

基于大数据分析的潜射武器健康状态评估

郑帮涛

(海军装备部, 北京 100086)

摘要: 针对潜射武器健康状态评估中的大数据分析问题, 对大数据概念以及特点进行了论述, 重点阐述了神经网络、群体智能和相关分析三种大数据分析的主要思想、应用范围, 提出了潜射武器健康状态评估的潜射武器大数据存储与处理、数据特征提取、健康状态评估技术路线, 最后指出了大数据分析在潜射武器健康状态评估中应用的可行性。

关键词: 大数据分析; 潜射武器; 健康状态评估; 深度神经网络

Health State Evaluation of Submarine-launched Weapons Based on Big Data Analysis

Zheng Bangtao

(Naval Department of Equipment, Beijing 100036, China)

Abstract: As to health state evaluation for submarine-launched weapon based on big data analytics, this paper discusses concept and characteristic of big data, especially the main idea and application scope of three big data analysis methods, including neural network, swarm intelligence and correlation analysis. Technical route of combining health state evaluation for submarine-launched weapons health state evaluation and big data analysis technology is proposed, including big data storage and process, big data feature extraction and health state evaluation. Finally, application feasibility of big data analysis in health state evaluation for submarine-launched weapon is analyzed.

Keywords: big data analysis; submarine-launched weapon; health state evaluation; deep neural network

0 引言

随着潜射武器的快速发展, 部队实战化训练越来越频繁, 暴露出越来越多的可靠性问题。装备一旦出现故障, 一般采取基层级维修和基地级维修两种方式, 基层级维修范围比较窄, 基地级维修时间长、距离远, 这两种维修保障方式已经不能完全满足现代战争对潜射武器的快速反应需求。

装备主动维修或基于状态的维修成为解决这一问题的有效途径, 而装备健康管理是装备主动维修的关键技术。潜射武器系统是一类复杂综合系统, 包括电子设备、机械设备、光学设备等, 要对武器系统进行健康管理就需要大量数据做支撑。

随着潜射武器作战、试验的频率不断加大, 装备数据量的爆炸式增长, 如何分析大数据, 充分挖掘大数据的潜在价值实现装备的健康管理, 为装备管理部门和部队服务, 成为需要深入探讨的现实问题。本文将探讨大数据分析技术应用于潜射武器健康管理的关键问题, 为后续潜射武器综合保障研究提供支持。

1 大数据概念及特点

大数据指在可容忍的时间内用传统信息技术和硬件工具难以对其进行获取、管理、处理和分析的数据集合^[1]。大数据一般指数据量达到 PB 级甚至 EB 级的大规模数据。大数据具有以下 4 大特征:

1) 体量庞大, 指大数据所包含的数据量庞大, 这是判断

一个问题是否属于大数据问题的最基本特征。随着潜射武器试验、作战训练、装备保障的频率加快, 数据量急剧增加, 虽然目前数据量尚未达到体量浩大的规模, 但是可以预见在不久的将来, 潜射武器的数据量将会达到大数据要求。

2) 来源复杂, 指大数据所包含的数据来源广泛、形式多样。例如, 潜射武器数据来自试验、训练、作战, 数据包括: 作战方案、维修数据、作战使用数据等多模态非结构化数据, 也包含如电压测试数据、温湿度等结构化数据。

3) 增长速度快。例如, 海洋作战环境等大规模系统每时每刻在产生大规模数据。

4) 价值稀疏。指大数据中的有用数据量少。例如, 对于装备测试数据, 真正有价值的是一些故障数据或奇异点数据, 但是这些数据在装备正常状态下数据量极少。然而, 这些价值稀疏的数据正是大数据研究的目标与意义。

从上述大数据特点可以发现, 潜射武器当前的数据规模暂时还达不到“体量浩大”, 但是, 来源复杂、增长速度快、价值稀疏的特点基本满足。随着后续潜射武器综合保障频次的不断加深, 尤其是作战环境数据的迅速膨胀, 潜射武器的数据规模会越来越大。

