	有效验证系统在极端情况下的性能表现和可靠性。	现存文章没有具体给出设计有效的雪崩验证用例的方法

综上所述,基于验证用例生成的黑盒测试是对于电网调度自动化系统进行测试验证的有效手段,目前研究趋势也正向自动化测试、结合仿真平台的测试方向发展。但目前研究中缺乏对于验证用例生成方法的具体说明,部分文献中采用的人工设计验证用例、生成两两组合覆盖性验证用例的方法生成的验证用例对于状态空间覆盖性不够,不适用于需要完备探索高维耦合状态空间中所有异常区域的鲁棒性测试。面向电网调度自动化系统鲁棒性的验证用例生成方法还亟需进一步研究。

### 3 面向系统鲁棒性测试的验证用例生成框架

#### 3.1 鲁棒性的定义

如引言章节所述,IEEE组织对于一般软件系统鲁棒性的定义为:在压力条件以及扰动下系统保持正确的能力。在宏观尺度上,ISO/IEC组织给出系统鲁棒性定义为:系统在任何情况下都保持其性能水平的特性<sup>[15]</sup>。因此,面对系统鲁棒性测试的需求,应重点探讨系统是否在任何异常、极端以及压力输入下,都可以保持正常功能,因此需要重点生成使得系统出现输出错误或功能异常的输入和参数条件作为验证用例。

具体而言,一般使用系统输出保持不变(或保持正确时),输入能承受的最大扰动作为系统鲁棒性的度量指标,该指标又被称为系统的鲁棒半径。

#### 3.2 面向系统鲁棒性测试的验证用例生成问题建模

如3.1节所述,面对电网调度自动化系统鲁棒性测试的需求,需要重点探索使得系统出现输出错误的输入和参数条件。为了能够完备验证系统对于所有异常、极端输入的处理能力,对于系统进行摸底探底,需要明确所有可能使得系统性能失效的状态空间区域。但对于一个高维状态空间,失效区域的分布可能是高度非线性、不规则以及离散的,生成对这些区域具有覆盖性的验证用例非常困难,可以通过生成系统性能良好与性能失效交界区域的验证用例(后文统称为边界验证用例),用尽可能少的验证用例,勾勒系统所有失效区域的分布。

本文将被测电网调度自动化系统视为一个黑盒系统;所有影响系统输出的因素构成状态空间;将系统输出划分为不同性能模式,并使用离散整数代表系统性能模式,例如使用1代表系统性能良好(输出正确),使用0代表系统性能失效(输出错误),性能模式构成性能空间;将状态空间中系统性能良好和性能失效之间的交界区域定义为性能边界区域;验证用例的搜索与生成将在状态空间中进行,其中,边界验证用例位于性能边界区域中,对于这些边界验证用例,状态空间中的微小改变将使得所对应的性能模式发生突变。通过生成对于系统性能边界具有覆盖性的验证用例,可以完备勾勒出

系统所有失效区域的分布。下面将给出相关定义的具体数学建模。

定义1:被测电网调度自动化系统

依据黑盒测试的思想,电网调度自动化系统可被视为从输入空间到输出空间的映射 $f$ :

$$f: \mathbf{X}^p \rightarrow \mathbf{Y}^q \quad (1)$$

其中: $\mathbf{X}^p \subset \mathbb{R}^p$ 为 $p$ 维输入向量,由所有对于系统输出会产生影响的因素构成,如输入变量、参数设置等,输入向量张成电网调度自动化系统的输入空间。 $\mathbf{Y}^q \subset \mathbb{R}^q$ 为 $q$ 维输出向量,构成电网调度自动化系统的输出空间。

更具体而言,对于电网调度自动化系统,输入向量 $\mathbf{X}^p$ 是多模态数据,包括厂站输入数据与模型参数变量。厂站输入数据为时序数据矩阵,假设厂站数目为 $N$ ,时间步长为 $T$ ,则厂站输入数据为 $N \times T$ 的矩阵;模型参数变量为单一变量,可能为连续参数,也可能为离散状态。输出向量 $\mathbf{Y}^q$ 为电网调度自动化系统输出的调节指令。

