

M 现代计算机

XIANDAI JISUANJI

第29卷第10期 (总第778期)

半月刊 (1984年创刊)

2023年5月25日出版

主管单位 中山大学
主办单位 广州中山大学出版社有限公司
出版单位 广东现代计算机杂志社有限公司
发行 广东省报刊发行局 (全国公开发行)
印刷 广州一龙印刷有限公司
社长 黄少伟
主编 石玉珍
编委 邹岚萍 熊锡源 李文 石玉珍 梁嘉璐
地址 广州市海珠区新港西路135号
中山大学内 (510275)
电话 020-84112089 (编辑部)
网址 www.moderncomputer.cn
电子邮箱 tougao@moderncomputer.cn

ISSN 1007-1423
CN 44-1415/TP

邮发代码: 46-121
定价: 30.00元



邮局订刊二维码



现代计算机
官方网站二维码



M 现代计算机

第29卷第10期 (总第778期)

2023年5月

2023年5月 第29卷

第10期 (总第778期)



ISSN 1007-1423
CN 44-1415/TP

MODERN COMPUTER

现代计算机



中山大学出版社 主办

中国期刊数据库CNKI全文收录期刊
中国学术期刊（光盘版）收录期刊
中文科技期刊数据库全文收录期刊
中国核心期刊（遴选）数据库收录期刊
中国学术期刊综合评价数据库收录期刊

- ◆ 研究与开发：计算机发展和软、硬件开发的理论研究
- ◆ 图形图像：重点为与图形图像相关的理论及实践研究
- ◆ 开发案例：基于某方面的计算机开发案例研究与分析
- ◆ 实践与经验：计算机应用的实例及心得

版权声明

1. 本刊版权属于杂志社所有，其他报刊或网站如需转载，须经本刊同意，注明转载自本刊并付作者稿酬。
2. 本刊来稿恕不退还，请自留底稿。请勿一稿多投。来稿文责自负，严禁抄袭。对侵犯他人版权或其他权利的稿件，本刊概不承担连带责任。
3. 对所投稿件，本刊编辑有权根据刊物的需要进行删改或调整。
4. 凡是刊登在本刊的稿件，即表示作者同意稿件在《现代计算机》网站、中国期刊数据库CNKI、中国学术期刊（光盘版）、中文科技期刊数据库、中国科技期刊（遴选）数据库、中国学术期刊综合评价数据库等媒体发布。

目次

研究与开发

- 基于多尺度循环网络的运动模糊图像复原方法 张甜, 卢振坤, 纪佳奇, 刘胜 (1)
- 基于故障树与贝叶斯网络的汽车故障诊断方法 李星 (9)
- 项目申报书形式审查算法 张彬, 程健峰 (15)
- 基于层聚合机制的密集场景人群计数方法 席梦园, 严华 (20)
- 基于Stacking的个人薪资预测研究 林心慧, 陈超 (25)
- 智慧图书馆建设背景下构建图书智能即时剔旧体系的应用研究 罗可 (30)

实践与经验

- 基于主观逻辑和拍卖算法的公证人组跨链交易机制 廖政, 魏博众, 罗健炜 (34)
- 多关系型数据库向图数据库转换与同步方法 郭良银 (42)
- 水声网络路由协议实现与测试 张振儒, 杜秀娟 (50)
- 数据元治理框架及建议 程红云, 方亮, 陈美 (57)
- 基于规则引擎的DXF与空间矢量数据互转方法研究 李辉勤, 龙杰 (63)

开发案例

- 低代码赋能电信运营商行业数字化转型 赵龙, 冯强中, 汪伟伟, 许磊, 解宇恩 (67)
- 基于知识图谱的法理知识服务系统设计与实现 莫城为, 温怡烽, 邓邦鹏 (74)
- 基于微信小程序的在线答题系统的设计与实现 岑俊慷, 刘卫 (81)
- 基于STM32的物联网智能家居控制系统 肖榆瀚, 柯玮翔 (87)
- 汽车零配件订单智能管理系统开发设计 黄梅香 (93)
- 基于Funcode平台的《华夏宝藏》游戏设计与开发
..... 李晓蕾, 王守佳, 刘钰周, 赵晨贺, 黄瀚霖, 王猛 (98)
- 一种电波产品业务平台智能监控系统 刘晓雷, 刘书志, 李静静 (104)
- 基于Web的“湘赣边”线上全景漫游系统开发 邱望, 熊钰, 梁莉菁 (109)
- 软件测试管理系统的研究与应用 付春子, 唐海涛, 徐进 (113)
- 无线充电智能小车设计 徐琛浩, 胡艺轩, 丁纵辽, 周飞雪, 张岚, 赵春锋 (117)

Modern Computer

(Vol. 29, No. 10; May. 25, 2023)

CONTENTS

Research and Development

Motion blurred image restoration method based on multi-scale recurrent network	(1)
Automotive fault diagnosis technology based on fault tree and Bayesian network	(9)
Formal examination algorithm for project application materials	(15)
Crowd counting in dense scenes via layer aggregation mechanism	(20)
Prediction of personal salary based on Stacking fusion model	(25)
Research on application of intelligent and real-time weeding system for books in the construction of smart library	(30)

Practice and Experience

Cross-chain trading mechanism of notary group based on subjective logic and auction algorithm	(34)
Conversion and synchronization method from multi-relational database to graph database	(42)
Realization and test in lake of routing protocol for underwater acoustic networks	(50)
Governance framework and suggestions of data element	(57)
Research on the method of converting DXF and space vector data based on rule engine	(63)

Development Solution

Low-code enabling the digital transformation of telecom operators	(67)
Design and implementation of the legal knowledge server system based on knowledge graph	(74)
The design and realization of the online answering system based on Wechat small program	(81)
Intelligent home control system for IoT based on STM32	(87)
Development and design of auto parts order intelligent management system	(93)
Design and development of “Huaxia Treasure-Trove” game based on Funcode platform	(98)
An intelligent monitoring system for radio product business platform	(104)
Development of “Xiang-Gan Bian” online panoramic roaming system based on Web	(109)
Research and application of software test management dystem	(113)
The design of wireless charging intelligent car	(117)

研究与开发

文章编号: 1007-1423(2023)10-0001-08

DOI: 10.3969/j.issn.1007-1423.2023.10.001

基于多尺度循环网络的运动模糊图像复原方法

张甜^{1,2}, 卢振坤^{1,2*}, 纪佳奇^{1,2}, 刘胜^{1,2}

(1. 广西民族大学电子信息学院, 南宁 530006; 2. 广西高校智慧无人系统与智能装备重点实验室, 南宁 530006)

摘要: 针对目前图像去模糊恢复细节不好、泛化性能不高的问题, 提出了一种基于多尺度循环网络的运动模糊图像复原算法, 在编码端将多尺度特征融合模块和残差密集连接模块融合, 引入多尺度残差密集型连接模块。同时引入一种注意力机制并将融合了注意力机制的多尺度残差密集连接模块作为网络的基本结构。实验结果表明, 与生成对抗网络相比, 在 GOPRO 数据集的 PSNR 和 SSIM 最大分别提升了 4.13 dB 和 0.0254 dB, 和近年来效果最明显的 SRN 相比, Kohler 数据集上的 PSNR 和 SSIM 分别提升了 0.31 dB 和 0.0179 dB, 具有更好的泛化性能。

关键词: 多尺度特征融合模块; 残差密集型连接模块; 注意力机制

0 引言

在现代社会, 图像是一个重要的交流用具。但在拍照中出现相机轻微抖动、拍摄对象的移动等现象都会造成图像原有像素点改变, 从而导致图像变模糊, 这个过程称作运动模糊。因此, 模糊图像复原技术变得十分重要^[1]。图像复原^[2]是使用图像的有关知识理论, 将模糊图像恢复到清晰图像的过程。

图像模糊用公式表示为

$$B = I \otimes K + N \quad (1)$$

式中: B 代表模糊图像, I 代表清晰图像, K 代表模糊核或点扩散函数 (point spread function, PSF), N 代表噪声, \otimes 代表卷积操作^[3]。图像复原方法^[4]主要包括盲去模糊复原和非盲去模糊复原。非盲去模糊是在模糊核明确的前提下恢复图像; 盲去模糊是在模糊核不明确的前提下恢复图像。代表性的去模糊方法是: 空间域算法 Lucy-Richardson 方法^[5]、频率域算法维纳滤波^[6]、基正则化^[7], 这些方法都是利用数学模型

确定 K 之后还原图像。但是噪声对其结果有影响。非盲去模糊一开始恢复的方法^[8,9]是使用各式各样的约束条件, 引用图像先验知识用来表示图像模糊, 但事实上非均匀运动模糊基本都是不能预测的, 因此这些方法并不实用。

近几年, 深度学习兴起并广泛应用到图像处理等领域中。首先 Sun 等^[10]使用一种新的基于卷积神经网络 (convolutional neural network, CNN) 的方法去除图像模糊, 该方法在实验中虽然达到了较好的去模糊效果, 因为它采用相同方法估计模糊核, 只能应用在特定的模糊结构中。之后 Nah 等^[11]提出了一个多尺度的卷积神经网络, 应用于动态场景去模糊, 取得了非常好的效果。Kupyn 等^[12]提出一种新的生成对抗网络 (deblur generative adversarial network - v2, DeblurGAN-v2), 对比 Nah 等^[11]提出的去模糊方法, 此方法去模糊效果大大提高。最后 Tao 等^[13]提出一种多尺度循环网络 (scale-recurrent network, SRN), 相比 Kupyn 等^[12]提出的方法有了更好的

收稿日期: 2023-02-02 修稿日期: 2023-02-25

基金项目: 国家自然科学基金资助项目 (61561008); 广西自然科学基金 (2018GXNSFAA294019)

作者简介: 张甜 (1999—), 女, 陕西黄陵人, 硕士, 研究方向为图像处理、计算机视觉; *通信作者: 卢振坤 (1979—), 男, 广西南宁人, 教授, 博士, 研究方向为数字图像处理、计算机视觉、深度学习、探测与成像; 纪佳奇 (1997—), 男, 江苏徐州人, 硕士, 研究方向为深度学习、图像处理; 刘胜 (1998—), 男, 湖南衡阳人, 硕士, 研究方向为人群计数、目标检测

去模糊效果。

以上方法虽然都取得很好的效果，但存在恢复细节不好以及泛化性能低的问题。因此针对上述问题，本文做了如下工作：

(1) 在编码端引入多尺度特征融合模块(multi-scale feature extraction block, MFEB)，增加感受野、提取图像更多不容易被发现的特征信息。

(2) 提出多尺度残差密集连接模块(multi-scale residual dense block, MRDB)，有效结合图像局部信息和全局信息，保证模型训练时信息的完整性。

(3) 引入注意力机制模块并且用融合了注意力机制的多尺度残差密集连接模块作为网络的基本结构，精准采集物体的信息。

1 基于多尺度循环网络的方法

本文网络模型由三个尺度(三层)构成，采用从粗到细的网络学习方法，借鉴高斯金字塔的分辨率分配，输入端模糊图片分别为 64×64 、 128×128 、 256×256 三个不同的分辨率，如图1(a)所示，其中每一层都是由编码网络、长短时记忆网络(long short term memory, LSTM)^[14]和解码网络构成。输入模糊图片首先经过编码端，编码端是由三个相同的单元构成，如图1(b)所示，每个单元都是由卷积、两个MRDB模块以及通道和空间注意力机制(convolutional block attention module, CBAM)组成的基本结构，在MRDB中使用MFEB来增加感受野。编码端对图片进行特征提取后，输入LSTM模型中，再通过解码端对图片进行恢复(解码端结构和编码端一样)，第一层的输出结果经过上采样，与下一层的输入图像作为第二层的输入端。使用跳跃连接融合特征信息，避免随着网络层数的增加信息丢失。防止模型训练时间过长，训练过程中整个网络共同使用权值参数，所有卷积核的大小都为 5×5 ，除非特殊说明。

1.1 残差密集连接模块

如图2所示，残差密集连接模块(residual dense block, RDB)是在残差模块(residual block, RB)和密集连接模块(dense block, DB)的基础上进行改进的。使用RDB模块有效结合浅层和深

层信息。RDB还可以将上一个RDB模块的状态融合到当前RDB的每一个卷积层，也可以和之后的RDB建立联系，将从上一层和当前层学习到的特征信息传输到之后每一层，形成RDN网络。如图2(c)所示，RDB是对每一个卷积层的输出结果进行结合，充分利用每一层的特征信息，最后通过 1×1 的卷积调整通道数目，减少参数量和计算量，提高图像质量。

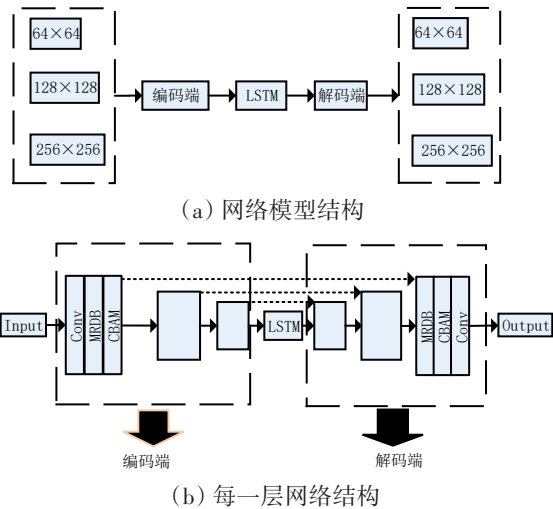


图1 本文网络模型

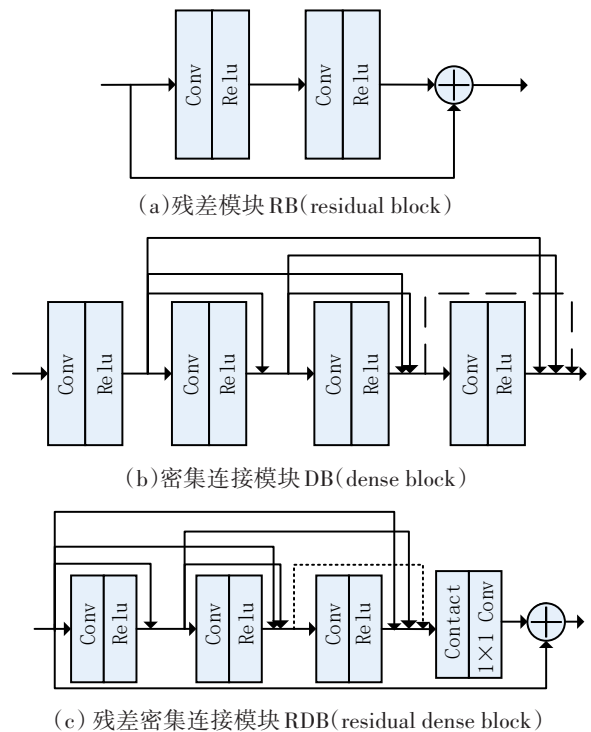


图2 前期网络结构

1.2 多尺度残差密集连接模块

基于RDB模块提出一种新的网络模块MRDB，MRDB是在原有的RDB模块上进行改进的，为了获取更多的特征信息，提高训练效率，将RDB中的卷积模块改进为MFEB模块。

MFEB结合了不同的尺度信息，使用 3×3 和 5×5 的尺度信息，不仅能感知局部信息还能感知全局信息。如图3所示，第一层是一个 3×3 的卷积，第二层是两个 3×3 卷积的拼接，用两个 3×3 卷积代替一个 5×5 卷积。与一个 5×5 卷积相比，两个 3×3 卷积叠加使用不仅增加了非线性特性，还有效减少了参数数量和计算量。对比如表1所示，采用两个 3×3 卷积代替一个 5×5 卷积是有利于模型训练的。