2 大数据分析技术

为了充分挖掘大数据的潜在价值, 不同应用领域专家学者对深度学习、群体智能、相关分析、粒计算、深度森林等进行了深入研究。本文将对典型大数据分析技术进行分析, 论述各类分析技术的应用领域, 为大数据分析技术在潜射武器健康状态评估中的应用提供基础。

收稿日期: 2017-11-10; 修回日期: 2017-12-20。

作者简介: 郑帮涛(1969-), 男, 湖北襄阳人, 高级工程师, 主要从事导弹测试与控制方向的研究。

2.1 深度神经网络

随着计算机技术的快速发展，深度神经网络在大数据分析中的应用领域不断拓宽。近年来，出现了众深度自动编码器、深度卷积神经网络等多种神经网络模型，这些模型在图像分类、手写字体识别等领域得到了广泛应用。本文简要介绍各类神经网络的基本原理及应用范围，探讨在潜射武器健康状态中应用的可行性。

2.1.1 深度卷积神经网络

深度卷积神经网络是在传统卷积神经网络的基础上，结合深度学习思想而提出的一类深度学习神经网络模型。与传统卷积神经网络相比，深度卷积神经网络的卷积层、池化层数量有所增加，相应的各层神经元数量也会增加，结构如图 1 所示^[2]。

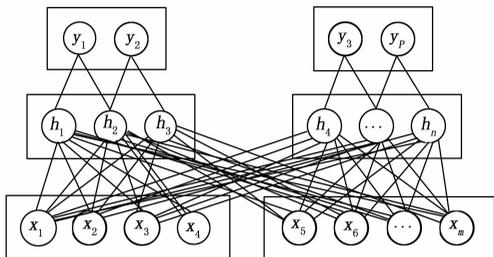


图 1 卷积神经网络拓扑结构图

由于深度卷积神经网络可以实现权值共享，大大减少了调节参数数量，训练速度也会大大提高。目前，深度卷积神经网络模型主要有并行深度卷积神经网络^[3]、基于改进 Fisher 准则的深度卷积神经网络^[4]等。深度卷积神经网络应用领域主要有：图像分类、人脸识别、色织物缺陷检测、船舶识别。

潜射武器健康状态评估本质上也是分类问题，从深度卷积神经网络的机理和应用领域可以看出，深度卷积神经网络比较适合用于解决图像识别问题，可以考虑将深度卷积神经网络用于潜射武器图像识别，然后根据识别的结果评估健康状态。

2.1.2 深度自动编码器

传统自动编码器是一种三层神经网络模型，包括输入层、隐含层和输出层。深度自动编码器在传统自动编码器隐含层上增加了深度，形成了一个由若干曾隐含层、输入层和输出层组成的深度自动编码器，结构如图 2 所示。

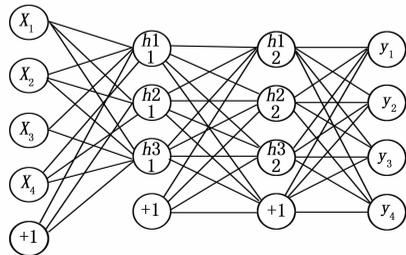


图 2 深度自动编码器网络拓扑结构图

深度自动编码器的训练一般分为预训练和微调两个步骤。预训练主要是根据应用领域，将神经元连接权值和其他参数限定在一定范围内，为后续的微调减少训练时间。微调通常使用 BP 算法等梯度下降算法对整个深度自动编码器的参数进行调整，以获得最优的训练结果。目前，深度自动编码器在特征提

取、故障诊断^[5]、图像识别等领域有广泛应用。故障诊断与特征提取是潜射武器健康状态评估的重要内容，深度自动编码器可用于复杂装备的特征提取和故障诊断。

2.1.3 无限深度神经网络

一般的深度神经网络主要用于处理静态数据，实现模式识别、分类和预测。但是，对于动态数据，尤其是“bus driver”问题却显得无能为力。扩展了时间维度的无限深度神经网络却可以轻松解决动态数据的特征提取和建模。