定义2:状态空间与性能空间

状态空间由电网调度自动化系统输入空间中的重要因素所构成,为输入空间的子集,记为 $\mathbf{A}^n$ , $\mathbf{A}^n \subset \mathbf{X}^p$ ,且 $n \leq p$ , $\mathbf{A}^n$ 中元素记为 $\mathbf{x}$ 。状态空间可以通过专家经验确定,也可以通过敏感性分析等方法确定。每一个验证用例对应状态空间中的一个向量。

输出空间可以依照系统的不同执行结果分为多个互不交叉的性能模式:

$$P: P \subset \mathbf{Y}^q, \cup_i P_i = \mathbf{Y}^q, P_i \cap P_j = \phi \\ \forall \mathbf{x} \in \mathbf{A}^n, f(\mathbf{x}) \in P_i \quad (2)$$

性能空间由性能模式张成,记为 $P$ 。性能空间可以看作是状态空间对应的响应空间,即状态空间一个输入会对应一个唯一的性能模式。通常,性能模式取值为一系列离散值,表示系统的不同执行情况。例如,使用性能模式1和0分别代表系统性能良好与失效。具体而言,对于电网调度自动化系统,依据系统输出的调度指令来判断目前指令是否正确,如果正确,则性能模式 $P$ 为1,否则性能模式为0。

定义3:性能边界区域及性能边界验证用例

性能边界区域为状态空间中微小扰动导致输入所对应的性能模式发生变化的区域。将性能模式 $P_a$ 和 $P_b$ 间的性能边界区域记为 $B_{a,b}$ , $B_{a,b} \in \mathbf{A}^n$ , $B_{a,b}$ 定义为:

$$\forall \mathbf{x}_a \in B_{a,b}, \exists \mathbf{x}_b \in B_{a,b}, \\ f(\mathbf{x}_a) \in P_a, f(\mathbf{x}_b) \in P_b, \\ a \neq b, |\mathbf{x}_a - \mathbf{x}_b| < W_b \quad (3)$$

其中: $W_b$ 为预先定义的性能边界宽度。一般而言, $W_b$ 的取值可以与状态空间范围成正比,例如取值为 $0.05 \times$ 状态空间范围。

边界验证用例是位于性能边界区域中的验证用例，其主要特征是当对其施加一个大小为  $W_b$  的扰动后，相应性能模式值可能会发生改变。

性能边界验证用例记为  $x_p$ ， $x_p$  满足下列条件：

$$\begin{aligned} \forall x_p \in A^n, \exists \Delta x \in A^n, \\ x'_p = x_p + \Delta x, \\ \text{s.t. } f(x_p) \in P_a, f(x'_p) \in P_b, \\ a \neq b, |\Delta x| < W_b \end{aligned} \quad (4)$$

图 1 给出了性能边界区域以及性能边界验证用例的示意图。

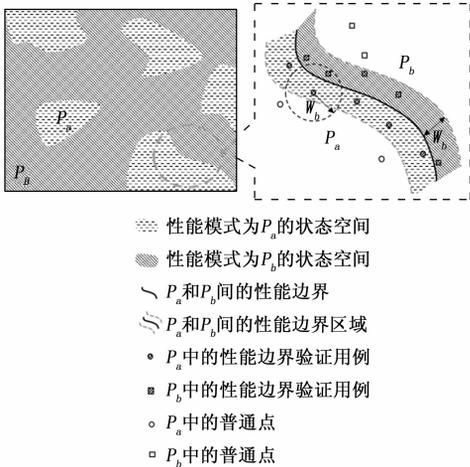


图 1 性能边界以及性能边界验证用例示意图

### 3.3 面向系统鲁棒性测试的验证用例生成算法框架

为了可以完备勾勒出被测系统状态空间中所有失效区域的分布，生成的验证用例需要覆盖所有的性能边界区域。面对状态空间维数灾难以及边界分布未知的难题，生成的边界验证用例需要具有：1) 覆盖性：验证用例需要对于状态空间中的所有性能边界区域实现覆盖，以对于系统鲁棒性进行准确评估。2) 准确性：所生成的验证用例应集中于性能边界区域中，减少远离边界的无效验证用例数目。3) 高效性：边界验证用例生成算法需要快速高效，以尽快生成大量边界验证用例，减少测试成本。