表1 3×3 与 5×5 卷积对比

卷积核大小	参数个数	图片为 32×32 的感受野
3×3 卷积核	18	28
5×5 卷积核	25	28

参照RDB网络原理。如图3所示，RDB是从每一层的卷积中提取信息；而MRDB是从融合后的多尺度模块中提取信息，获取的特征信息更全面，将每一层MFEB的输出都输入到下一层，防止信息传输的过程中破坏了信息的完整性，最大化利用特征信息，有利于训练更深的网络模型。最后采用 1×1 的卷积核降低通道数目，减少参数数量和计算量。

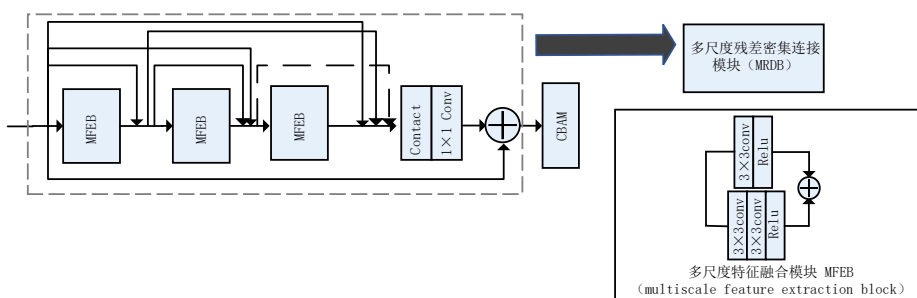


图3 融合注意力机制的多尺度残差密集连接模块

1.3 融合注意力的多尺度残差密集连接网络

本文采用CBAM^[15]注意力机制，如图4所示。CBAM注意力机制是一种简单且轻量型的公用模型，可以和所有卷积神经网络拼接，有效提升卷积神经网络的性能，模型开支非常小，广泛使用性已证明。

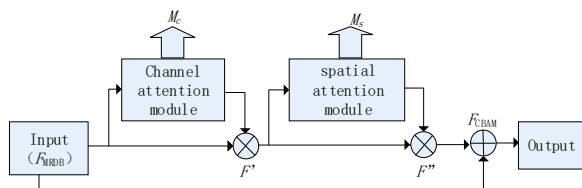


图4 CBAM注意力机制

如图3所示，使用融合了注意力机制的多尺度残差密集连接模块作为整体网络的基本结构。利用RDN实现原理，由融合了注意力机制的多尺度残差密集连接模块组成的基本结构如图5所

示， N 、 N^{-1} 、 N^{+1} 分别表示本模块、上模块、下一模块的状态， F_{N-1} 表示第 N^{-1} 层的输出结果， $F_{N,M}$ 表示第 N 层第 M 个MFEB模块的输出结果， $F_{N,LF}$ 表示第 N 层contract之后 1×1 卷积处理的结果， F_N 表示第 N 层MRDB的输出结果， F_{N+1} 表示 N^{+1} 层的输出结果， F_{CBAM} 表示注意力机制输出结果。则第 N 层、第 $M2$ 个网络层输出为

$$F_{N,M2} = \delta(W_{N,M2}[F_{N-1}, F_{N,M1}]) \quad (2)$$

同理可得出：

$$F_{N,M1} = \delta(W_{N,M1}[F_{N-1}]) \quad (3)$$

$$F_{N,M3} = \delta(W_{N,M3}[F_{N-1}, F_{N,M1}, F_{N,M2}]) \quad (4)$$

$$F_N = F_{N,LF} + F_{N-1} \quad (5)$$

$$F_N = F_N + F_{CBAM} \quad (6)$$

上式中： δ 代表ReLU激活函数， $W_{N,M2}$ 代表第 N 个、第 $M2$ 个网络层的权重， LF 表示 1×1 卷积的相关操作。图5中第 N 层融合了注意力机制的多尺度残差密集连接模块与 N^{-1} 和 N^{+1} 模块都存在

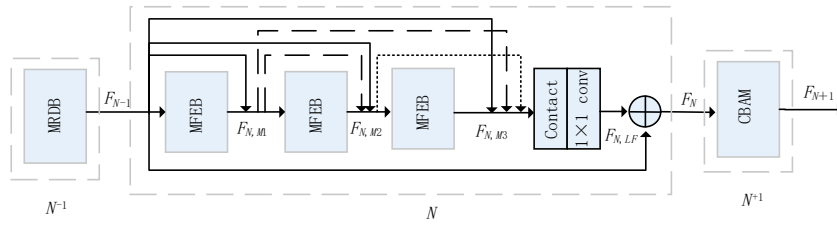


图5 由融合了注意力的多尺度残差密集连接模块组成的基本结构

联系，不仅保存了前馈特征、提取了局部特征，还将产生的结果融入后端网络。由融合注意力机制的多尺度残差密集连接模块组成的基本结构，在图像恢复中可以高效精确地提取物体信息。

1.4 损失函数

本文的损失函数为

$$L = \sum_{i=1}^n \frac{W_i}{N_i} \|I^i - I_{gt}^i\|_2^2 \quad (7)$$

使用欧式距离(Euclidean distance)损失函数，式中 I^i 和 I_{gt}^i 分别是第 i 个尺度上的输出结果值和实际结果。 W_i 为第 i 个尺度上的权重， N_i 为第 i 个尺度上归一化的元素数。

2 实验验证

2.1 实验数据集

2.1.1 GOPRO 数据集

GOPRO 数据集是由 Nah 等^[11]提出的，该数据集由成对的模糊和真实的图像构成，图像分辨率为 1280×720，一共 3214 张图像对，2013 对用作训练集，1111 对用作测试集。是运动模糊最常用的数据集，因此本文使用该数据集作为训练集和测试集。

2.1.2 Kohler 数据集

Kohler 数据集是由 Kohler 等^[16]提出的，该数据集创建了 12 个模糊核，使用该模糊核对图像处理，图像总数为 48 张，是测试图像盲去模糊的基准数据集，因此本文使用该数据集作为测试集。

2.2 实验参数设置

本文实验运行环境如下：操作系统为 Windows10、GPU 为 Nvidia GeForce RTX2070、内存为 16 GB、学习框架为 TensorFlow。为了实验结果有效，本文的所有实验都是在上述环境上实

现。在整个模型训练过程中，使用 Adam 求解器，参数设置为 $\beta_1=0.9$ ， $\beta_2=0.999$ 。学习率为初始的 0.0001 呈指数衰减到 1×10^{-6} 。实验表明，训练 2000 epoch 就可以达到收敛状态，整个训练过程需要 108 个小时。每次随机选取 2 张模糊图片，将图片裁剪为 256×256 像素大小后作为网络输入。由于本文网络是全卷积网络，所以可以输入任何尺寸的图片，训练集和测试集都为 720×1280 大小的图像。

2.3 实验评价指标

本文选用峰值信噪比(peak signal to noise ratio, PSNR)、结构相似度(structural similarity, SSIM)和图像恢复时间(单张图像的平均值)作为图像恢复评价指标。

2.3.1 PSNR

$$PSNR = 20 \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{MSE}} \right) \quad (8)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I_{(i,j)} - K_{(i,j)}]^2 \quad (9)$$

式中： MAX_I 是图片最大像素值； I 为真实图像； K 为恢复的图像； $m \times n$ 代表图像 I 和 K 的分辨率。 $PSNR$ 越大，说明图像恢复效果越好，和真实图像越像。

2.3.2 SSIM

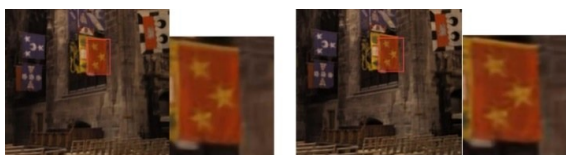
$$SSIM(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\delta_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\delta_x^2 + \delta_y^2 + C_2)} \quad (10)$$

式中： μ_x 和 μ_y 分别代表真实图像和恢复图像的平均值， δ_x 和 δ_y 分别代表真实图像和恢复图像的标准差， δ_{xy} 代表真实图像和恢复图像的协方差， C 为常数。 $SSIM$ 越接近 1，说明和真实图像越接近。

2.4 实验结果及分析

2.4.1 注意力机制对实验影响

为了证明引入注意力机制的有效性，在实验参数保持相同的前提下做了两组对比实验，其中一组是本文提出的网络结构，另一组是去除注意力机制的网络结构。结果的评价指标为 *PSNR*、*SSIM* 和图像恢复时间，选用 Kohler 数据集作为测试集。两者在 Kohler 数据集上的实验结果如图 6 所示。



(a) 有注意力机制

(b) 无注意力机制

图 6 有无注意力机制对比

如图 6 所示，明显能看到添加注意力机制算法恢复的五角星图案边缘以及轮廓更加清晰，恢复的细节更好。

表 2 为有无注意力机制算法在 Kohler 数据集上的数据结果，可以看出，有注意力机制的 *PSNR* 和 *SSIM* 更好，虽然其图像恢复时间滞后了 0.01 s。但总体来说，*PSNR* 和 *SSIM* 得到了提升，并且图像恢复细节提升，因此滞后的 0.01 s 可以忽略不计。两者训练时间相同，图像恢复时间也相差不大，说明两者网络模型的大小、参数量差不多，在此前提下引入注意力机制的 *PSNR* 和 *SSIM* 更好，证明它有更好的恢复效果和泛化性能。

表 2 引入 CBAM 注意力机制对实验结果的影响

算法	<i>PSNR</i> /dB	<i>SSIM</i> /dB	恢复时间/s
本文算法	19.83	0.7684	1.59
去除 CBAM 的算法	19.80	0.7650	1.58

2.4.2 实验分析及对比

为了验证本实验的图像复原效果，在 GOPRO 数据集和 Kohler 数据集上进行测试。本文选用了近几年模糊图像复原较好的算法与之比较，选用 Kupyn 等^[17]提出的生成对抗网络 (deblur generative adversarial network, Deblur-

GAN)、Kupyn 等^[12]提出的 DeblurGAN-v2 算法，包括 Inception-ResNet-v2 和 Mobile Net 两种网络 (分别表现出了最好的运动模糊图像还原效果和最快的恢复时间)、Tao 等^[13]提出的图像复原效果最好的 SRN 网络。

图 7 所示为上述算法以及本文方法在 GOPRO 数据集上的测试结果，从图 7(b)、(c) 可以看出，DeblurGAN 网络和 DeblurGAN-v2(Inception-ResNet-v2) 网络恢复的车牌数字模糊，道路的线条没有恢复出来。从图 7(d) 可以看出，DeblurGAN-v2(Mobile Net) 恢复的车牌数字边缘不平滑。但可以看出本文方法和 SRN 恢复图像的效果好于其他生成对抗网络算法，如图 7(f) 所示，虽然本文实验恢复的字母边缘还是不太平滑，比 SRN 略差一些，但它能将 SRN 复原不出来的道路线条恢复出来，因此本文方法在恢复某些细节方面做的更好，和 SRN 算法达到了不相上下的效果。如图 8 所示，选取 Kohler 数据集上两张不同的图片进行对比，可以看出第一幅图钟表长针下面罗马数字的恢复情况，明显本文方法恢复的罗马数字边缘比较平滑，其他方法恢复的罗马数字边缘出现了重影；图片中数字“5”的恢复，不论是细节纹理还是边缘，本文达到最好的效果。从第二幅国旗图中也能看出来，本文恢复的图像不管是在轮廓还是色调方面最接近清晰图像，具有很好的复原效果。因此，本文方法在 Kohler 数据集上获得了最好的去模糊效果。

如表 3 所示，在 GOPRO 数据集上，所提算法的结果优于一系列生成对抗网络算法，*PSNR* 取得了最好的结果，*SSIM* 虽然和 DeblurGAN-v2(Inception-ResNet-v2) 相比有 0.0413 dB 的差距，但 *PSNR* 比它高 0.88 dB，并且在 Kohler 数据集上取得高于它的结果。在 Kohler 数据集上，所提方法的 *PSNR* 和 *SSIM* 与生成对抗网络算法相比 *PSNR* 和 *SSIM* 最大提升了 0.04 dB 和 0.0128 dB。

和近年来效果最明显的 SRN 算法相比。在 GOPRO 数据集上，虽然本文的 *PSNR* 和 *SSIM* 还和 SRN 存在 1 dB 和 0.0603 dB 的差距，但本文恢复图像细节优于它；加之在 Kohler 数据集上，本文算法的 *PSNR* 和 *SSIM* 比 SRN 分别提升了 0.31 dB 和 0.0179 dB。对比算法的恢复图像时