无限深度神经网络 (infinite deep neural networks, infinite DNNs) 是在传统反馈式神经网络的基础上，扩展了时间维度，其结构如图 3 所示^[6]。无限深度神经网络主要应用领域包括机器翻译、自然语言理解、动态数据特征提取等。

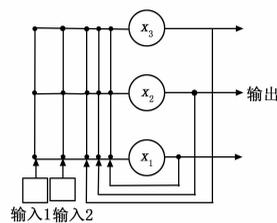


图 3 无限深度神经网络拓扑结构图

潜射武器系统在测试、发射过程中会产生大量实时时间序列动态数据，无限深度神经网络可根据装备动态数据，实时评估装备的健康状态。

2.2 群体智能算法

群体智能算法包括：遗传算法 (genetic algorithm, GA)、粒子群优化 (particle swarm optimization, PSO)、人工免疫系统 (artificial immune system, AIS) 等，这类智能算法可以近似求解一些 NP 问题，还可用于约简数据规模^[7]。

2.2.1 遗传算法

遗传算法的基本原理在众多文献均有描述，本文不再赘述。遗传算法在大数据分析方面的应用有两种方式：1) 单独采用遗传算法实现大数据处理，如数据属性约简^[7]；2) 遗传算法与其他算法结合，生成优化的分类算法或拟合算法，如神经网络与遗传算法结合实现数据挖掘^[8]等。

2.2.2 粒子群优化算法

粒子群优化算法采用速度更新算子实现粒子的位置更新。与遗传算法相比，粒子群优化算法中个体具有记忆功能，没有选择、交叉、变异算子，算法更加简单、收敛速度更快。粒子群优化算法在大数据中的应用方式与遗传算法相似，主要是和其他模式识别算法相结合，实现大数据的处理，主要应用领域包括数据分类或聚类^[9]，数据离散化^[10]等。

2.2.3 人工免疫系统

人工免疫系统是对免疫系统原理的模拟算法，包括克隆选择算法、阴性选择算法、免疫进化算法等。以克隆选择算法为例，采用克隆、变异算子实现群体优化，克隆选择算法在保持种群多样性方面具有优势。人工免疫系统在大数据分析中的应用主要以数据分类^[11]、分级聚类^[12]等。

增强决策力和流程优化能力是大数据分析的主要目的之一，而装备健康状态评估问题本质上是最优化问题。在大数据环境中，优化问题不可避免地涉及到更多的决策变量和优化目

标, 形成更为复杂的多目标优化问题。与传统方法相比, 这类算法不需要目标函数的梯度信息, 不受目标函数的形式和性质的限制, 可以用于优化任意形式的目标函数, 优化过程独立于具体的应用领域, 不易受到具体领域的约束, 因此, 其具有比传统优化方法更广泛的应用范围。

2.3 大数据相关分析技术

相关分析在传统数据分析中的应用范围非常广泛。随着大数据时代的到来, 相关分析技术如何解决大数据分析问题, 以及大数据的相关分析应用成为大数据分析领域关注的热点问题之一^[13]。目前, 常见的大数据相关分析分为两类: 一类是基于互信息和距离的变量间的复杂相关关系分析; 另一类是基于矩阵计算的属性相关分析。

2.3.1 基于互信息的相关分析

基于互信息的相关分析是以信息熵理论为基础, 度量信息之间的影响程度关系。

设随机向量 (X, Y) 的联合概率分布为 p_{ij} , 则 (X, Y) 的二维联合熵为:

$$H(X, Y) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log(p_{ij}) \quad (1)$$

假定 X 和 Y 的边际分布分别为 $p_{i \cdot}$ 和 $p_{\cdot j}$, 可定义在已知 Y 的条件下 X 的条件熵:

$$H(X/Y) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log\left(\frac{p_{ij}}{p_{\cdot j}}\right) \quad (2)$$

同理, 可得在已知 X 的条件下 Y 的条件熵:

$$H(Y/X) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log\left(\frac{p_{ij}}{p_{i \cdot}}\right) \quad (3)$$