为了可以在高维状态空间中高效地定位分布未知、形状未知、离散程度未知、区域大小未知的所有潜在的性能边界区域，并在性能边界区域产生验证用例，算法需要具有很强的对于状态空间的探索能力以及对于边界区域的开发能力，同时，应尽量减少测试成本与测试时间。

为了实现使用尽可能少的验证用例覆盖尽可能多的不规则性能边界区域的目的，学术界通常采用代理模型辅助的自适应采样框架来生成边界验证用例<sup>[16-18]</sup>。基于代理模型辅助的自适应采样框架流程如图 2 所示。性能边界自适应探索过程由代理模型构建和自适应采样两个

步骤构成。一方面，自适应采样方法已经被证明可以有效在感兴趣的不规则区域生成采样点（状态空间的采样点就是验证用例）；另一方面，通过一系列 {验证用例, 性能模式} 构造的代理模型，可以替代被测系统在采样过程为采样点提供模拟性能模式，减少调用被测系统的同时，为采样过程提供梯度等信息，将黑盒被测系统白盒化以及可解释化，提高自适应采样的效率。自适应采样在状态空间中进行性能边界自适应搜索，产生验证用例，并提交被测系统获得验证用例对应性能模式，一系列 {验证用例, 性能模式} 对将用于下一次迭代中训练代理模型，以指导下一次迭代的自适应采样过程。在自适应采样与代理模型的交互迭代中，代理模型越来越准确，自适应采样生成的验证用例也越来越收敛于真实性能边界区域，迭代结束后，自适应采样步骤将输出获得的所有边界验证用例。

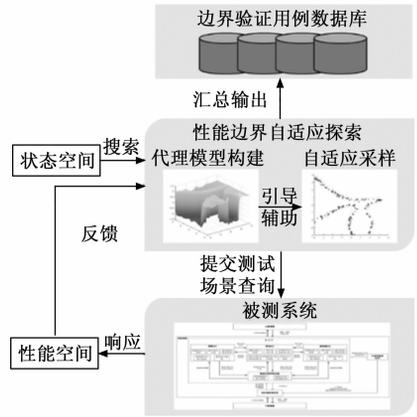


图 2 基于代理模型辅助的自适应采样框架图

具体而言，使用基于代理模型辅助的自适应采样框架生成电网调度自动化系统的边界验证用例时，需要采取以下 6 个步骤：初始采样、初始代理模型构建、新验证用例的生成、代理模型的更新、验证用例的合并以及停机准则的判断。首先，通过均匀采样、随机采样等方法获得初始验证用例并构建初始代理模型，为后续的迭代步骤提供空间的基本信息；接着，当未达到停机准则时，调用自适应采样方法生成新的验证用例，并调用被测系统获得新增验证用例的真实性能模式；然后，进行代理模型的更新以及验证用例的合并；最后，当达到停机准则后，退出迭代，输出算法产生的所有验证用例，即为最终生成的验证用例。算法步骤如图 3 所示。

在整个算法迭代过程中，如何生成新的验证用例以及如何更新代理模型是算法的核心组成部分。对于自适应采样步骤，一般流程为：生成候选验证用例；调用代理模型获得模拟性能模式；依据验证用例状态空间信息以及性能空间信息计算合适的采样指标；最后依据采样指标以某种方式保留精英验证用例，即位于边界区域的