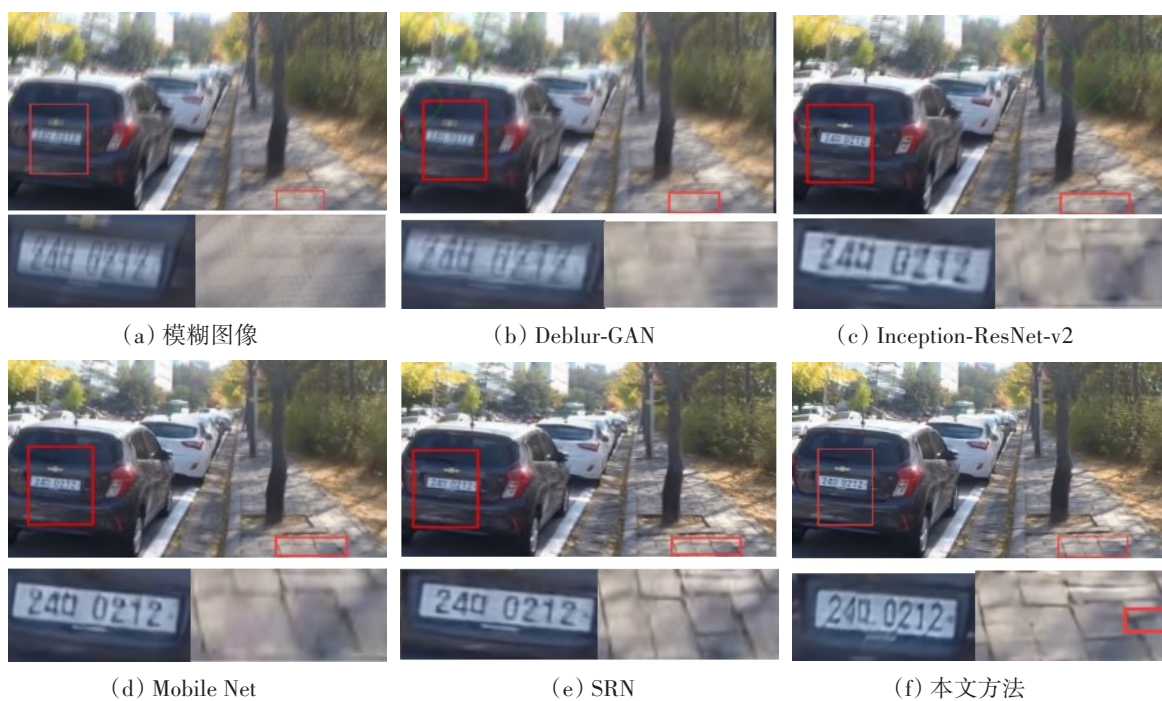


图7 GOPRO数据集对比结果

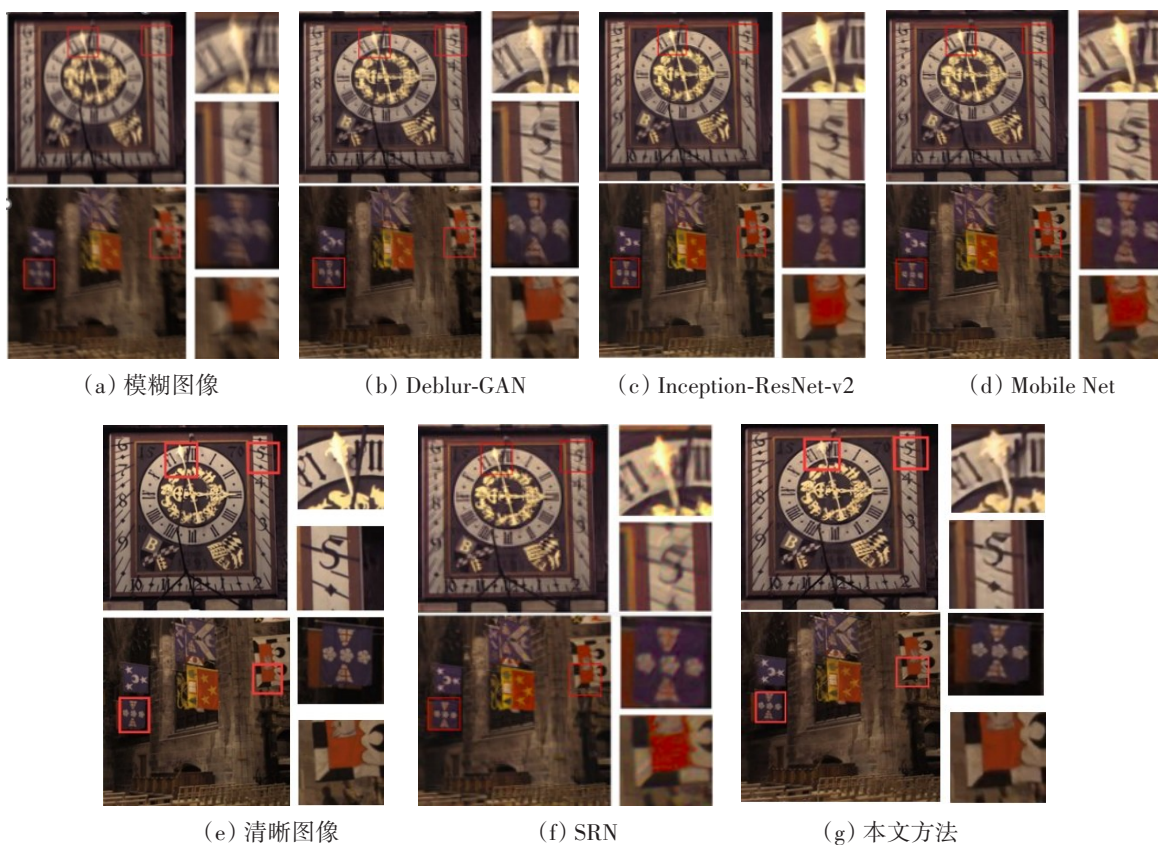


图8 Kohler数据集对比结果

表3 数据结果对比

数据集	不同算法	PSNR/dB	SSIM/dB	复原时间/s
GOPRO	SRN	30.15	0.9331	1.383
	DeblurGAN	25.02	0.8474	1.504
	DeblurGAN-v2(Inception-ResNet-v2)	28.27	0.9141	1.083
	DeblurGAN-v2(Mobile Net)	25.96	0.8731	0.821
	本文方法	29.15	0.8728	1.603
Kohler	SRN	19.52	0.7505	1.375
	DeblurGAN	19.75	0.7447	1.492
	DeblurGAN-v2(Inception-ResNet-v2)	19.50	0.7469	1.082
	DeblurGAN-v2(Mobile Net)	19.79	0.7556	0.811
	本文方法	19.83	0.7684	1.592

间,由于本文引入了多尺度残差密集连接模块,充分利用了浅层信息并将其与深层信息结合,因此整体网络模型会变大,复原时间略高于其他算法。

综合各方面评价指标,本文使用GOPRO数据集作为网络的训练集,但在使用和训练集不同的Kohler数据集上取得了更好的数据结果。由此说明,所提方法具有更好的泛化性能。

3 结语

针对目前运动模糊图像复原算法中存在恢复的图像细节不够清楚,网络泛化性能低的问题。本文通过使用MFEB模块,采用多尺度网络提取潜在信息;构建基于RDB的MRDB模块;使用CBAM注意力机制并用融合了注意力机制的多尺度残差密集连接模块作为网络的基本结构,提出了一种新的多尺度循环网络的运动模糊图像复原方法。在加深网络的同时最大化利用输入图像的信息,有效地提取物体及其周边的特征信息。

实验结果表明,本文方法在恢复图像的细节上略胜一筹;在GOPRO数据集和Kohler数据集上取得较为理想的结果,尤其是在Kohler数据集上达到了最好的效果,证明本文算法具有更好的泛化性能。

参考文献:

- [1] 王玉全,隋宗宾.运动模糊图像复原算法综述[J].微型机与应用,2014,33(19):54-57.
- [2] 黄正源,谢维成,黄化入,等.图像运动模糊还原技术综述[J].计算机工程与应用,2020,56(24):28-34.
- [3] 李思俭,樊祥,程正东.红外视频帧运动模糊复原技术研究[J].激光与红外,2018,48(3):389-395.
- [4] 刘平,刘传才,张佳洛.残差分离卷积神经网络的图像去运动模糊技术[J].计算机与数字工程,2020,48(8):1975-1980.
- [5] 闻毅.Lucy-Richardson算法在模糊图像复原中的应用[J].智能计算机与应用,2015,5(1):95-97.
- [6] 刘泽坤,张涛.维纳滤波算法在散焦模糊图像复原中的应用研究[J].信息通信,2020(8):63-64.
- [7] 王宗跃,夏启明,蔡国榕,等.自适应图像组的稀疏正则化图像复原[J].光学精密工程,2019,27(12):2713-2721.
- [8] SHENG B, LI P, FANG X, et al. Depth-aware motion deblurring using loopy belief propagation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(4): 955-969.
- [9] WHYTE O, SIVIC J, ZISSERMAN A, et al. Non-uniform deblurring for shaken images [J]. International Journal of Computer Vision, 2012, 98(2): 168-186.
- [10] SUN J, CAO W F, XU Z B, et al. Learning a convolutional neural network for non-uniform motion blur removal [C] // Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2015: 769-777.
- [11] NAH S, KIM T H, LEE K M. Deep Multi-scale convolutional neural network for dynamic scene deblurring [C] // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2017: 257-265.
- [12] KUPYN O, MARTYNIUK T, WU J R, et al. DeblurGAN-v2: deblurring (orders-of-magnitude)

- faster and better [C]//Proceedings of the 2019 IEEE Conference on International Conference on Computer Vision, Piscataway: IEEE, 2019: 8877-8886.
- [13] TAO X, GAO H Y, SHEN X Y, et al. Scale-recurrent network for deep image deblurring [C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2018: 8174-8182
- [14] ZHAO Z, CHEN W, WU X, et al. LSTM network: a deep learning approach for short-term traffic forecast [J]. IET Intelligent Transport Systems, 2017, 11(2): 68-75.
- [15] WOO S H, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Proceedings of the 2018 IEEE Conference on European Conference on Computer Vision, Cham: Springer, 2018: 3-19
- [16] KOHLER R, HIRSCH M, MOHLER B, et al. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database [C]//Proceedings of the 2012 IEEE Conference on European Conference on Computer Vision, Cham: Springer, 2012: 27-40.
- [17] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. DeblurGAN: blind motion deblurring using conditional adversarial networks [C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2018: 8183-8192.

Motion blurred image restoration method based on multi-scale recurrent network

Zhang Tian^{1,2}, Lu Zhenkun^{1,2*}, Ji Jiaqi^{1,2}, Liu Sheng^{1,2}

(1. College of Electronic Information, Guangxi Minzu University, Nanning 530006, China;

2. Key Laboratory of Intelligent Unmanned System and Intelligent Equipment, Education Department of Guangxi Zhuang Autonomous Region, Nanning 530006, China)

Abstract: Aiming at the problem that the image deblurring-restoring details is not good and the generalization performance is not high. A motion blur image restoration algorithm based on multi-scale recurrent network is proposed. At the encoder, the Multi-scale Feature Extraction Block and Residual Dense Block are fused, and the Multi-scale Residual Dense Block is introduced. At the same time, an attention mechanism is introduced and the multi-scale residual dense block module integrated with attention mechanism is used as the basic structure of the network, which is beneficial to extract the channel and spatial feature information of the image. The experimental results show that compared with the GAN, the *PSNR* and *SSIM* of the GOPRO data set have increased by 4.13 dB and 0.0254 dB, respectively. Compared with the most effective SRN in recent years, the *PSNR* and *SSIM* of the Kohler data set have been improved respectively. 0.31 dB and 0.0179 dB, with better generalization performance.

Keywords: multi-scale feature fusion module; residual dense connection module; attention mechanism

文章编号: 1007-1423(2023)10-0009-07

DOI: 10.3969/j.issn.1007-1423.2023.10.002

基于故障树与贝叶斯网络的汽车故障诊断方法

李 星*

(西南交通大学制造业产业链协同与信息化支撑技术四川省重点实验室, 成都 611756)

摘要: 针对传统汽车故障诊断过程中不能高效快速定位故障原因而造成维修成本和维修时间增大的问题, 经过对基于故障树与贝叶斯网络的汽车故障诊断方法进行研究, 分析汽车故障数据, 构建故障树模型并将其转化为贝叶斯网络, 进行贝叶斯网络反向推理, 确定故障诊断顺序, 提高故障诊断的准确率, 快速响应服务需求, 帮助维修人员快速定位故障发生的位置以及故障原因, 从而提高维修效率, 提升服务质量。

关键词: 故障诊断; 故障树分析法; 贝叶斯网络

0 引言

汽车产业以其在国民经济中的重要地位和对经济增长的重要贡献被列为国家的战略性竞争产业^[1]。汽车产业可分为汽车售前产业和汽车售后服务市场, 在国外发达国家的汽车市场中, 汽车售后服务市场利润占比能达到汽车行业总利润的 60%~70%, 而我国目前在汽车售后服务市场的利润占比不足 40%^[2], 潜力巨大。维修保养业务是汽车售后服务市场的重要业务之一^[3], 然而就目前而言, 汽车维修行业仍存在如服务品质良莠不齐、售后服务效率低、故障定位不明确、同一故障出现多次返修等问题^[4]。

智能故障诊断是为了弥补传统故障诊断方法的不足而出现的新的故障诊断方法^[5], 其中, 基于知识的故障诊断方法是当前研究和应用的热点, 其表现形式多为专家系统、故障树分析、支持向量机、实例推理等^[6], 在这之中, 具有直观、清晰、逻辑性强等特点的故障树分析法是故障诊断领域的主要分析方法之一^[7]。然而传统的故障树分析法建立的单一故障诊断模型已不能满足当前复杂系统的故障诊断需求, 常采取的方法是将故障树与其他算法相结合的故

障诊断方法。柳扬^[8]将故障树最小割集划分与概率神经网络模型结合起来创建了水泵水轮机故障诊断模型; Wang 等^[9]建立了 TCMS 故障的动态故障树(dynamic fault tree, DFT)模型, 并将其转化为贝叶斯网络模型; 陈洪转等^[10]采用模糊集合论的方法确定条件概率等参数, 构建了基于故障树的复杂装备模糊贝叶斯网络推理故障诊断模型; Majdi 等^[11]对各种安全概念和 E/E 体系结构进行建模, 构建了动态故障树; 褚景春等^[12]在构造的多层结构故障树模型基础上引入基于贝叶斯准则的概率神经网络, 构建基于故障树和概率神经网络的风力发电机转速超限故障的智能诊断模型。

在企业长期的生产与服务过程中积累了大量的数据信息, 提取关键数据并进行清洗和分析, 构建汽车故障树能够实现数据的有效利用和清晰地表现出车辆故障间的关系。本文研究基于故障树与贝叶斯网络的故障诊断方法, 快速定位故障发生的位置以及故障原因, 从而提高维修效率。同时在外出维修时, 维修人员在得到准确故障信息后能备齐工具和配件, 从而减少不必要的费用, 提高客户满意度。

收稿日期: 2023-02-19 修稿日期: 2023-03-03

作者简介: *通信作者: 李星(1998—), 女, 四川德阳人, 硕士, 研究方向为云服务平台、供应链管理、故障诊断, E-mail: lx33612022@163.com

1 故障树分析法

故障树是一种逻辑推理图^[13]，是倒立的多叉树结构，由事件和逻辑门构成。事件用来描述系统和元、部件故障的状态，逻辑门把事件联系起来，表示事件之间的逻辑关系。

1.1 故障树定性分析

故障树定性分析阶段主要任务为确定故障树的所有最小割集。

割集是能够诱发顶事件的基本事件的集合，即一个割集代表了系统发生故障的一种可能性，当割集中的基本事件同时出现时，故障必然出现。如果将割集中的基本事件任意删除一个，余下的组合不能使得顶事件发生，此时的割集就是最小割集^[14]。求取最小割集的研究方法较多，最常用的是上行法(Semanderes 算法)和下行法(Fussel-Vesely 算法)^[15]。本文采用下行法，按照遇到与门就增加割集阶数(割集所含基本事件数目)，遇到或门就增加割集个数的规则，从顶事件开始，由上到下逐层寻找并用输入事件取代输出事件进行横向列表，遇到与门就将输入事件横向并列写出，遇到或门就将输入事件在下一列纵向展开，直到故障树的最底层，得到的表格的最后一列的每一行就是故障树的割集，再进行比较和合并消元，剩下的就是故障树的全部最小割集。

1.2 故障树定量分析

故障树定量分析主要任务是顶事件的概率计算和基本事件与最小割集的重要度分析。

故障树的顶事件代表的是系统故障，因此顶事件发生的概率 $P(T)$ 就是系统的不可靠度 $F_s(t)$ ，数学表示式为

$$P(T) = F_s(t) = G[F(t)] \quad (1)$$

公式(1)中， $F(t) = [F_1(t), F_2(t), \dots, F_n(t)]$ ，其中 $F_i(t)$ 表示元、部件不可靠度，即第 i 个基本事件的发生概率 P_i ，有 $F_i(t) = P_i$ 。

在大多数情况下基本事件会重复地出现在几个最小割集中，此时利用最小割集求解顶事件发生概率时就应该采用相容事件的概率公式：

$$\begin{aligned} P(T) &= P(K_1 \cup K_2 \cdots \cup K_{N_k}) \\ &= \sum_{i=1}^{N_k} P(K_i) - \sum_{i < j=2}^{N_k} P(K_i K_j) + \\ &\quad \sum_{i < j < k=3}^{N_k} P(K_i K_j K_k) + \cdots + \\ &\quad (-1)^{N_k-1} P(K_1, K_2, \dots, K_{N_k}) \end{aligned} \quad (2)$$

公式(2)中， K_i, K_j, K_k 表示第 i, j, k 个最小割集， $P(K_i)$ 表示最小割集 K_i 发生概率， N_k 表示最小割集数。

其中最小割集 K_i 发生概率计算公式为

$$P(K_i) = \prod_{x_j \in K_i} P_j \quad (3)$$

公式(3)中， x_j 表示第 j 个基本事件， P_j 表示第 j 个基本事件发生概率。

但在实际的工作中，如果按照公式(2)精确计算故障树顶事件概率会收敛得非常快，并且当最小割集数足够大时，会产生“组合爆炸”问题。可以看出公式(2)共有 $2^{N_k} - 1$ 项，其中起主要作用的是第一项和第二项，所以计算顶事件发生概率的公式可近似表示为

$$P(T) = \sum_{i=1}^{N_k} P(K_i) - \sum_{i < j=2}^{N_k} P(K_i K_j) \quad (4)$$

基本事件的重要度分析包括结构重要度分析、概率重要度分析和关键重要度分析三个方面，其中基本事件关键重要度代表着基本事件的概率变化率对顶事件发生概率的变化率的影响程度，即第 i 个元、部件故障率变化引起的系统故障率的变化率大小，是从敏感度和概率双重视角来衡量各基本事件的重要度标准，比结构重要度和概率重要度更合理更具有实际意义，因此本文在进行基本事件重要度分析时主要进行的是基本事件的关键重要度分析。用 $I_i^{CR}(t)$ 表示基本事件的关键重要度，其计算公式为

$$\begin{aligned} I_i^{CR}(t) &= \lim_{\Delta F_i(t) \rightarrow 0} \left(\frac{\Delta g[F(t)]}{g[F(t)]} \bigg/ \frac{\Delta F_i(t)}{F_i(t)} \right) \\ &= \frac{F_i(t)}{g[F(t)]} \cdot \frac{\partial g[F(t)]}{\partial F_i(t)} = \frac{F_i(t)}{F_s(t)} \cdot \Delta g_i(t) \end{aligned} \quad (5)$$