一般而言, $H(X) - H(X, Y)$ 表示已知 $Y(X)$ 的情况下 $X(Y)$ 信息量的变化程度。显然, 若变化程度较小则表明 $Y(X)$ 对 $X(Y)$ 的影响较小, 也就是说, X 和 Y 相关性弱。反之, 说明 X 和 Y 相关性强。这个差值称为互信息, 记为:

$$I(X, Y) = H(X) - H(X, Y) \quad (4)$$

互信息可以直观判断非线性信息间的影响程度, 从而给出信息相关性刻画度量。主要应用领域包括属性相关分析、故障诊断等。潜射武器系统信息类型繁多, 各类数据的相关性难以直观判断, 可利用互信息相关分析技术实现多类数据的潜射武器健康状态特征提取。

2.3.2 基于矩阵计算的相关系数

对于高维、多变量数据, 采用成熟的矩阵计算方法实现数据属性的相关分析是一种有效的技术途径。基于矩阵的计算方法是将高维数据以矩阵的形式描述, 并采用正则化方法实现属性的相关分析。RV 系数是矩阵计算中的基本相关系数, 是在皮尔逊系数的基础上重新构造的相关系数。RV 的计算依赖于构造矩阵 L, M , 这两个矩阵根据具体问题构造不同的矩阵, 具有较强的通用性和泛化能力。

3 大数据分析在潜射武器健康管理中的应用

3.1 潜射武器数据分析

随着部队潜射武器作战训练任务的不断增加, 潜射武器武器系统产生了大量数据, 包括: 试验数据、测试数据、环境数据、故障数据、维修数据、作战训练数据等。

数据类型包括: 数字、文字、图像、视频等, 且由于记录

格式不统一, 数据类型也不统一。数据的存储方式主要包括: 电子介质存储和纸质存储。电子介质主要包括测试数据、各类图像、视频等。纸质数据主要包括各类测试数据、环境数据、故障数据、维修数据、作战训练数据等。

3.2 潜射武器健康状态评估

潜射武器健康状态是综合保障中的关键问题之一。潜射武器的健康状态评估与装备的使用时间、使用频率、装备测试数据、环境数据、维修数据等都有关系。另一方面, 潜射武器武器系统由电子设备、机械设备、光学设备等组成, 涉及多个学科领域, 要评估潜射武器的健康状态是一项综合工程。

健康状态评估需要利用大量装备基础数据, 在这些数据的基础上, 对装备各个部件状态进行评估和预测。通过上述分析可以发现, 神经网络方法比较适于装备状态健康评估。但是由于健康状态数据繁多, 可以使用相关分析技术和粒计算方法对数据进行属性约简或剔除冗余数据。在神经网络模型中, 卷积神经网络更适于健康状态评估。基本技术路线如图 4 所示。

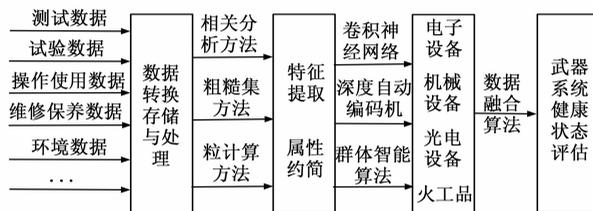


图 4 潜射武器健康状态评估技术路线图

3.2.1 潜射武器大数据存储与处理

由于潜射武器数据来源不一、数据量大, 数据格式、数据类型、数据精度等均有差异。所以, 首先需要对各类数据进行处理, 主要处理方式包括: 数据的归一化、数据格式的统一、数据的提取、数据的统一表示等。

数据存储采用 B/S 架构, 层次化设计分为硬件层、虚拟化云平台、分布式计算平台、数据服务平台和平台管理监控。虚拟化云平台层由云平台服务集群组成, 将硬件资源转化为多个虚拟计算机, 构成虚拟机网络。分布式计算平台运行分布式文件系统、分布式数据库及分布式计算框架, 为装备信息的存储和访问提供底层支撑。数据服务平台上构建装备信息数据仓库将通过各种来源和方式获得的数据根据其使用需求进行统一的抽取、转换并存放, 并向上层应用软件提供统一的访问接口, 为用户提供数据服务。平台管理监控用于监视云平台服务集群软硬件运行状态, 进行集群的管理维护, 并制定相关管理策略。