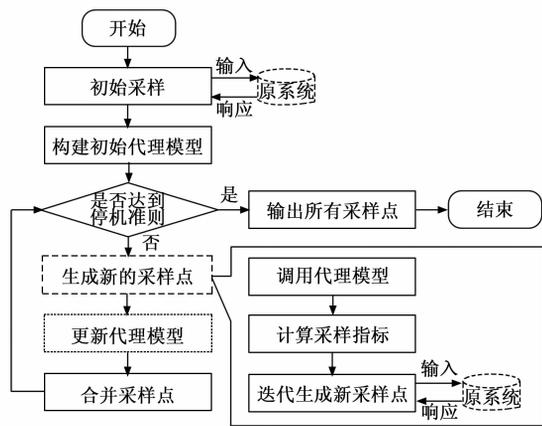


图 3 基于代理模型辅助的自适应采样框架流程图

验证用例。其中, 设计合适的采样指标以及生成新验证用例的方法是算法需要重点考虑的。采样指标通常由全局探索指标和局部开发指标构成, 以反应验证用例所处位置是不是感兴趣的区域为目标, 是采样效果好坏的核心。面对性能边界验证用例生成问题, 采样指标通常需要综合考虑验证用例附近性能模式的变化——以反应验证用例是否位于性能边界附近, 以及验证用例附近已产生的验证用例的密度——以反应该区域是否已经被探索过。新验证用例的生成过程依据采样指标决定从候选验证用例中选择哪些作为精英验证用例, 新验证用例生成方法决定了算法的收敛速度, 可以采用启发式搜索、接受—拒绝采样等一系列方法。

对于代理模型更新步骤, 代理模型基于统计学及数学推导, 构建 {验证用例, 性能模式} 之间的映射关系, 对于一个新的验证用例, 代理模型可以替代被测模型, 给出精确的性能模式值, 从而加快验证用例的生成过程, 减少昂贵的被测模型的调用次数。目前有多种机器学习模型都可以作为代理模型, 但存在各自的优缺点。例如, Kriging 模型与高斯过程回归模型 (GPR, gaussian process regression)<sup>[19]</sup> 精度较高, 且可以提供不确定度的估计, 被广泛用作代理模型, 但构建和调用的时间和空间复杂度较大, 不适用于高维问题。支持向量机模型 (SVM, support vector machine)<sup>[20]</sup> 更加适用于高维问题, 且可以提供分类面信息, 但对于异常值比较敏感。神经网络模型 (ANN, artificial neural network) 可以获得反向梯度, 更适用于高维大数据量问题, 但可解释性较差。在解决问题时, 应当依据数据特性, 选择合适的代理模型, 在模型构建精度与构建效率之间取得平衡。

### 3.4 鲁棒性评估指标

边界验证用例生成完成后, 需要对于电网调度自动化系统给出量化的鲁棒性评估。依据 3.1 节鲁棒性的定义, 本文结合鲁棒半径的概念与性能边界验证用例, 提

出平均边界距离作为系统的鲁棒性度量指标。

鲁棒半径指系统输出保持不变时, 输入所能承受的最大扰动。对于某个状态空间的输入  $x$  而言, 在对其施加扰动时, 系统输出保持不变时输入承受的最大扰动强度, 就是  $x$  到系统性能边界的最近距离。性能边界验证用例可以作为系统性能边界的表征, 因此, 本文使用输入  $x$  到距离其最近的性能边界验证用例的距离作为  $x$  到性能边界最近距离的近似。将性能边界验证用例集合记为  $S$ , 输入  $x$  处的鲁棒半径可以定义为:

$$\forall x \in A^n, R_x = \min_{x_p \in S} \|x - x_p\|_p \quad (5)$$

对于被测电网调度自动化系统而言, 使用随机采样方法获得正确执行的一系列输入样本  $T$ , 使用这些样本鲁棒半径的均值作为系统鲁棒性的度量指标:

$$R_{MUT} = \frac{1}{N} \sum_{i=1}^N R_{x_i}, x_i \in T, T \subset A^n$$

## 4 实验与分析

本章对于第三章中所提到的各代理模型的维度适用性展开分析, 对比高斯过程分类模型 (GPC, gaussian process classification)、支持向量机模型 (SVM) 以及神经网络模型 (ANN) 在不同维度的分类问题上的拟合精度与速度。

### 4.1 数据集

由于面向系统鲁棒性测试的验证用例生成算法框架中, 代理模型的功能是拟合状态空间连续变量与离散性能模式之间的关系。因此选择构造具有离散分类标签的数据集来对各代理模型进行测试。数据集中输入到输出标签的构造方法如图 4 所示。样本每个维度取值范围均为  $[-1, 1]$ , 若样本位于任何一个圆 (超球体) 内, 则标签为 1, 否则标签为 0。该构造方法具有很强的灵活性, 可以拓展到任意维度。