公式(5)中， $\Delta g_i(t)$ 表示基本事件概率重要度，

计算公式为

$$\Delta g_i(t) = \frac{\partial F_s(t)}{\partial F_i(t)} = \frac{\partial P(T)}{\partial F_i(t)} \quad (6)$$

为了直观地观察最小割集对顶事件发生的贡献程度以确定可能性最大的故障原因集合并与之后贝叶斯网络反向推进进行比较，本文还进行了最小割集重要度的计算，用 I_i^k 表示最小割集重要度，其计算公式为

$$I_i^k = \frac{P(K_i)}{P(T)} \quad (7)$$

2 基于故障树的贝叶斯网络模型

贝叶斯网络以概率论知识为基础^[14]，本质上是一个有向无环图，是决策理论与概率推理相结合的产物，可用符号 $B(G, P)$ 表示，其中 G 表示贝叶斯网络拓扑结构， P 表示贝叶斯网络的概率分布^[16]。

2.1 故障树与贝叶斯网络转化关系

在故障树到贝叶斯网络的映射中，故障树事件转化为贝叶斯网络节点，故障树的逻辑门则与贝叶斯网络的有向边和条件概率相对应^[10,14]。本文只涉及到故障树的与门和或门，因此此处只给出这两类逻辑门的转化关系。

故障树与门与贝叶斯网络的转化关系如图1所示。

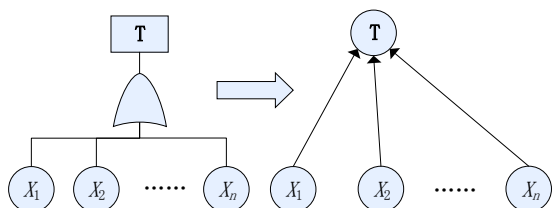


图1 故障树与门与贝叶斯网络的转化

只有所有父节点变量状态均为1时，子节点变量状态为1的概率才会等于1，否则等于0，概率分布为

$$\begin{cases} P(T = 1 | X_1 = 1, X_2 = 1, \dots, X_n = 1) = 1 \\ P(T = 1 | \text{else}) = 0 \end{cases} \quad (8)$$

故障树或门与贝叶斯网络的转化关系如图2所示。

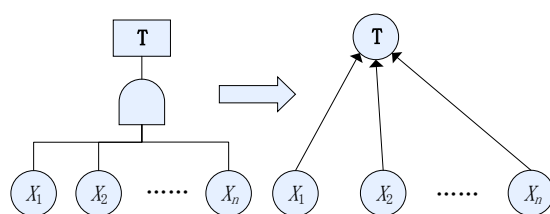


图2 故障树或门与贝叶斯网络的转化

只有所有父节点变量状态均为0时，子节点变量状态为1的概率才会等于0，否则等于1，概率分布为

$$\begin{cases} P(T = 1 | X_1 = 0, X_2 = 0, \dots, X_n = 0) = 0 \\ P(T = 1 | \text{else}) = 1 \end{cases} \quad (9)$$

2.2 贝叶斯网络双向推理

不同于故障树只能单向推理，贝叶斯网络还可以进行双向推理^[17]，即正向推理和反向推理，其中反向推理又被称为诊断推理或者逆向推理，是从故障诊断的角度出发，将网络叶节点(故障树的顶事件)作为证据变量，置于某一固定取值，根据其所处故障状态进行概率更新，计算根节点和其他非根节点的处于各个故障状态的后验概率，这些概率更加符合实际情况。当叶节点 T 处于 T_p 的故障状态下时，根节点 X_j 处于 X_j^i 状态的后验概率计算公式如下：

$$\begin{aligned} P(X_j = X_j^i | T = T_p) &= \frac{P(X_j = X_j^i, T = T_p)}{P(T = T_p)} \\ &= \frac{\sum_{X_1, X_2, \dots, X_n} P(X_1, X_2, \dots, X_j = X_j^i, \dots, X_n, T = T_p)}{P(T = T_p)} \end{aligned} \quad (10)$$

3 实验验证

3.1 故障树模型构建与分析

结合某公司的汽车故障及检修情况，对基于故障树与贝叶斯网络的故障诊断方法进行实验验证。针对“发动机不能启动”故障现象构建初始的故障树模型如图3所示，各编号对应事件名称如表1所示。

利用下行法求解故障树最小割集，可得：

发动机不能启动故障树的一阶最小割集有：

$\{X_1\}$ 、 $\{X_2\}$ 、 $\{X_3\}$ 、 $\{X_4\}$ 、 $\{X_5\}$ 、 $\{X_6\}$ 、 $\{X_7\}$ 、

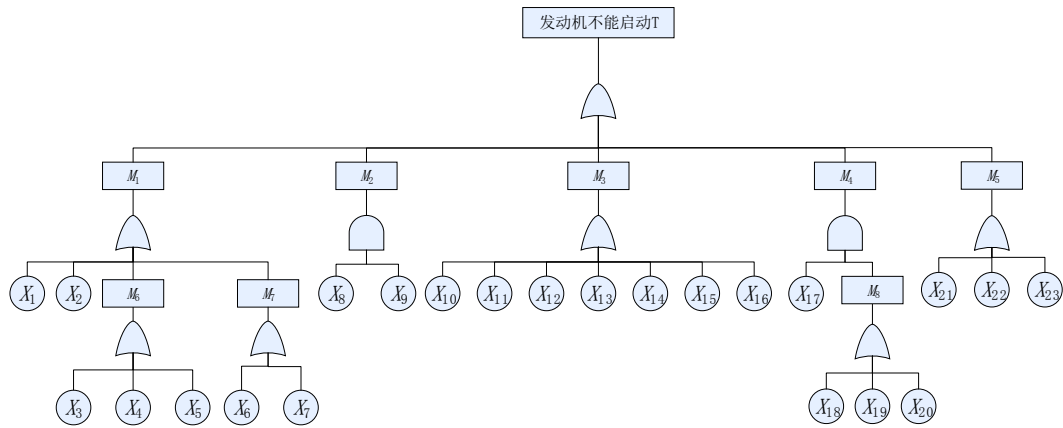


图3 发动机不能启动的故障树模型

$\{X_{10}\}$ 、 $\{X_{11}\}$ 、 $\{X_{12}\}$ 、 $\{X_{13}\}$ 、 $\{X_{14}\}$ 、 $\{X_{15}\}$ 、 $\{X_{16}\}$ 、 $\{X_{21}\}$ 、 $\{X_{22}\}$ 、 $\{X_{23}\}$;

发动机不能启动故障树的二阶最小割集有:

$\{X_8, X_9\}$ 、 $\{X_{17}, X_{18}\}$ 、 $\{X_{17}, X_{19}\}$ 、 $\{X_{17}, X_{20}\}$ 。

分别根据公式(5)和公式(7)计算各基本事件的关键重要度和所有最小割集的重要度,结果从大到小排序。基本事件关键重要度结果如表2所示,最小割集重要度结果如表3所示,结合表1内容可以看出,“燃油滤清器过脏”“喷油嘴堵塞”“蓄电池存电不足”的重要度最高,对顶事件的影响最大。

表1 故障树事件编号与名称对应表

事件编号	事件名称	事件编号	事件名称
M_1	启动系统故障	X_9	点火器故障
M_2	点火系统故障	X_{10}	喷油嘴堵塞
M_3	燃油喷射系统故障	X_{11}	燃油泵不工作
M_4	进气系统故障	X_{12}	泵油压力过低
M_5	ECU故障	X_{13}	燃油管泄漏变形
M_6	蓄电池故障	X_{14}	断路继电器断开
M_7	起动机故障	X_{15}	燃油压力调节器工作不良
M_8	怠速控制阀故障	X_{16}	燃油滤清器过脏
X_1	电路总保险丝断开	X_{17}	空气流量计故障
X_2	点火开关故障	X_{18}	怠速控制阀控制线路故障
X_3	蓄电池存电不足	X_{19}	怠速控制阀空气管破裂
X_4	电极桩柱夹松动	X_{20}	怠速控制阀接头漏气
X_5	电极桩柱氧化严重	X_{21}	固定脚螺栓松动
X_6	起动机线路断路	X_{22}	电子元件焊脚接头松脱
X_7	线路连接器接触不良	X_{23}	电阻、电容原件失效
X_8	点火线圈工作不良		

表2 基本事件关键重要度

序号	基本事件编号	关键重要度	序号	基本事件编号	关键重要度
1	X_{16}	0.0810	13	X_5	0.0211
2	X_{10}	0.0743	14	X_{13}	0.0120
3	X_3	0.0706	15	X_{21}	0.0088
4	X_1	0.0687	16	X_{22}	0.0056
5	X_4	0.0631	17	X_{23}	0.0032
6	X_2	0.0604	18	X_{17}	0.0026
7	X_{15}	0.0469	19	X_8	0.0017
8	X_{12}	0.0355	20	X_9	0.0017
9	X_{14}	0.0321	21	X_{20}	0.0009
10	X_7	0.0304	22	X_{18}	0.0009
11	X_6	0.0287	23	X_{19}	0.0007
12	X_{11}	0.0219			

表3 最小割集重要度

序号	最小割集	重要度	序号	最小割集	重要度
1	$\{X_{16}\}$	0.1664	12	$\{X_{11}\}$	0.0483
2	$\{X_{10}\}$	0.1538	13	$\{X_5\}$	0.0465
3	$\{X_3\}$	0.1467	14	$\{X_{13}\}$	0.0268
4	$\{X_1\}$	0.1431	15	$\{X_{21}\}$	0.0197
5	$\{X_4\}$	0.1324	16	$\{X_{22}\}$	0.0125
6	$\{X_2\}$	0.1270	17	$\{X_{23}\}$	0.0072
7	$\{X_{15}\}$	0.1002	18	$\{X_{17}, X_{20}\}$	0.0021
8	$\{X_{12}\}$	0.0769	19	$\{X_8, X_9\}$	0.0012
9	$\{X_{14}\}$	0.0698	20	$\{X_{17}, X_{19}\}$	0.0003
10	$\{X_7\}$	0.0662	21	$\{X_{17}, X_{18}\}$	0.0001
11	$\{X_6\}$	0.0626			

3.2 基于故障树的贝叶斯网络模型构建与推理

根据图1与图2故障树与贝叶斯网络的转化关系，将图3发动机不能启动的故障树模型中的顶事件、中间事件和基本事件转化为贝叶斯网络节点，并通过有向弧进行连接，形成发动机不能启动的贝叶斯网络模型，如图4所示。

故障树基本事件的发生概率就是贝叶斯网络根节点状态为“1”时的先验概率，如表4所示。而非根节点的先验概率应当是按照公式(8)与公式(9)的逻辑计算得到的条件概率，如表5所示。

表4 贝叶斯网络根节点先验概率

节点编号	$P(X_i = 1)$	$P(X_i = 0)$	节点编号	$P(X_i = 1)$	$P(X_i = 0)$
X_1	0.080	0.920	X_{13}	0.015	0.985
X_2	0.071	0.929	X_{14}	0.039	0.961
X_3	0.082	0.918	X_{15}	0.056	0.944
X_4	0.074	0.926	X_{16}	0.093	0.907
X_5	0.026	0.974	X_{17}	0.042	0.958
X_6	0.035	0.965	X_{18}	0.027	0.973
X_7	0.037	0.963	X_{19}	0.024	0.976
X_8	0.055	0.945	X_{20}	0.028	0.972
X_9	0.038	0.962	X_{21}	0.011	0.989
X_{10}	0.086	0.914	X_{22}	0.007	0.993
X_{11}	0.027	0.973	X_{23}	0.004	0.996
X_{12}	0.043	0.957			

表5 贝叶斯网络非根节点先验概率

节点编号	$P(M_i = 1)$	$P(M_i = 0)$	节点编号	$P(M_i = 1)$	$P(M_i = 0)$
M_1	0.342	0.658	M_5	0.022	0.978
M_2	0.002	0.998	M_6	0.172	0.828
M_3	0.310	0.690	M_7	0.071	0.929
M_4	0.004	0.996	M_8	0.087	0.913

将贝叶斯网络叶节点“发动机不能启动”的概率设置为100%，使其成为证据变量，进行贝叶斯网络的反向推理，完成概率更新，得到贝叶斯网络根节点的后验概率如表6所示，通过此过程得到的概率比先验概率更符合实际情况。

表6 贝叶斯网络根节点后验概率

节点编号	$P(X_i = 1)$	$P(X_i = 0)$	节点编号	$P(X_i = 1)$	$P(X_i = 0)$
X_1	0.143	0.857	X_{13}	0.027	0.973
X_2	0.127	0.873	X_{14}	0.070	0.930
X_3	0.147	0.853	X_{15}	0.100	0.900
X_4	0.132	0.868	X_{16}	0.166	0.834
X_5	0.047	0.953	X_{17}	0.045	0.955
X_6	0.063	0.937	X_{18}	0.027	0.973
X_7	0.066	0.934	X_{19}	0.024	0.976
X_8	0.057	0.943	X_{20}	0.028	0.972
X_9	0.040	0.960	X_{21}	0.020	0.980
X_{10}	0.154	0.846	X_{22}	0.013	0.987
X_{11}	0.048	0.952	X_{23}	0.007	0.993
X_{12}	0.077	0.923			

将更新后的概率重新代入公式(7)求取最小割集的重要度，并从大到小排序，结果如表7所示。和表3中的结果相比，最小割集的后验重要度较先验重要度在数值上普遍有所上升，排序总体变化不大： $\{X_{17}, X_{20}\}$ 下降了三个位次， $\{X_{17}, X_{19}\}$ 上升了两个位次， $\{X_{17}, X_{18}\}$ 上升了一个位次，其余最小割集位次不变，“燃油滤清器过脏 $\{X_{16}\}$ ”“喷油嘴堵塞 $\{X_{10}\}$ ”，“蓄电池存电不足 $\{X_3\}$ ”仍然是排序前三的最小割集，是最容易造成“发动机不能启动”现象发生的故障原因集合。

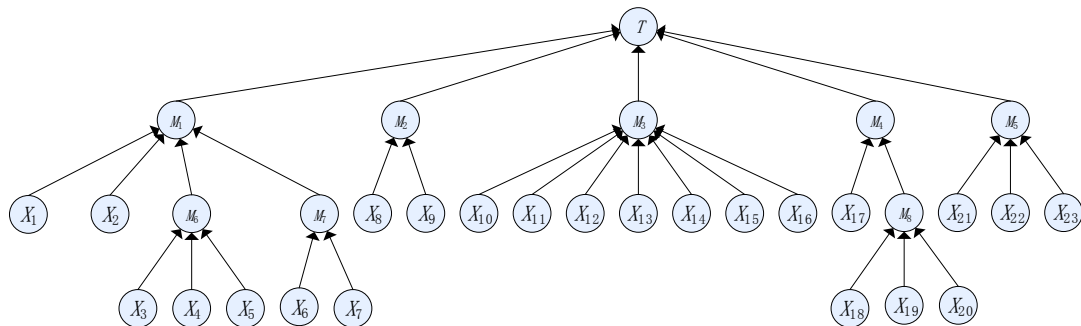


图4 发动机不能启动的贝叶斯网络模型

表7 最小割集后验重要度

序号	最小割集	重要度	序号	最小割集	重要度
1	$\{X_{16}\}$	0.2970	12	$\{X_{11}\}$	0.0859
2	$\{X_{10}\}$	0.2755	13	$\{X_5\}$	0.0841
3	$\{X_3\}$	0.2630	14	$\{X_{13}\}$	0.0483
4	$\{X_1\}$	0.2558	15	$\{X_{21}\}$	0.0358
5	$\{X_4\}$	0.2361	16	$\{X_{22}\}$	0.0233
6	$\{X_2\}$	0.2272	17	$\{X_{23}\}$	0.0125
7	$\{X_{15}\}$	0.1789	18	$\{X_{17}, X_{19}\}$	0.0014
8	$\{X_{12}\}$	0.1377	19	$\{X_8, X_9\}$	0.0010
9	$\{X_{14}\}$	0.1252	20	$\{X_{17}, X_{18}\}$	0.0010
10	$\{X_7\}$	0.1181	21	$\{X_{17}, X_{20}\}$	0.0004
11	$\{X_6\}$	0.1127			

4 结语

本文针对传统故障诊断过程中出现的故障定位时间长、维修效率低等问题研究了基于故障树与贝叶斯网络的故障诊断方法,分析汽车维修服务信息和汽车故障数据,以“发动机不能启动”故障现象作为顶事件建立故障树模型,并根据转化原理将其转化为贝叶斯网络进行反向推理,克服了贝叶斯网络构建困难和故障树只能单向推理的缺点,同时后验概率有效降低参数不确定性,比先验概率更具有实际意义,得到的诊断结果也更为精确。

未来的研究工作将融合更全面的汽车维修服务数据进行模型训练,进一步提高故障诊断的精确性,同时将开展汽车故障诊断系统的开发工作并进行实际应用。

参考文献:

- [1] 工信部装备工业司.《中国制造2025》解读之:推动节能与新能源汽车发展[EB/OL].(2023-02-09).
http://www.gov.cn/zhuanti/2016-05/12/content_5072762.
- [2] 吴志轩.面向汽车快修业的数字化维修辅助系统的设计与实现[D].大连:大连理工大学,2016.
- [3] 杨雯迪.基于数据空间故障文本处理的维修知识服

务技术研究[D].成都:西南交通大学,2021.