3.2.2 潜射武器数据特征提取

由于大数据固有的价值稀疏性, 根据不同的应用领域, 需要从海量数据中提取特征。特征提取方法分为两类: 属性约简和数据变换, 包括相关分析、粗糙集、粒计算^[14]、小波变换等。

相关分析可实现数据属性之间的关联性计算, 提取与决策相关的属性, 删除与决策相关性弱的属性, 实现属性约简, 进而为后续的状态评估提供支持。相关分析适用于量化数据的处理。

粗糙集是一类针对模糊数据的属性约简方法^[15], 该方法

保持相对分类能力不变的条件下, 删除冗余的、不必要的属性或属性值, 达到知识简化的目的。

3.2.3 潜射武器健康状态评估

潜射武器系统是由电子设备、机械设备、光电设备等多类设备组成的复杂综合大型系统, 由于各类装备失效机理、数据等方面存在较大差异, 装备健康状态评估的方法也有所差异。

以电子设备为例, 说明健康状态评估技术路线 (如图 5 所示)。根据数据类型的不同, 电子设备的健康状态评估采用不同的技术路线。

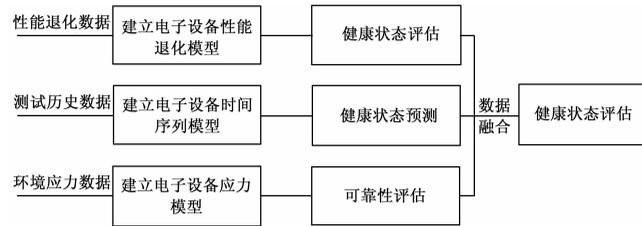


图 5 电子设备健康状态评估技术路线

如果退化数据丰富, 可采用深度神经网络、证据推理等方法建立电子设备性能退化模型, 并进行健康状态实时评估。如果电子设备历史测试数据比较丰富, 可采用神经网络、支持向量机等方法建立时间序列预测模型, 实时预测装备健康状态。如果环境应力数据比较丰富, 可采用应力分析法建立可靠性评估模型, 实现设备的可靠性评估, 这种方法比较适合于长期贮存的导弹武器系统。如果各类数据都比较丰富, 则可以采用数据融合的方法实现健康状态评估。

4 结束语

潜射武器的健康状态评估是一个复杂的系统性工程, 也是装备管理部门和部队十分关心的问题, 是潜射核力量形成的重要支撑因素。随着大数据分析技术的迅速发展以及潜射武器各类数据的不断增加, 利用大数据分析技术解决潜射武器健康状态评估问题已经成为一个行之有效的办法。通过上述分析可以发现, 有些大数据分析技术比较适合潜射武器的健康状态评估, 有些分析技术适合于数据处理。虽然当前潜射武器的各类数据

暂时还达不到大数据的规模, 但是应用大数据分析技术解决当前潜射武器系统的健康状态评估成为一个有效的可选方案。

参考文献:

- [1] 程学旗, 靳小龙, 王元卓, 等. 大数据系统和分析技术综述 [J]. 软件学报, 2014, 25 (9): 1889-1908.
- [2] 周飞燕, 金林鹏, 董 军. 卷积神经网络研究综述 [J]. 计算机学报, 2017 (40): 1-23.
- [3] 王伟凝, 王 励, 赵明权, 等. 基于并行深度卷积神经网络的图像美感分类. 自动化学报, 2016, 42 (6): 904-914.
- [4] 孙艳丰, 齐光磊, 胡永利. 基于改进 Fisher 准则的深度卷积神经网络识别算法 [J]. 北京工业大学学报, 2015, 41 (6): 835-841.
- [5] 孙 文, 邵思羽, 严加强. 基于稀疏自动编码深度神经网络的感应电动机故障诊断 [J]. 机械工程学报, 2016, 52 (9): 65-71.
- [6] 张 蕾, 章 毅. 大数据分析的无限深度神经网络方法 [J]. 计算机研究与发展, 2016, 51 (1): 68-79.
- [7] 章 毅, 郭 泉, 王建勇. 大数据分析的神经网络方法 [J]. 工程科学与技术, 2017, 49 (1): 9-18.
- [8] 郑 斌. 基于改进遗传算法的不完整大数据填充挖掘算法 [J]. 微电子学与计算机, 2016, 41 (2): 35-38.
- [9] 邹心遥, 陈敬伟, 姚若河. 采用粒子群优化的 SVM 算法在数据分类中的应用 [J]. 华侨大学学报 (自然科学版), 2016, 37 (2): 171-174.
- [10] 董跃华, 刘 力. 基于自适应改进粒子群优化的数据离散化算法 [J]. 计算机应用, 2016, 36 (1): 188-193.
- [11] 洪 铭, 柳培忠, 骆炎民. 一种基于多智能体策略的人工免疫神经网络数据分类方法 [J]. 计算机应用研究, 2017, 34 (1): 151-156.
- [12] 张 健, 王笑竹. 人工免疫聚类算法在城市交通流量分析中应用 [J]. 中央民族大学学报 (自然科学版), 2017, 26 (1): 56-60.
- [13] 梁吉业, 冯晨娇, 宋 鹏. 大数据相关分析综述 [J]. 计算机学报, 2016, 39 (1): 1-18.
- [14] 苗夺谦, 李德毅, 姚一豫, 等. 不确定性与粒计算. 北京: 科学出版社, 2011.
- [15] 郭 平, 王 可, 罗阿理, 等. 大数据分析中的计算智能研究现状与展望. 软件学报, 2015, 26 (11): 3010-3025.

(上接第 245 页)

接收数据时, 播放器将从光纤接口读取数据并传输到 FPGA 的 GTP 模块, 图像数据接收控制逻辑再将 GTP 数据解析缓存到内部 FIFO 中, 然后通过 SRIO 控制器, 将 FIFO 内部的数据传输到 DSP 的内部 RAM 中, 并在传输一定的数据量后发送中断到 DSP 端口通知数据到达, DSP 在收到通知中断后, 从内部 RAM 中把数据传输到外部 DDR2 存储器上, 在一帧时序内, 将每次的数据组合成一幅完整的图像并缓存在 DDR2 上。发送图像数据时, 播放器首先从 DSP 的 DDR2 空间把数据传输到内部 RAM 上, 再通过 SRIO 端口, 把数据发送到 FPGA 的内部 FIFO 中, FPGA 的 FIFO 在达到一定的数据量后自启动读取操作, 从 FIFO 中输出数据到 GTP 端口, 通过 GTP 把数据发送到光纤接口上, 最终实现了图像数据的发送。

4 结束语

本文设计了一种采用 FPGA 和 DSP 技术实现的红外仿真

图像高速实时数据流播放器, 用于光电经纬仪红外仿真系统的虚拟目标与真实背景图像的实时融合处理, 实现了高带宽、高帧频、高分辨率的图像流数据的实时输出, 解决了仿真目标与真实背景融合的大数据量运算问题。在实际应用中, 取得了较好的视觉效果, 该技术能够广泛应用于其他光学仿真系统。

参考文献:

- [1] Gonzalez R C, Woods R E, 阮秋琦, 等译. Digital Image Processing / 数字图像处理 [M]. 北京: 电子工业出版社, 2002.
- [2] 于正林, 谭 微, 姜 涛. 飞行目标跟踪训练视景仿真技术研究 [J]. 计算机仿真, 2011, 28 (4): 379-383.
- [3] 王习文, 陈 娟, 王秋平. 基于视景仿真的光电经纬仪模拟训练器 [J]. 吉林大学学报 (工学版), 2011, 41 (2): 509-513.
- [4] 王习文, 陈 娟, 王秋平. 光电经纬仪跟踪飞行目标的一种仿真方法研究 [J]. 计算机仿真, 2009, 26 (4): 38-42.
- [5] 田 耕, 徐文波. Xinlinx FPGA 开发实用教程 [M]. 北京: 清华大学出版社, 2010.