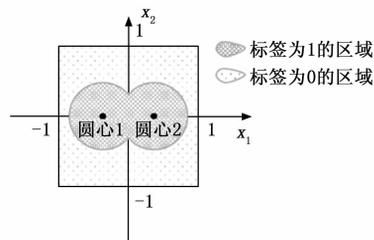


图 4 数据集构造方法

依据此方法分别构造 2 维、4 维、8 维、15 维和 20 维状态空间, 训练数据集与测试数据集均依照均匀分布在状态空间中随机产生, 两者互不交叠。在测试代理模型时, 使用训练集拟合模型, 然后使用测试集对模型精度进行测试。各维度下训练数据集与测试数据集信息如表 2 所示。表中“正类: 负类”指状态空间中标签为正类的区域与标签为负类的区域的大概比值。

表 2 代理模型拟合精度与速度比较

状态空间维度	训练样本数目	测试样本数目	正类 : 负类
2	200	100	7 : 13
4	800	400	2 : 3
8	3 200	1 600	1 : 1
15	20 000	5 000	9 : 14
20	50 000	10 000	9 : 13

### 4.2 代理模型配置

GPC 模型缩放系数取值为 1, 为了提高非线性拟合能力, 选择径向基函数为核函数 (RBF, radius based function), 核参数取值为 1。超参数保持固定, 不进行优化。

SVM 模型惩罚系数取值为 1, 同样选择 RBF 函数作为核函数。

ANN 模型有 2 层隐藏层。第一层隐藏层有 100 个神经元, 第二层隐藏层有 50 个神经元。使用 Relu 激活函数。

### 4.3 评价指标

使用精度指标与时间指标分别度量代理模型的拟合精度和拟合速度。

精度指标为代理模型在测试集上的分类准确率, 记为 acc, 假设测试集为  $S_{Test}$ , 则 acc 的计算方式如式 (7) 所示。其中  $y_i^{pred}$  指代理模型对于测试样本的预测标签,  $y_i^{true}$  指测试样本的真实标签。分类准确率为代理模型在测试集上标签预测正确的样本的比例:

$$acc = \frac{|\{ \forall x_i \in S_{Test}, y_i^{pred} = y_i^{true} \}|}{|S_{Test}|} \quad (7)$$

其中:  $|\cdot|$  代表集合中元素数目。精度指标范围为  $[0, 1]$ , 越接近 1, 说明代理模型拟合精度越高。

时间指标记为  $Time_{train}$  和  $Time_{test}$ , 分别代表各代理模型的训练时间和推理时间, 单位为秒 (s)。时间指标范围为  $[0, \infty)$ , 时间指标越小, 说明代理模型拟合速度越快。

### 4.4 实验结果

在不同维度数据集下, GPC、SVM 和 ANN 的拟合精度与速度对比如表 3 所示。

### 4.5 对比分析

从问题维度的角度分析, 对于 2 维和 4 维的低维问题而言, GPC 模型展示出较大的优势。此时, GPC 模型的拟合速度可以接受, 精度较高。而 SVM 模型与 ANN 模型虽然拟合速度更快, 但是精度不够。尤其是对于 ANN 模型, 在 2 维 200 个训练点下, 展现出较低的精度。随着状态空间维度的升高, GPC 的训练时间指数增加, SVM 和 ANN 的训练时间虽然也在增加, 但较为缓慢。并且由于训练样本数目增多, ANN 和 SVM 模型精度有了较大的提升。当问题维度达到 15 维以上时, GPC 已无法在可以接受的时间内完成拟合,