- [4] F6汽车科技大数据研究院.2020中国汽车后市场维保行业调研报告[J].汽车维修与保养,2021(2):76-78.
- [5] 肖伟.面向数字化汽车快修的智能故障诊断与维修系统研究[D].大连:大连理工大学,2017.
- [6] 王晓峰,毛德强,冯尚聪.现代故障诊断技术研究综述[J].中国测试,2013,39(6):93-98.
- [7] 吕瑞,孙林夫.基于多源信息融合故障树与模糊Petri网的复杂系统故障诊断方法[J].计算机集成制造系统,2017,23(8):1817-1831.
- [8] 柳场.基于故障树分析的水泵水轮机故障诊断方法研究及应用[D].武汉:华中科技大学,2019.
- [9] WANG C, WANG L, CHEN H, et al. Fault diagnosis of train network control management system based on dynamic fault tree and Bayesian network[J]. IEEE Access, 2021, 9:2618-2632.
- [10] 陈洪转,赵爱佳,李腾蛟,等.基于故障树的复杂装备模糊贝叶斯网络推理故障诊断[J].系统工程与电子技术,2021,43(5):1248-1261.
- [11] MAJDI G, SEBASTIAN J, JOOST-PIETER K, et al. Safety analysis for vehicle guidance systems with dynamic fault trees[J]. Reliability Engineering and System Safety, 2019, 186:37-50.
- [12] 褚景春,王飞,汪杨,等.基于故障树和概率神经网络的风电机组故障诊断方法[J].太阳能学报,2018,39(10):2901-2907.
- [13] 罗航.故障树分析的若干关键问题研究[D].成都:电子科技大学,2011.
- [14] 柳昕汝.山区高速公路伤亡事故故障树及贝叶斯网络模型[D].哈尔滨:哈尔滨工业大学,2019.
- [15] 曾声奎,赵廷弟,张建国,等.系统可靠性设计分析教程[M].北京:北京航空航天大学出版社,2001.
- [16] JUN H B, KIM D. A Bayesian network-based approach for fault analysis[J]. Expert Systems with Applications, 2017, 81:332-348.
- [17] 熊娜,梅天宇,李飞.基于区间贝叶斯网络的复杂系统可靠性分析[J].现代防御技术,2021,49(6):96-103.

(下转第19页)

文章编号: 1007-1423(2023)10-0015-05

DOI: 10.3969/j.issn.1007-1423.2023.10.003

项目申报书形式审查算法

张彬¹, 程健峰^{2*}

(1. 绵阳师范学院传媒学院, 绵阳 621000; 2. 四川大学计算机学院, 成都 610065)

摘要: 根据项目申报书扫描件的文本较为规范的特点, 采用 MSER 算法进行文字内容的检测, 然后利用 PaddleOCR 进行公司名称识别, 实验结果表明召回率达到 91.38%, 精确度为 98.56%, F1 值为 94.83%。采用 CRNN 深度神经网络对印章内容检测, 并根据印章文字分布的空间特征, 在 CNN 网络中采用圆形卷积核, 对比传统的方形卷积核, 识别效果得到一定提升, 召回率为 98.90%, 精确度为 98.39%, F1 值为 98.64%。实验结果表明, 文章提出的算法可以有效进行项目申报书扫描件中涉及印章匹配的形式审查, 对提升机关部门的工作效率具有重要意义。

关键词: 图像处理; 印章识别; 文档检测; 深度学习; 字符识别

0 引言

近年来, 国家发布了不少产业政策以推动高新技术产业和国家战略科技发展, 其中涉及大量的项目申报环节, 尤其为市一级的项目主管部门带来了海量的项目申报书扫描件的审查工作, 其中印章检查是形式审查环节的最主要内容, 传统方式主要由人工挨个查看印章是否缺失, 以及申报单位名称及印章公司名称是否匹配, 效率极为低下, 因此探讨采用机器学习算法来辅助印章检测, 可以极大地提高项目主管部门的工作效率, 具有较好的经济效益和社会效益。

肖进胜等^[1]利用带角度信息的联结文本提议网对印章文字区域进行检测, 并使用贝塞尔拟合文本区域, 最后利用注意力转移机制对文字区域进行识别。林汉斌等^[2]使用 YOLOv3 对文档中的印章进行检测和定位, 利用 DBNet 对印章的文字进行检测, 利用 CRNN 进行文字识别, 实现对文档中的印章高效精确的识别。戴俊峰等^[3]根据印章元素通常成圆环状排列的特点, 对中文印章图像进行极坐标展开, 克服印章文

字方向不统一的问题, 利用 CTPN+CRNN 网络进行文字的检测与识别, 对 CTPN 网络中的文本构造算法进行改进, 实现对印章内容快速准确的识别。

本文提出了一种基于 MSER 和深度学习的印章审查算法^[4-8]。主要采用圆形卷积核的 CRNN 网络, 一方面识别申报书扫描件中特定位置的公司名称, 另一方在公司名称的一定邻域内进行印章检测并识别印章中的公司名称, 然后进行字符匹配, 完成印章审查环节。

1 文档检测与印章检测识别

本文主要是采用 MSER 算法识别申报书扫描件中需要盖章的公司名称, 然后采用改进的 CRNN 网络进行环状印章文字的检测与识别。总体流程如图 1 所示。

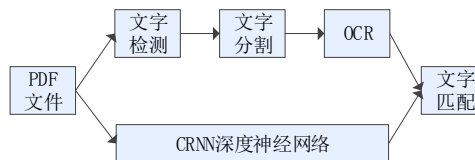


图 1 算法总体流程

收稿日期: 2022-12-09 修稿日期: 2023-02-09

作者简介: 张彬(1975—), 女, 四川绵阳人, 学士, 教师, 研究方向为计算机多媒体技术; *通信作者: 程健峰(1999—), 男, 重庆人, 硕士研究生, 研究方向为算法设计, E-mail: 731590520@qq.com

1.1 基于MSER算法的文字检测

最大稳定极值区域(maximally stable extremal regions, MSER)算法公式如下:

$$q(i) = \frac{|R(i + \Delta) - R(i)|}{|R(i)|} \quad (1)$$

其中: $R(i)$ 表示阈值为 i 时的某一连通区域, Δ 为灰度阈值的微小增加量, $q(i)$ 为阈值是 i 时的区域 $R(i)$ 的变化率, $|R(i)|$ 表示区域 $R(i)$ 的面积。当 $q(i)$ 为局部极小值时, 说明 $R(i)$ 区域的变化非常小, 则 $R(i)$ 可以被认为是最大稳定极值区域。在多重阈值的情况下, 所选区域与背景对比显著, 正是项目申报书扫描件中的文字特征, 也是MSER算法的固有特征。

算法主要思想: MSER对灰度图像取阈值进行二值化处理, 阈值从0到255依次进行递增, 阈值的递增类似于分水岭算法中的水平面上升。具体来说, 图像中灰度值的不同就对应地势高低的不同, 每个阈值都会生成一个二值图。在得到的所有二值图像中, 图像中的某些连通区域变化很小, 甚至没有变化, 则该区域就被称为最大稳定极值区域。在项目申报书扫描件中, 由于文字区域的灰度值是接近一致的, 而且和文字周边像素的灰度值差别较大, 因此在阈值持续增长的一段时间内它们都不会被覆盖, 直到阈值涨到文字本身的灰度值时才会被淹没, 所以文字区域可以作为最大稳定极值区域。

1.2 文字分割

文字检测完成后, 申报书扫描件中的文字区域即被矩形框标记, 此时还须进行文字分割, 减少复杂背景的扰乱。本文采用基于纵横投影的方法进行文字块的切割。先进行横向切割, 横向扫描会得出每一行黑点数量, 如图2所示。

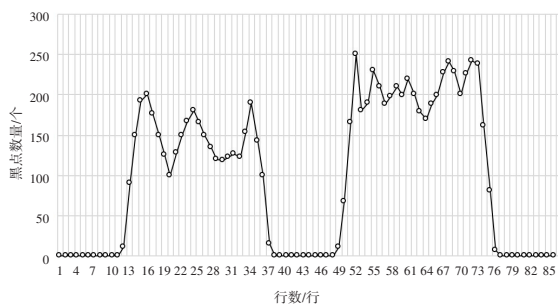


图2 横向切割统计图

在图2中, y 轴数值不为0的区域即文字所在的地方, 取值为0的区域即每行之间的距离。通过以下规则就可以得出每行文字的起点和终点, 用以定位该行文字区域: ①若前一个数为0, 则标记第一个不为0的坐标; ②若前一个数不为0, 则标记第一个为0的坐标。形象地说就是从出现第一个非空白行(开始有字)到出现第一个空白行(没有字)这一段区域就是文字所在的区域。

接着进行纵向扫描。纵向扫描与横向扫描的基本原理相同。针对前述横向扫描出的区域, 再进行纵向扫描, 得出单个字符。同样利用横向扫描的思路, 对纵向扫描得出的数据进行分割, 即可得出单个的文字。

1.3 OCR识别

在确定文字区域后, 即可以使用OCR工具进行文字识别。本文采用开源OCR引擎PaddleOCR, 内部处理过程如图3所示。

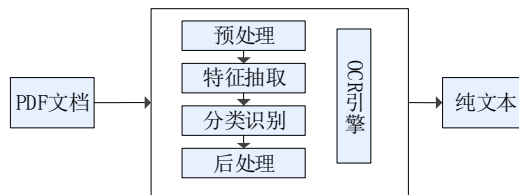


图3 内部处理过程

PaddleOCR是通用型工具, 针对项目申报书特点, 本文采用了基于radon变换的图像矫正、对比度增强等预处理技术以提高准确率。

1.4 印章检测与公司名称识别

对于项目申报书, 正文内容的格式较为规范, 因此用常见的OCR工具即可较为准确地完成识别任务, 但是对于印章的检测和识别, 可能存在印章部分模糊、与正文内容重叠、角度倾斜等问题, 用传统的OCR工具识别准确性较差, 因此本文采用CRNN深度神经网络来实现端到端的印章检测与公司名称识别。CRNN网络结合了CNN与RNN, 将特征提取、序列预测和转录结合到一个网络结构中, 从而能够进行端到端的训练和预测。

CRNN 网络包括三部分：特征提取网络、序列预测网络和转录层。其中特征提取网络由卷积神经网络构成，主要作用是从图片中提取高维特征；序列预测网络由双层 BiLSTM 网络组成，主要功能是利用循环神经网络对特征提取网络输出的特征序列和标签分布间的关系进行学习，得到初步的预测结果；转录层则主要利用自然语言处理 CTC 模块，整合初频的预测结果，得出最终的预测结果，如图 4 所示。

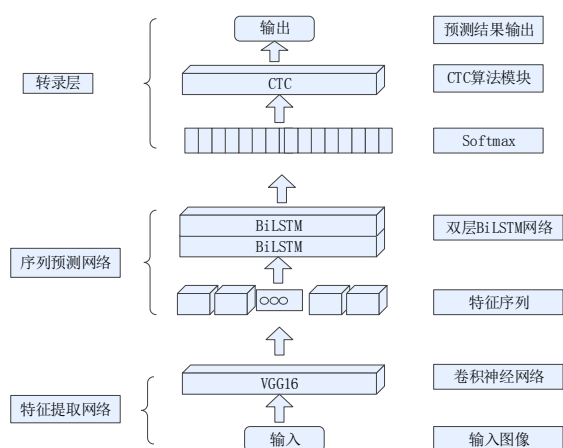


图 4 网络结构

考虑到项目申报书里公司印章通常为圆形的特征，本文中的卷积神经网络采用圆形卷积核：

$$O_j = \sum_{s \in S, r \in R} W_s \left(\sum_{s \in S} B(s, j+r) I_s \right) \quad (2)$$

由于圆形卷积核的接受场不完全是整数的位置，所以需要双线性插值的方法来计算相应位置的特征值，因此圆形卷积核的计算量要大于传统的方形卷积核。

转录层的 CTC 模块即为模型对应的损失函数：

$$L(S) = \sum_{(x,y) \in D} -\log(p(y|x)) \quad (3)$$

式(3)中， x 为输入序列； y 为输出序列； D 为训练集； p 为后验概率。为了降低计算复杂度，需使用动态规划前向传播算法。在预测阶段，CRNN 要找到最大概率对应的输出，即：

$$y^* = \arg \max_y p(y|x) \quad (4)$$

1.5 文字匹配

识别出的印章公司名称需要与项目申报书扫描件中邻近区域内识别的公司名称匹配，确保印章的准确有效。本文采用基于文本特征的相似度计算方法，主要流程如下：

(1) 将文本转换为 feature vectors。使用 TF-IDF 得到 feature vectors，向量维度为词典大小，向量的每一维是词典中该位置的词在文本计算的 TF-IDF 值。

(2) 利用 feature vectors 计算文本间的相似度。基于两个文本的特征向量，使用余弦相似度计算他们的相似度：

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (5)$$

2 实验结果与分析

本文作者将历年来存档的 356 份项目申报书扫描件分成两组，第一组 300 份用于训练，第二组 56 份用于测试。

2.1 检测效果评价指标

本文采用 DetEval^[9] 评价方法来对实验效果进行分析判断。精确度、召回率、F1-score 公式依次为

$$\begin{cases} R(G_i, D_j) = \frac{Area(G_i \cap D_j)}{Area(G_i)} \\ P(G_i, D_j) = \frac{Area(G_i \cap D_j)}{Area(D_j)} \\ F(G_i, D_j) = \frac{2 \times Area(G_i \cap D_j)}{Area(G_i) + Area(D_j)} \end{cases} \quad (6)$$

式中： D 表示算法输出的 bbox 集合； G 表示 Ground Truth bbox 集合。

2.2 申报书公司名称检测效果

因为本文要解决的问题较为特殊，待检测文本有关键词后缀“公司”，因此仅需要在文字检测算法中加入“公司”之后的紧邻位置无其它字符的约束条件即可，检测效果较好。

根据实验结果，在申报书扫描件出现模糊不清、页面倾斜等问题时，依然能较好地检测

出公司名称。定量分析来看,在测试集上本文算法的召回率达到91.38%,精确度达到98.56%,F1指标为94.83%,另外本文还对比了另外两款OCR工具:EasyOCR、ChineseOCR,如表1所示。

表1 OCR性能对比

Algorithm	Recall/%	Precision/%	F1-score/%
EasyOCR	92.01	92.59	92.30
Chineseocr	92.19	96.76	94.42
PaddleOCR	91.38	98.56	94.83

从结果来看,PaddleOCR效果更好。而且对本文要解决的问题来说,精确度的重要性更高。

2.3 印章公司名称检测效果

本文对比了传统方形卷积核与圆形卷积核的印章公司名称识别效果,如表2所示。

表2 不同卷积核性能对比

Type of convolution kernel	Recall/%	Precision/%	F1-score/%
Square	93.71	95.69	94.69
Circular	98.90	98.39	98.64

实验结果表明,圆形卷积核因为插值计算导致计算量适当增大,但是召回率和精确度更高。最终的识别效果如图5所示。



图5 识别效果

3 结语

本文根据项目申报书扫描件的文本特点,以及印章公司名称的空间特征,对文本内容、印章内容检测与匹配算法进行了对比研究和一

定的优化,尤其是圆形卷积核的采用,取得了相对较好的效果。但是针对项目申报书中没有出现公司名称但仍需盖章的情况,以及骑缝章的形式审查要求,还未找到相应的解决方案,可作为下一步的研究方向。

参考文献:

- [1] 肖进胜,赵陶,熊闻心,等.角度优化网络的印章文字检测与识别算法[J].电子与信息学报,2021,43(11):3327-3334.
- [2] 林汉斌,陈元宝,王慧冰.基于目标检测和图像处理的印章识别[J].IT经理世界,2021(12):73-74.
- [3] 戴俊峰,杨天,熊闻心.基于极坐标转换的中文印章文字识别[J].计算机工程与设计,2021,42(11):3174-3180.
- [4] LIN H, YANG P, ZHANG F. Review of scene text detection and recognition[J]. Archives of Computational Methods in Engineering, 2020, 27 (2) : 433-454.
- [5] WANG Q, HUANG Y, JIA W, et al. ConvLSTM with focused attention for scene text recognition[J]. Science China Information Sciences, 2020, 63 (2) : 31-44.
- [6] SHI B, BAI X, YAO C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition[J]. Pattern Analysis & Machine Intelligence, 2017, 39(11): 2298-2304.
- [7] SU B, LU S. Accurate recognition of words in scenes without character segmentation using recurrent neural network[J]. Pattern Recognition: the Journal of the Pattern Recognition Society, 2017, 63:397-405.
- [8] LONG S, HE X, YAO C. Scene text detection and recognition: the deep learning era[J]. International Journal of Computer Vision, 2020, 129(1):161-184.
- [9] KARATZAS D, SHAFAIT F, UCHIDA S, et al. ICDAR 2013 robust reading competition[C]//Proceedings of the 12th International Conference on Document Analysis and Recognition, Washington, DC, USA, 2013:1484-1493.

Formal examination algorithm for project application materials

Zhang Bin¹, Cheng Jianfeng^{2*}

(1. School of Communication, Mianyang Normal University, Mianyang 621000, China;

2. School of Computer Science, Sichuan University, Chengdu 610065, China)

Abstract: In view of the relatively standardized characteristics of the project application materials, the MSER algorithm to detect and recognize characters, then the company name is recognized by PaddleOCR. The experimental results show that the recall rate can reach 91.38%, the accuracy is 98.56%, F-measure is 94.83%. In this paper, CRNN deep neural network is used to detect the seal content. Given the spatial characteristics of seal text distribution, circular convolution kernel is used in CNN network, the recognition effect of which is better than the traditional square convolution kernel. According to the experimental results, the recall rate can reach 98.90%, the accuracy is 98.39%, F-measure is 98.64%. The experimental results show that the algorithm proposed in this paper can effectively carry out the formal examination involving seal matching in the scanned documents of project application, and is of great significance to improve the work efficiency of the government department.

Keywords: image processing; seal recognition; document detection; deep neural network; OCR

(上接第14页)

Automotive fault diagnosis technology based on fault tree and Bayesian network

Li Xing*

(Sichuan Provincial Key Laboratory of Manufacturing Industry Chain Collaboration and Information Support Technology,

Southwest Jiaotong University, Chengdu 611756, China)

Abstract: Aiming at the problem that the cause of the fault cannot be efficiently and quickly located in the traditional automobile fault diagnosis process, resulting in the increase of maintenance cost and maintenance time, after studying the automobile fault diagnosis method based on fault tree and Bayesian network, the automobile fault data is analyzed, the fault tree model is constructed and converted into Bayesian network, Bayesian network reverse reasoning is carried out, the fault diagnosis sequence is determined, the accuracy of fault diagnosis is improved, the service demand is quickly responded, and the maintenance personnel are helped to quickly locate the location of the fault and the cause of the fault. So as to improve maintenance efficiency and improve service quality.

Keywords: fault diagnosis; fault tree analysis; Bayesian networks

文章编号: 1007-1423(2023)10-0020-05

DOI: 10.3969/j.issn.1007-1423.2023.10.004

基于层聚合机制的密集场景人群计数方法

席梦园, 严 华*

(四川大学电子信息学院, 成都 610064)

摘要: 人群计数是计算机视觉中的一项基本而重要的任务, 旨在估计图像中的人群数量, 但计数精度受到剧烈尺度变化的影响。提出了一个基于层聚合机制的人群计数方法, 通过聚合层间特征获得密集场景中的人群多尺度信息。在四个公共人群计数数据集进行的大量实验证明, 该方法能够很好地在密集场景中进行准确的人群计数。

关键词: 人群计数; 多尺度; 层聚合机制

0 引言

人群计数任务的主要目标是预测图像或视频中的人数。因为它在如视频监控、交通控制、公共安全、城市规划、体育事件、疫情防控、细胞计数等现实场景中有广泛的应用价值, 这一领域的研究最近受到了极大的关注。然而, 由于不同人群场景中的巨大尺度变化, 准确的人群计数依然十分具有挑战性。具体地说, 单幅图像内的不同区域以及图像间乃至不同数据集间的人头尺度都存在十分大的差异。因此, 提高人群计数准确性的关键方法是如何更好地提取多尺度特征。传统的基于检测的方法主要适用于比较稀疏的人群计数。拥挤场景下的人群密度大, 人之间的遮挡非常严重, 所以计数效果较差。基于回归的方法^[1]通过提取低级特征和学习回归模型, 来映射特征到人群数量的关系。但这种方式只能回归人群的数量关系, 无法反映人群空间上的密度分布状态。深度学习凭借优秀的特征学习能力, 被广泛应用于计算机视觉任务中。基于深度学习的人群计数方法利用预测密度图的方法得到了更好的预测结果, 但图像中人群的剧烈尺度变化依然影响了计数的准确性。

本文提出了一种基于层聚合机制的人群计数方法, 通过逐步聚合分层特征获得人群多尺度信息, 其包括主干网络、五个堆叠的层聚合模块和由七个卷积层组成的后端网络。具体而言, 我们采用预训练的 VGG-16^[2]的前十个卷积层作为生成低级特征的主干网络, 然后, 用 1×1 的卷积处理这些特征来压缩特征的通道数, 再传递给五个堆叠的层聚合模块。层聚合模块有三个卷积层, 通过一系列聚合操作、通道注意力模块和残差连接来提取多尺度特征, 后端网络最后预测出人群分布的密度图。与其它方法相比, 本文所提出的方法具有较准确的人群计数结果。

1 算法简介

即使在深度卷积神经网络中, 利用单层的特征图生成密度图也是不够的。我们提出的方法通过聚合和重用 CNN 中固有的层次特征, 降低计数误差高质量密度图。本节首先介绍所提方法的总体框架, 然后介绍所提出的层聚合机制的详细结构。

1.1 算法总体框架

本文提出的方法如图 1 所示, 主要包括三

收稿日期: 2022-12-05 修稿日期: 2023-02-07

作者简介: 席梦园(1997—), 男, 四川南充人, 硕士, 研究方向为机器学习、计算机视觉与模式识别; *通信作者: 严华(1971—), 男, 四川渠县人, 博士, 教授, 研究方向为智能信息系统, E-mail: yanhua@scu.edu.cn

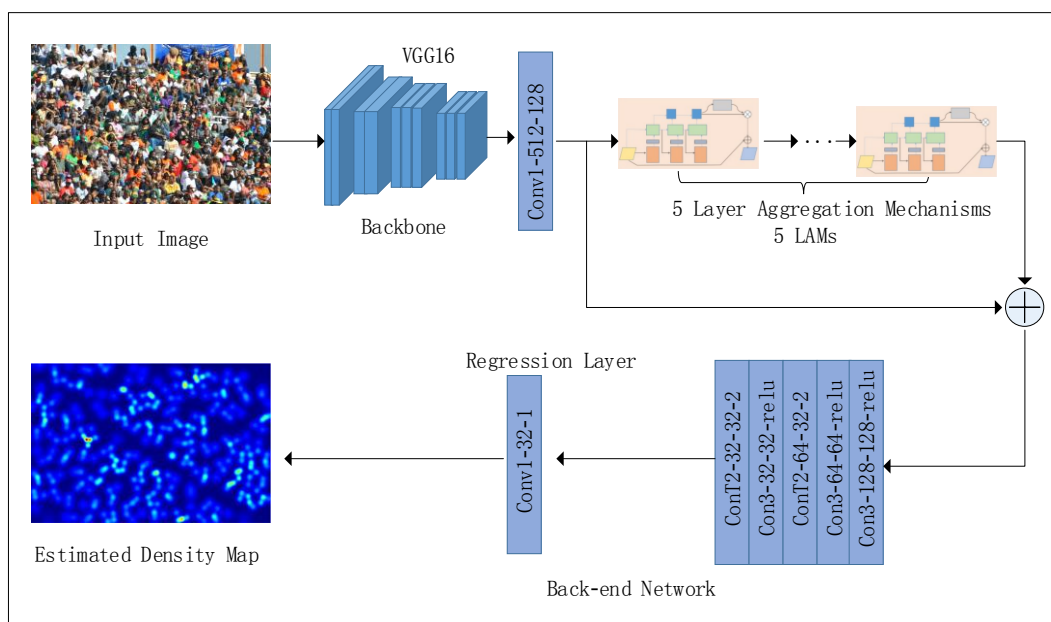


图1 本文算法流程

个部分：作为低级特征提取器的主干网络，五个层聚合机制聚合不同层的特征，以及由七个卷积层组成的后端网络，用于回归密度图。输入图像中的总人数可以通过生成的密度图像素求和得到。遵循现有方法的思想，我们采用了VGG-16的前10层卷积层作为我们方法的骨干，其强大的迁移学习及其扩展的灵活性能很好地应用在人群计数任务中。VGG-16的所有卷积滤波器都采用 3×3 的小卷积核。通过小卷积核的堆叠，我们可以获得与大卷积核同样大小的感受野（即，两个堆叠的 3×3 的卷积与一个 5×5 的卷积具有相同的感受野），并且具有更少的参数。同时，更多的非线性层可以加深网络并学习更复杂的表示。在经过三个池化层后，主干网络输出的特征图大小是原始输入图像的 $1/8$ 。为了防止增大网络复杂度，我们首先通过 1×1 卷积压缩特征图的通道数，然后传递给依次堆叠的五个层聚合机制以提取多尺度特征和丰富的层间信息。后端网络中采用了四个卷积层和三个转置卷积来回归最终的预测密度图。其中，三个转置卷积层能将密度图恢复到与输入图像相同大小的分辨率。人群的数量可以在密度图上求像素之和得到。

1.2 层聚合机制

层聚合机制主要由三个空洞卷积以及一个通道聚合模块组成。在此基础上，迭代的特征融合旨在进一步挖掘隐藏在原始网络中的信息。层聚合机制的细节结构如图2所示。在深度卷积网络中即使存在高级的语义特征，但缺乏多尺度信息的提取，根据最终单层卷积层输出的特征图来预测密度图的准确性是有限的。在神经网络中，研究者们广泛认为浅层特征可以提供丰富的空间信息，而深层提取的高级特征包含抽象的语义信息。因此聚合这些神经网络中不同深度的层可以获取丰富的人群特征信息，包含多尺度信息，从而改进模型对人群位置和人群密度的预测。通过聚合具有不同感受野的不同深度的层提取人群的多尺度信息，而聚合通道间的特征可以更加关注到对密度图预测重要的通道^[3-4]。

本文所提出的层聚合机制通过这两种聚合方式获得丰富的多尺度信息，从而在较拥挤的人群场景中也能准确地预测出人群密度图。如图2所示，输入特征首先通过空洞卷积^[5]和非线性激活函数。然后，通过跳跃连接，融合层间特征。聚合中间特征后，把聚合后的特征送入下一个卷积作为输入，如图2所示。这将传播

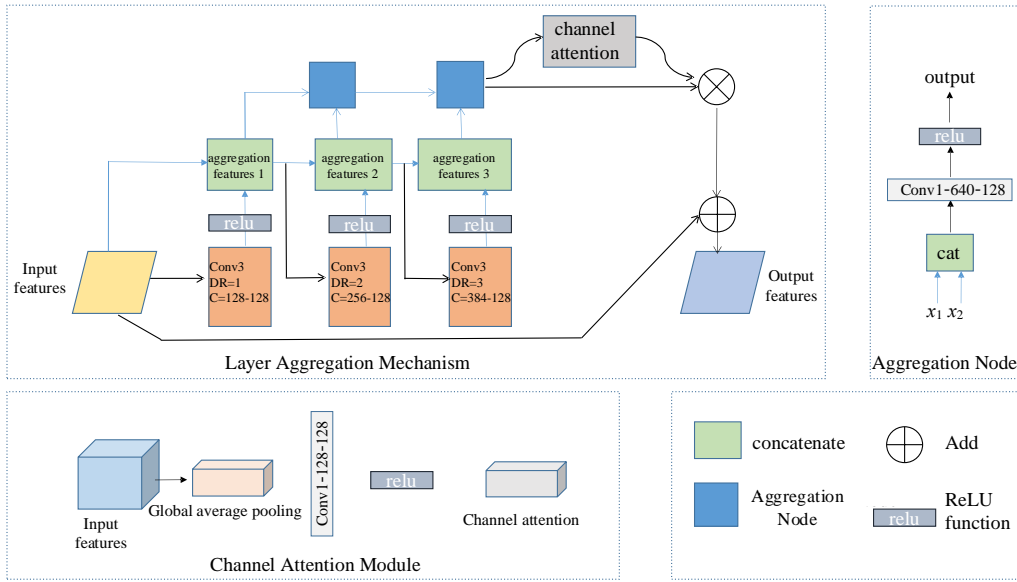


图2 层聚合机制结构

所有先前块的聚合特征，便于更好地保存信息。跨层聚合 P 的过程可表示为如下公式：

$$P: \begin{cases} b_1 = \text{cat}(\sigma(c_1(f_i)), f_i) \\ b_2 = \text{cat}(\sigma(c_2(b_1)), b_1) \\ b_3 = \text{cat}(\sigma(c_3(b_2)), b_2) \end{cases} \quad (1)$$

其中： c_1, c_2, c_3 表示空洞率分别为 $\{1, 2, 3\}$ 的空洞卷积； σ 表示 ReLU 函数， cat 表示通道拼接操作； f_i 是主干网络提取的低级特征压缩通道数后的特征，作为第一个聚合块的输入； b_1, b_2, b_3 为聚合块的输出，分别送入后续融合节点进一步迭代融合。如图2所示，在逐步聚合不同层之后，引入融合节点来对这些聚合特征进行迭代重用。为了避免融合节点的复杂度开销，所有节点使用 1×1 的卷积将不同输入信道通道数转换为与层聚合机制的输入通道数相同的128，可以降低计算复杂性并便于后续连接。融合节点的主要功能是组合和压缩其输入特征。该过程描述如下：

$$F_{fn}(x_1, x_2) = \sigma(\text{conv}(\text{cat}(x_1, x_2))) \quad (2)$$

其中： x_1, x_2 是融合节点的输入， conv 是一个 1×1 卷积，用于改变通道数量。然后，融合节点学习选择和映射重要特征。最后，将这些融合特征发送到通道聚合模块。不同通道的特征图对人群密度的估计具有不同的重要性。通道聚合