表 3 代理模型拟合精度与速度比较

状态空间维度	模型	acc	$Time_{train}$ (s)	$Time_{test}$ (s)
2	GPC	<b>0.980 0</b>	0.257 0	0.000 0
	SVM	0.870 0	<b>0.002 0</b>	<b>0.000 0</b>
	ANN	0.600 0	0.008 1	0.000 0
4	GPC	<b>0.960 0</b>	3.251 9	0.003 0
	SVM	<b>0.960 0</b>	<b>0.020 1</b>	0.003 5
	ANN	0.955 0	0.092 9	<b>0.000 0</b>
8	GPC	0.940 0	111.798 2	0.049 1
	SVM	0.930 6	0.376 2	0.058 2
	ANN	<b>0.957 5</b>	<b>0.357 4</b>	<b>0.001 0</b>
15	GPC	—	—	—
	SVM	<b>0.945 8</b>	15.174 9	0.983 0
	ANN	0.938 6	<b>1.066 2</b>	<b>0.001 4</b>
20	GPC	—	—	—
	SVM	<b>0.948 2</b>	94.073 4	4.427 4
	ANN	0.928 6	<b>4.247 9</b>	<b>0.004 1</b>

表中“—”表示在该情况下, 模型的训练时间超过 15 分钟, 未再进行比较。

SVM 的拟合精度虽然可以得到保证, 但拟合时间也进一步增加。此时 ANN 模型的优势进一步凸显, 可以在很快的训练时间和几乎不变的推理时间内, 完成对于数据的拟合, 并且保持可以接受的精度。

从不同代理模型适用性的角度分析, 可以得到如下结论: GPC 适合于低维问题, 可以保证最优的精度以及可以接受的训练时间, 但其复杂度随着维度上升呈指数变化, 无法胜任高维数据的拟合。SVM 的复杂度与支持向量个数有关, 相较于 GPC 更适合于高维问题, 在中高维度下, SVM 具有较好的精度和较快的拟合时间。但若问题维度高于 20 维, SVM 的拟合时间也将超出 2 分钟。ANN 适合于高维问题, 训练与推理速度快, 具有很强的非线性拟合能力, 但需要足够的训练样本以保证其精度。

## 5 研究挑战与未来应用

第三章所述面向系统鲁棒性测试的边界验证用例生成方法框架的有效性已经在其他复杂系统 (如无人机路径规划系统、船舶自动避碰系统等) 的边界验证用例生成问题中得到了验证, 但将该框架应用于为电网调度自动化系统生成边界验证用例时依然存在一些挑战。结合第一章所介绍的电网调度自动化系统的特点与测试难点, 本文将关键挑战与可能的解决方法总结如下。

### 1) 状态空间维度过高及时序特性:

对于电网调度自动化系统而言, 对于输出会产生影响的因素可能有成千上万个, 即状态空间的维度可能达到上万维。在一个上万维的状态空间中搜索边界验证用例是不切实际的。一方面, 空间维数的增长造成覆盖性能边界区域所需的验证用例数目急剧增长, 例如, 每个

维度如果仅使用 2 个验证用例覆盖, 10 000 维的状态空间需要  $2^{10\,000}$  个验证用例进行覆盖, 这是一个超过  $10^{3\,010}$  的天文数字; 另一方面, 在如此高维的状态空间中, 常见欧式距离、汉明距离等距离度量方式都将失效, 无法准确衡量验证用例距离边界的远近。

因此, 必须采用一系列降维方法降低状态空间的维度。首先, 应当借助专家经验, 分析电网调度自动化系统所面临的输入数据特征, 分析变量之间的耦合关系, 将对系统具有耦合作用的变量提取为一组, 分组进行测试, 减少状态空间维度; 然后, 可以借助一系列特征选择方法, 进一步选择出对于系统性能具有更大影响的因素构成状态空间。一系列降维方法后, 状态空间可能还是高维的, 此时, 可以选择神经网络等更适合于高维空间的拟合的方法构建代理模型, 利用对抗样本生成等方法寻找边界验证用例, 使得问题可求解。

#### 2) 故障模式隐蔽与异常区域占比小:

为了能够对于电网调度自动化系统的鲁棒性进行全面的评估, 边界验证用例应当覆盖所有性能失效区域的边界。但正是由于故障模式隐蔽, 加之搜索空间巨大, 带来了性能边界搜索困难与测试预言问题。

针对搜索空间巨大、失效区域较小、故障模式隐蔽的问题, 可以从以下两个方面尝试解决。首先, 可以借助专家经验, 进行空间尺寸的缩减, 从已知的、简单的故障模式出发进行搜索, 逐步扩大搜索区域与搜索维度。其次, 可以借助生成式模型(如变分自编码器与对抗生成式网络), 基于少量边界验证用例, 生成更多样化的边界验证用例。

针对测试预言问题, 即需要人工判断系统对于验证用例的处理是否正确而带来的巨大测试成本和无法实现自动化测试的问题, 可以采取蜕变测试的思想, 不直接将系统输出是否正确作为性能模式, 而将系统输出是否发生改变作为性能模式的判定标准, 一旦系统的输出发生了改变, 性能模式一定发生了从性能良好到性能失效或性能失效到性能良好的转变。

#### 3) 高测试效率需求:

考虑到状态空间维度很高, 覆盖系统性能边界所需的验证用例数目激增, 极大增加了测试成本与测试时间。一方面, 生成验证用例时, 需要获得验证用例的性能模式, 需要反复调用电网调度自动化系统, 带来很大的测试成本; 另一方面, 执行大量的验证用例同样需要消耗时间。

可以从两个角度尝试解决该问题。首先, 需要尽量减少验证用例的数目。可以考虑结合黑盒测试与白盒测试的优点, 将电网调度自动化系统的各级功能进行拆分, 将大的黑盒系统拆分为各子黑盒系统, 对于每个子功能分别生成验证用例, 缓解状态空间各维度耦合带来

的验证用例数目指数增长的问题。其次, 需要加快验证用例的生成速度。除了使用代理模型技术外, 可以充分利用电网调度自动化系统的各级仿真模型, 减少调用真实电网调度自动化系统的次数。

#### 4) 验证用例覆盖性评价困难:

由于电网调度自动化系统的真实性能边界以及性能失效区域分布未知, 存在难以验证算法生成的边界验证用例对于边界的覆盖性、难以验证所生成的验证用例是否能够实现对于电网调度自动化系统的完备测试、是否没有遗漏性能失效区域的问题。

针对这一问题, 可以借助蒙特卡洛的思想, 依据专家经验, 生成极端或异常用例, 通过计算边界验证用例对于随机生成的极端用例的覆盖性, 作为算法生成的验证用例覆盖性的评判标准。

## 6 结束语

本文重点对于面向电网调度自动化系统鲁棒性的验证用例生成方法进行了研究。首先, 论述了电网调度自动化系统鲁棒性测试的必要性; 其次, 对于电网调度自动化系统的特点和测试难点进行了分析; 然后, 按照白、灰、黑 3 种测试方法以及人工/自动验证用例生成方法对于现有电网调度自动化系统的测试验证方法现状进行了总结, 现有方法集中于代码测试以及功能测试, 缺乏对于验证用例生成方法的具体说明, 无法适用于电网调度自动化系统的鲁棒性测试; 接着, 介绍了面向系统鲁棒性测试的验证用例生成框架, 分别给出了鲁棒性的定义、面向系统鲁棒性测试的验证用例生成问题建模、面向系统鲁棒性测试的验证用例生成算法框架以及鲁棒性评估指标; 针对框架中涉及到的各代理模型的维度适用性进行了实验分析, 对比了 GPC、SVM 及 ANN 在不同状态空间维度下的拟合精度与速度; 最后, 结合电网调度自动化系统的测试难点以及验证用例生成算法框架, 具体分析了使用该框架为电网调度自动化系统生成验证用例时所面临的难点以及可能采取的解决方法与研究方向。

#### 参考文献:

- [1] Q/GDW 11205—2014, 电网调度自动化系统软件通用测试规范 [S]. 北京: 中国电力出版社, 2014.
- [2] 夏文岳, 李宇佳, 卫泽晨, 等. 基于多维场景组合验证的电网自动控制系统测试方法及相关系统: CN119575944A [P]. 中国: 2025-03-07.
- [3] 李宇佳, 杨清波, 李立新, 等. 电网调度自动化系统软件通用测试规范解读 [J]. 智能电网, 2015, 3 (2): 186-190.
- [4] 谢善益, 徐展强, 邓大为, 等. 电网调度自动化系统(主站系统)专业测试方法 [J]. 广东电力, 2005, (2): 13