模块可以根据每个通道的重要性给相应通道更高或更低的注意力权重，有助于精确的密度估计。如图2所示，通道聚合模块包括一个全局平均池化层，一个 1×1 卷积和一个 ReLU 激活函数。我们先通过全局平均池化将特征图转换为一维向量。 1×1 卷积和 ReLU 函数用于对不同通道的重要性进行学习并产生相应的注意力权重，然后应用于输入特征图。整个通道聚合过程可总结为

$$C(x) = \sigma(\text{Conv}(\text{GAP}(x))) \quad (3)$$

$$F_x = C(x) \otimes x \quad (4)$$

其中： x 是通道聚合模块的输入， GAP 代表全局平均池化层； conv 是一个 1×1 卷积， σ 表示 ReLU 函数； $C(x)$ 代表学习到的各通道权重， F_x 是经过通道聚合模块处理后的输出特征。最后，将层聚合机制的输入特征与 F_x 进行长距离的跳跃连接^[6]，得到层聚合机制最终的输出。

2 实验结果与分析

2.1 实验设置

在每个图像的不同位置裁剪出8个具有1/4大小的图像块。并且通过水平翻转来进行数据增强。本文的人群计数方法以端到端的方式进行训练。主干网络参数由VGG-16在ImageNet上预训练的结果初始化得到，其余可训练参数由

平均值为0，标准差为0.01的高斯分布随机初始化。遵照以往的方法^[7]，采用欧几里得损失函数衡量模型预测结果与标注数据的差距，使用固定学习率为0.00001的RMSprop优化器，以训练本文所提出的人群计数网络。采用基于几何自适应的高斯核得到训练密度图，我们在ShanghaiTech Part A/B, UCF-QNRF, UCF_CC_50四个公开数据集上进行了训练，数据集的统计信息如表1所示，其中图像数量表示为训练集+测试集；Min, Max, Ave, Total分别表示数据集中图像最小，最大，平均和总的人群数量。

2.2 实验结果

我们使用平均绝对误差(MAE)和均方根误

差(RMSE)作为评估指标。在四个人群计数公开数据集(ShanghaiTech Part A/B、UCF-QNRF、UCF_CC_50)上对所提出方法的有效性进行了评估，并与其他最新方法进行了比较。结果如表2所示，其中加粗数据代表最好的结果。在ShanghaiTech Part A和Part B部分，与最新的方法相比，本文方法达到了最佳的MAE，分别为56.5和7.0。在UCF-QNRF数据集上，我们还获得了优于其他方法的性能。UCF_CC_50是一个大规模的拥挤场景数据集，在此数据集中，本文方法预测误差比在其它数据集中大，但仍然凭借优秀的多尺度特征学习能力，实现了最高的计数精度。

表1 数据集概述

数据集	分辨率	图像数量	Min	Max	Ave	Total
ShanghaiTech Part A	不同	300+182	33	3139	501.4	241677
ShanghaiTech Part B	768×1024	400+316	9	578	123.6	88488
UCF-QNRF	不同	1201+334	49	12865	815	1251642
UCF_CC_50	不同	50	94	4543	1279.5	63974

表2 与其它最新方法在四个数据集上的结果对比

方法	ShanghaiTech Part A		ShanghaiTech Part B		UCF_CC_50		UCF-QNRF	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
MCNN ^[8]	110.2	173.2	26.4	41.3	377.6	509.1	277.0	426.0
Switch-CNN ^[9]	90.4	135.0	21.6	33.4	318.1	439.2	228.0	445.0
CP-CNN ^[10]	73.6	106.4	20.1	30.1	298.8	320.9	—	—
CSRNet ^[5]	68.2	115.0	10.6	16.0	266.1	397.5	—	—
SANet ^[11]	67.0	104.5	8.4	13.6	258.4	334.9	—	—
FF-CAM ^[4]	71.0	109.8	10.3	15.8	246.8	322.2	114.5	200.5
MFANet ^[12]	58.5	93.4	10.9	17.5	221.3	279.7	97.7	166.0
本文方法	56.5	94.5	7.0	10.7	156.1	186.6	96.9	163.4

3 结语

针对人群计数中人群剧烈尺度变化的问题，提出了一个基于层聚合机制的方法来提取多尺度特征，从而进行精确的人群计数。大量实验表明，本文方法在四个公开数据集上实现了极佳的性能。有效地聚合了层次特征，在未来的工作中，我们将探索并研究相应的轻量级网络设计。

参考文献：

- [1] 周成博,陶青川. 基于景区场景下的人群计数[J]. 现代计算机,2016(2):52-57.
- [2] SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. arXiv:1409.1556,2014.
- [3] 张菁,陈庆奎. 基于注意力机制的狭小空间人群拥挤度分析[J]. 计算机工程,2020,46(09):254-260,267.
- [4] 张宇倩,李国辉,雷军,等. FF-CAM:基于通道注意

- 机制前后端融合的人群计数[J]. 计算机学报, 2021 (2): 304-317.
- [5] LI Y, ZHANG X, CHEN D. CSRNet: dilated convolutional neural networks for understanding the highly congested scenes [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City: IEEE, 2018: 1091-1100.
- [6] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks [C] // Proceedings of the 14th European Conference, Amsterdam, The Netherlands, 2016: 630-645.
- [7] 付宇豪. 基于卷积神经网络的人群计数算法研究与应用[D]. 北京: 北京工业大学, 2019.
- [8] ZHANG Y, ZHOU D, CHEN S, et al. Single-image crowd counting via multi-column convolutional neural network [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas: IEEE, 2016: 589-597.
- [9] BABU SAM D, SURYA S, VENKATESH BABU R. Switching convolutional neural network for crowd counting [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii: IEEE, 2017: 5744-5752.
- [10] SINDAGI V A, PATEL V M. Generating high-quality crowd density maps using contextual pyramid CNNs [C] // Proceedings of the IEEE International Conference on Computer Vision, Venice: IEEE, 2017: 1861-1870.
- [11] CAO X, WANG Z, ZHAO Y, et al. Scale aggregation network for accurate and efficient crowd counting [C] // Proceedings of the European Conference on Computer Vision (ECCV), Munich: Springer, 2018: 734-750.
- [12] ZHU F, YAN H, CHEN X, et al. A multi-scale and multi-level feature aggregation network for crowd counting [J]. Neurocomputing, 2021, 423 (11): 46-56.

Crowd counting in dense scenes via layer aggregation mechanism

Xi Mengyuan, Yan Hua*

(College of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China)

Abstract: Crowd counting is a basic and important task in computer vision, which aims to estimate the number of people in images. However, the counting accuracy is affected by dramatic scale variation. In this paper, we propose a crowd counting method based on layer aggregation mechanism, which can obtain multi-scale information of crowd in dense scenes by aggregating inter-layer features. Extensive experiments carried out on four public crowd counting datasets demonstrate that the proposed method can accurately count people in dense scenes.

Keywords: crowd counting; multi-scale; layer aggregation mechanism

文章编号: 1007-1423(2023)10-0025-05

DOI: 10.3969/j.issn.1007-1423.2023.10.005

基于 Stacking 的个人薪资预测研究

林心慧, 陈超*

(四川轻化工大学计算机科学与工程学院, 宜宾 644000)

摘要: 薪水是求职者着重考虑的因素, 在信息飞速发展的时代, 求职者寻找工作时希望通过各种方法去合理地评估自己的薪酬区间, 而人们往往难以对自己的薪资水平进行合理预测, 因此合理评估能力, 进而找到适合自己的薪资标准就显得格外重要。基于这个现实问题, 通过对原始数据进行分析处理并结合 Stacking 算法, 即融合随机森林、梯度提升树、XGBoost、逻辑回归等模型, 对求职者的薪水进行一个合理的预测, 并且通过实验显示, 使用 Stacking 融合算法的模型具有更好的预测性能和泛化能力。

关键词: Stacking 融合模型; 个人薪水; 随机森林; 梯度提升树; XGBoost

0 引言

在互联网时代, 人们越来越倾向于通过互联网寻找工作机会, 而薪水是求职者着重考虑的因素, “凡事预则立, 不预则废” 这条原则也同样适用于挣扎于招聘市场的求职者们。但由于个人所接受信息的限制, 人们往往难以预测出自己可以争取的薪资范围, 低估或者高估自身能力, 进而导致在求职过程中遇到层层阻碍。因此, 建立高精度的薪水预测模型帮助求职者合理评估自身实力, 帮助雇主了解不同职位的市场价值, 通过薪水预测模型建立薪水预测页面, 招聘网站可以为求职者提供些许帮助, 并让用户更了解自身定位, 进而更加频繁地使用该招聘网站^[1-4]。由此可知, 薪水预测模型可以为求职者和雇主以及招聘平台三方提供便利。

近年来, 很多学者对个人薪水预测的研究主要包括基于统计学和基于机器学习的预测方法。潘博等^[4]使用 doc2vec 模型利用互联网招聘广告中的非结构化文本信息建立薪水预测模型。Khongchai 等^[5]使用 KNN 等模型对毕业生的薪水进行预测, 以此来激发学生学习动力。Martín 等^[6]使用基于决策树的集成模型对 IT 行业从业

人员的个人薪水进行预测。Jerrim^[7]使用线性回归模型预测学生薪水和家庭背景的关系。Hung 等^[8]提出了“公司、职业、公司”(COC)模型, 通过汇总工作审查和职位发布数据可以有效地预测相应工作岗位的工资、公司的通货膨胀和竞争力。

上述研究在对薪水的预测中多数使用的是单模型预测, 或者是对单一模型进行叠加。模型的叠加使得模型的预测性能有一定的提升, 为了进一步提升模型的预测性能, 本文将随机森林、梯度提升树、极限梯度提升树三个模型参与 Stacking 算法集成, 并使用集成后的模型对个人薪资水平进行预测, 最终通过实验证明 Stacking 多算法集成模型可以提高预测准确性, 拥有更好的预测效果。

1 Stacking 模型融合

1.1 模型介绍

Stacking 算法是模型融合常用的方法, 整个模型分为 Level 0 和 Level 1 两层, Level 0 层由多个基学习器组成, Level 1 层为元学习器。具体模型结构如图 1 所示。在 Stacking 中每一个参与

收稿日期: 2023-01-15 修稿日期: 2023-02-27

作者简介: 林心慧(1999—), 女, 硕士研究生, 研究方向为数据挖掘、数据可视化; *通信作者: 陈超(1980—), 男, 教授, 博士, 研究方向为高性能计算、云计算与大数据技术和智能信息处理, E-mail: chencao@suse.edu.cn

Stacking 模型融合的算法模型都为一个基学习器，放在了 Stacking 模型的 Level 0 层，并且根据 Level 0 层中每一个基学习器的输出得到一个新的特征输入，这个新的特征中每一列数据代表一个基学习器的输出结果，元学习器会对得到的新特征继续进行训练，使用 Stacking 模型融合方法最终输出的预测结果就是 Level 1 层中元学习器输出的预测结果^[9]。

Stacking 模型融合在解决回归问题的时候，Level 0 层的学习器要满足复杂度高、学习能力强的要求，如梯度提升树、XGboost、随机森林等，这样才可以更好地找出原始数据与标签的关系，即建立原始数据与标签之间的假设^[10-11]。Level 1 上的学习器采用的是可解释性强、较为简单的模型，如决策树、线性回归、逻辑回归等，因为 Level 1 上的算法的职责是融合各个基学习器所做出的假设，并最终输出融合模型的结果，相当于在寻找“最佳融合规则”，而非直接建立原始数据与标签之间的假设。Stacking 算法的本质就是让元学习器应用的算法找出不同基学习器之间的融合规则，是通过训练元学习器来融合不同基学习器的结果，进而得出更好的预测结果。

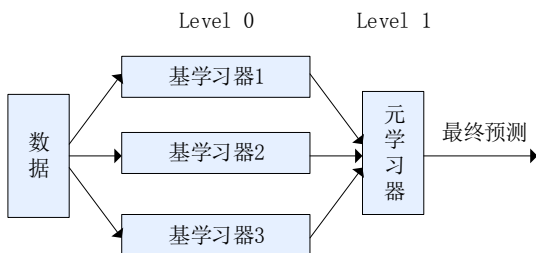


图 1 Stacking 模型结构

1.2 Stacking 融合训练流程

本文也采用 Stacking 模型融合思想，其中基学习器为我们提前筛选模型，元学习器为线性回归模型，以此构造一个 Stacking 集成模型，并用其对个人薪水进行预测。其主要想法是希望使用 Stacking 算法将选择的多个算法模型进行融合，进而构造一个新的融合模型，即 Stacking 集成模型，并希望 Stacking 集成模型能够吸收单一算法的优势，模型表现得更好，拥有更强大的预测能力，可以提高对个人薪水预测的准确性。

在基学习器训练完毕后，Stacking 模型开始进行融合训练，步骤如下：

(1) 将原始数据划分为训练集 M_{train} 和测试集 M_{test} 。

(2) 将训练集输入 Level 0 的个体学习器，分别在每个基学习器上进行交叉验证。在每个基学习器上，将所有交叉验证的验证结果纵向堆叠形成预测结果。假设预测结果为概率值，当融合模型执行回归或二分类任务时，该预测结果的结构为 $(M_{train}, 1)$ ，

(3) 要使用全部训练集重新对每一个个体学习器进行 fit 训练。只有个体学习器训练好了，后面测试的时候才可以用于预测。

(4) 将所有个体学习器的预测结果横向拼接，形成新特征矩阵。假设共有 N 个个体学习器，则新特征矩阵的结构 (M_{train}, N) ，并通过元学习器对其训练。具体流程如图 2 所示。

经过上面步骤，便得到了 Stacking 集成模型，后续将使用该模型对个人薪水进行预测，求职者可以使用该模型得到属于自己的薪资数据，进而帮助求职者合理地评估其薪水。

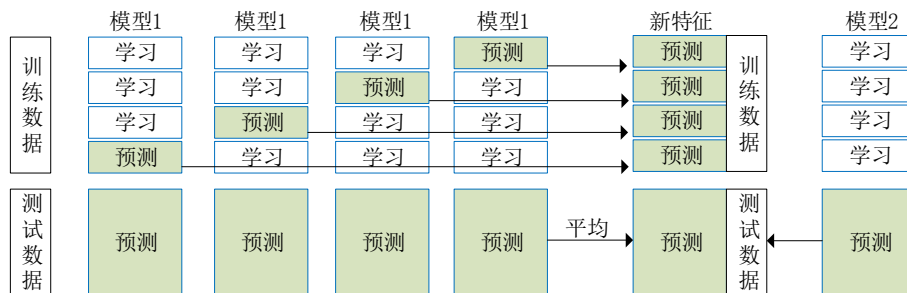


图 2 Stacking 融合训练流程

2 实验

2.1 数据描述

本论文使用的数据集选自 kaggle 官网公开数据集“Data Science and STEM Salaries”。来自顶级公司的 62,000 条薪资记录，这些数据是从 levels.fyi 中获得的。数据来源真实而且维度丰富，对数据分析起到了至关重要的作用。

该数据集记录了被调查者的性别(gender)、受教育程度(education)、就职公司(company)、所在地域(location)、工作经验(yearsofexperience)、职位等级(level)、职业名称(title)、入职目前所在公司的时间(timestamp)、基本工资(basesalary)、股票期权(stockgrantvalue)、奖金(bonus)、总工资(totallyearlycompensation)等个人信息。将使用被调查者的这些信息来进行个人薪水的预测。

2.2 数据预处理

本文所用数据集为公开数据集，可能会存在缺失值、重复值等问题，所以在使用前要进行数据预处理。

(1) 缺失值处理。为了让模型更容易从数据中学到知识，我们先要对原始数据进行缺失值处理。按照不同变量的缺失值占比情况，本文首先对缺失值比例较小的 company, level, tag 属性执行删除操作；而缺失值较多，并且对结果影响较大的属性则不可以直接执行删除空值的操作，对于这部分缺失值本文使用 Unavailable 将其补齐。

(2) 独热编码。独热编码即 One-Hot 编码，其方法是使用 N 位状态寄存器来对 N 个状态进行编码，每个状态都有它独立的寄存器位，并且在任意时候，只有一位有效^[12]，即 $[1,0,1]$ 或

$[1,0,0]$ 。在本文使用的数据集中，某些特征并不总是数值型数据，还有可能是字符型数据，但这些字符型的特征值无法直接放入机器学习的回归模型中，所以在数据被放入模型训练之前要对数据集进行 One-Hot 编码，例如 education 这个特征就有高中、本科、硕士、博士这 4 个值，那么对 education 这个特征进行独热编码，新构造的特性就会有 4 种结果，学历一高中、学历二本科、学历三硕士、学历四博士，再与原数据进行对比，对应位置相同的就为 1，否则就为 0。

(3) 预测标签选择。该数据集中的个人薪水存在两个特征，分别为“*totallyearlycompensation*”“*basesalary*”，在进行预测之前要确定本实验以哪一个薪水特征为最终的预测标签，对于这个问题，本文分别为“*basesalary*”特征和“*totallyearlycompensation*”绘制箱线图^[13]，如图 3 所示，进而查看数据的分散情况，“*basesalary*”和“*totallyearlycompensation*”都是偏斜的分布，有很多异常值。但“*totallyearlycompensation*”受异常值的影响更大，应该是股票期权和奖金交易会对最终薪酬产生极大的影响。因此，本文选择“*basesalary*”作为预测标签。

2.3 评价指标

采用统计误差标准中的平均百分比误差 (MAPE)、确定系数 (R^2) 作为模型预测能力的评价指标，计算公式如下：

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (1)$$

其中： \hat{y} 表示模型的预测值； y 为真实值；MAPE 的值越高说明模型预测的准确率越低，预测效果越差。

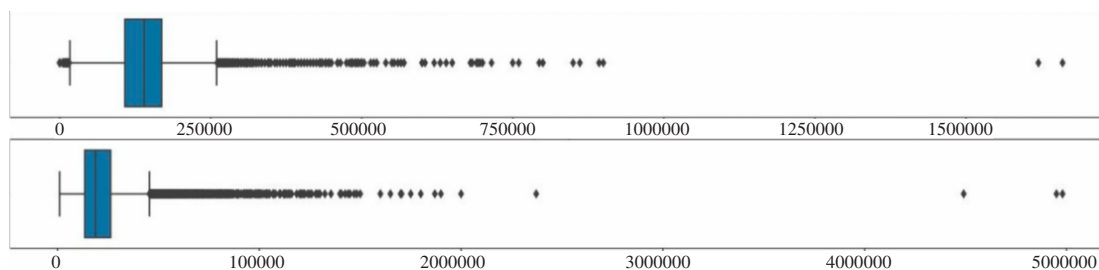


图 3 数据分布箱线图

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

其中： \bar{y} 表示均值， \hat{y} 表示模型的预测值， R^2 的值越高，表示模型性能越好。

2.4 基于单一模型的预测个人薪资结果对比

在本小节中使用随机森林、GBDT、XGBoost、AdaBoost等模型进行薪水预测实验，通过实验得出不同模型的性能，见表1，对比发现GBDT算法模型的MAPE值最低，因此预测结果准确性最高；相反，AdaBoost算法的MAPE值最高，因此预测该算法的预测结果准确性最低。

表1 不同模型的MAPE值

Model	MAPE/%
GBDT	13.59
XGBoost	13.65
随机森林	15.64
adaBoost	24.27

2.5 基于Stacking模型融合的个人薪水预测过程

在本小节中使用Stacking算法构造融合模型，首先筛选参与模型融合的基学习器，由于本文研究的是回归问题，所以在元学习器的选择上秉承着越简单越好的原则，因此将选择线性回归作为融合模型的元学习器，最终使用融合训练后产生的Stacking多算法集成模型对个人薪水进行预测。并得出模型预测薪水值和真实薪水值之间的仿真结果。

2.5.1 筛选基学习器

首先将数据集的80%作为训练集，20%作为测试集，为了筛选出适合作为基学习器的算法模型，本文分别使用训练集对GBDT、随机森林、Adaboost、XGBoost这四个算法的回归模型进行训练，比较不同模型的5折交叉验证得分情况。如表2所示，筛选出GBDT、随机森林、XGboost适合作为后续Stacking模型融合的基学习器，而Adaboost由于交叉验证得分太低，所以无法作为基学习器参与后续Stacking模型融合操作。

表2 不同模型的R²值

Model	R ²
GBDT	0.7575
XGBoost	0.7424
随机森林	0.6998
adaBoost	0.6209

2.5.2 实验流程

经过上面的模型选择，最终本文选择GBDT、XGBoost、随机森林作为基学习器，选择线性回归作为元学习器，展开模型融合工作。经过模型融合，构造出Stacking多算法集成模型，通过实验发现该模型的交叉验证的分数为0.8122，比任何一个基学习器单独建模的交叉验证分数都要高，说明经过Stacking模型融合操作，训练出来的模型拥有更好的泛化能力，薪水预测的准确性得到了提升。可以更好地帮助求职者评估自身能力，找到更满意的工作。

2.5.3 对比不同模型下的仿真结果

选取测试集中前100条预测结果，各模型预测结果与真实值对比情况，如图4所示。

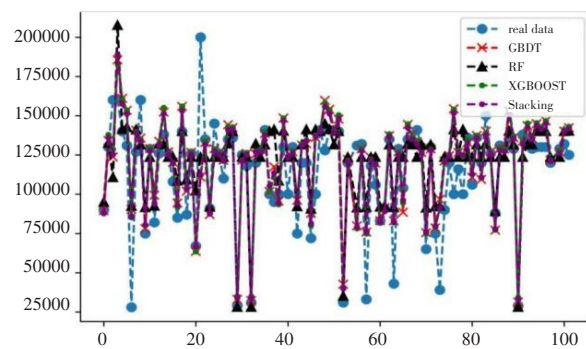


图4 预测与真实值的对比

由图4可知，使用Stacking多算法集成模型对真实值的拟合程度最高，该模型的预测结果最贴近真实值，对个人薪水的预测结果更为准确。

2.5.4 不同方法对薪水的仿真预测结果对比

使用不同模型进行个人薪水预测，得到的MAPE、R²值如表3所示。

表3 不同模型进行预测的MAPE、 R^2 值

Model	MAPE/%	R^2 %
GBDT	13.59	0.7575
XGBoost	13.65	0.7424
随机森林	15.64	0.6998
adaBoost	24.27	0.6209
Stacking集成模型	13.55	0.8122

3 结语

本文针对薪水预测问题, 构造了一个融合随机森林、XGBoost、GBDT三种算法的模型, 即Stacking多算法集成模型。最终通过实验与其他用作基学习器的单一模型进行对比, 结果表明: Stacking多算法集成模型在进行个人薪水预测时拥有更低的MAPE值, 以及更高的 R^2 值, 能更为准确地对个人薪水进行预测。进而说明使用Stacking方法可以提升模型的预测准确率。在日常研究中, 我们可以使用Stacking建立多算法集成模型, 进而设计出预测准确率更高的模型用于IT行业的薪水预测工作。

参考文献:

- [1] 朱婷婷. 数据类岗位的人才需求和薪资预测研究[D]. 武汉: 中南财经政法大学, 2020.
- [2] 师静谊. 数据分析师岗位薪酬分析[J]. 通讯世界, 2018, 25(12): 231-232.
- [3] 加春燕, 尚意展, 钟宇辉. 多元线性回归模型在软件行业薪资影响因素分析中的应用[J]. 北京工业职

业技术学院学报, 2019, 18(3): 32-35.

- [4] 潘博, 张青川, 于重重. Doc2vec在薪水预测中的应用研究[J]. 计算机应用研究, 2018, 35(1): 155-157.
- [5] KHONGCHAI P, SONGMUANG P. Random forest for salary prediction system to improve students' motivation[C]//2016 12th International Conference on Signal-Image Technology & Internet-Based Systems(SITIS), Naples, Italy, 2016.
- [6] MARTÍN I, MARIELLO A, BATTITI R, et al. Salary prediction in the IT job market with few high-dimensional samples: a Spanish case study[J]. International Journal of Computational Intelligence Systems, 2018, 11(1): 1192-1209.
- [7] JERRIM J. Do college students make better predictions of their future income than young adults in the labor force?[J]. Education Economics, 2015, 23(2): 162-179.
- [8] HUNG C C, LIM E P. On aggregating salaries of occupations from job post and review data[J]. IEEE Access, 2021, 9: 43422-43433.
- [9] 徐继伟, 杨云. 集成学习方法: 研究综述[J]. 云南大学学报(自然科学版), 2018, 40(6): 1082-1092.
- [10] 陈宇韶. 基于特征选择与改进Stacking算法的股价预测研究[D]. 衡阳: 南华大学, 2018.
- [11] 任红娟. 基于Stacking集成学习的网购客户流失预测研究[D]. 柳州: 广西科技大学, 2020.
- [12] 江琛, 夏晓玲, 吴伟, 等. 基于数据挖掘的上市公司高送转预测研究[J]. 湖北大学学报(自然科学版), 2021, 43(6): 698-705.
- [13] 赵颖祺, 陈玉虎, 张晰, 等. 一种基于箱线图的SAR图像舰船检测算法研究[J]. 中国海洋大学学报(自然科学版), 2021, 51(10): 130-140.

Prediction of personal salary based on Stacking fusion model

Lin Xinhui, Chen Chao*

(School of Computer Science and Engineering, Sichuan University of Science & Engineering, Yibin 644000, China)

Abstract: Salary is a factor that job seekers focus on. With the rapid development of information, job seekers hope to use various methods to reasonably evaluate their salary range when looking for a job. It is difficult for the job seekers to make reasonable predictions about their salary level, so it is reasonable. It is very important to evaluate your ability and then find the salary standard that suits you. Based on this real problem, we analyze the raw data and utilize the Stacking algorithm, that is, integrating random forest, gradient boosting tree, XGBoost, logistic regression and other models, a reasonable prediction of the salary of job seekers is made, and experiments show that using The model of Stacking fusion algorithm has better prediction performance and generalization ability.

Keywords: Stacking fusion model; personal salary; random forest; gradient boosted tree; XGBoost

文章编号: 1007-1423(2023)10-0030-05

DOI: 10.3969/j.issn.1007-1423.2023.10.006

智慧图书馆建设背景下构建图书智能即时剔旧体系的应用研究

罗可*

(邵阳学院图书馆, 邵阳 422000)

摘要: 智慧图书馆管理中, 馆藏剔旧是一个至关重要的环节。为了提高图书剔旧效率并减轻馆员工作负担, 提出了一种基于 OpenCV 技术和卷积神经网络 CNN 的图书即时智能剔旧系统。采用 OpenCV 技术对图书条形码图像进行分割, 然后运用卷积神经网络 CNN 进行图像识别, 将图书条形码与数据库进行比对, 并结合层次分析法计算得到图书的剔旧权重值, 最后根据设定的阈值将结果反馈给前端。设计和实现有效地提高了图书剔旧的效率, 为高校图书馆的智能化服务提供了创新思路。

关键词: 智慧图书馆; 馆藏剔旧; 层次分析法; CNN

0 引言

随着高校师生规模的不断扩大, 其馆藏资源和服务模式已经不能满足智慧图书建设的需求。图书馆的服务模式已经不仅仅单纯满足读者的基本信息需求和阅读需求, 逐渐向多样化、个性化的文化休闲需求转变, 这一转变使图书馆在承受着日益增大的空间压力的同时, 也在不断挤压着纸质图书的存储空间。

当前, 我国高校图书馆的纸质藏书量和年进书量是衡量学校办学水平的一个重要指标, 图书馆每年购买的图书数量都有上万册, 藏书数量也在不断增加, 但图书馆的藏书空间却是有限的, 图书剔旧成为当今图书馆界亟待解决的重大课题^[1]。随着网络信息技术的飞速发展, 智能移动设备的普及和电子图书的问世, 使得图书馆剔旧标准、剔旧方法得到了进一步的深化和扩展。本文以人工智能与图像识别技术为基础, 研发了一套智能即时图书剔旧系统, 以达到即时、精确的图书剔旧目的。

1 图书剔旧系统整体框架

1.1 图书剔旧系统功能描述

图书剔旧系统的功能主要是实现相对简单的图书条形码识别与智能剔旧处理。当用户通过智能设备中的拍照功能采集图书条形码图像, 将图像数据上传到云端服务器; 后台服务器会对图像数据进行预处理, 提取图片中的条形码来获取数据库与之对应的图书信息(如出版年份、借阅次数、价格等信息); 然后根据图书信息进行层次分析法剔旧量化, 并将剔旧量化信息反馈给客户端; 客户端接收到信息, 用户根据信息进行智能决策并进行剔旧系统处理。

1.2 系统整体设计

图书剔旧系统包含前端及后端两个部分: 前端为智能移动设备, 负责拍摄照片并收集条形码图像, 使用 WebSocket 方式将条形码图像数据发送给后端服务器, 最后根据后端反馈结果进行决策; 后端服务器得到图像数据后, 对数据流进行解析, 使用投影法技术对图像数字进

收稿日期: 2023-01-30 修稿日期: 2023-02-28

基金项目: 湖南省哲学社会科学基金项目(21YBA179): 人工智能技术在高校图书精准采购中的应用研究

作者简介: *通信作者: 罗可(1980—), 男, 湖南邵阳人, 硕士, 图书馆馆员, 研究方向为图书现代管理技术、计算机应用技术与开发, E-mail: luoke00@qq.com

行切割,生成相应的孤立的数字结果,然后采用CNN模型对数字识别结果进行处理,到数据库中匹配对应图书信息,将最终的量化后的文本结果发送给前端客户机,整体系统流程如图1所示。

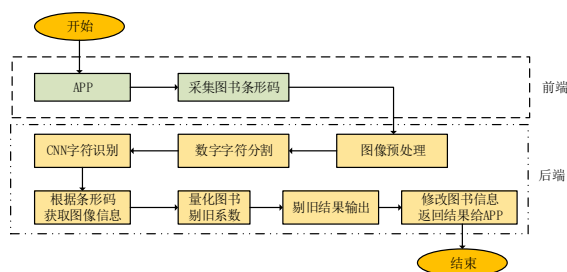


图1 图书剔旧系统整体流程

2 图书条形码预处理

一般而言,前端采集的原始图像都是彩色图像,为有效简化及保留图像的信息,以获取所感兴趣的区域,通常采用灰度化和二值化方法处理彩色的图像,以此来增强图书条形码定位,加速条形码中数字识别。

2.1 图像灰度化处理

图书条形码的图像方式会以一个二维函数 $f(x, y)$ 表示,每一个坐标位置即为一个像素,彩色图像中每个像素由红色(R)、绿色(G)及蓝色(B)三原色组成,如图2(a)所示。在计算机中每种颜色用8位表示,其灰度级范围为 $[0, 255]$,共256级。

一般而言,如果以彩色图像来处理会提取过多的特征,增加分析的误判率,所以将彩色图像转换成灰度图像,使得图像中的RGB色素皆为相同。彩色图像转换灰度图像的公式如下。

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (1)$$

根据所提供的公式(1),经过灰度化处理,原始条码图像(如图2(a)所示)以一种新的视觉呈现方式展现,如图2(b)所示。这种处理方法将图像转化为灰度级别,去除了颜色信息,使得条码图像更加突出、清晰,并便于后续的分析 and 处理。这种技术在图像处理领域中被广泛应用,为条码识别和相关应用提供了更可靠和准确的数据基础。

2.2 图像二值化处理

在图像采集过程中,由于受拍摄位置、角

度、天气等外部因素的影响,图像的灰度不均匀、对焦不精确、采样质量差、噪声等都会影响到图像的视觉质量。本文采用OSTU方法^[2]设定阈值,将灰度值大于阈值的设为灰度极大值,小于阈值的设为灰度极小值,进而实现二值化。二值化图像只呈现0和255的两种像素值图像,图2(c)为二值化后图书条形码图像。

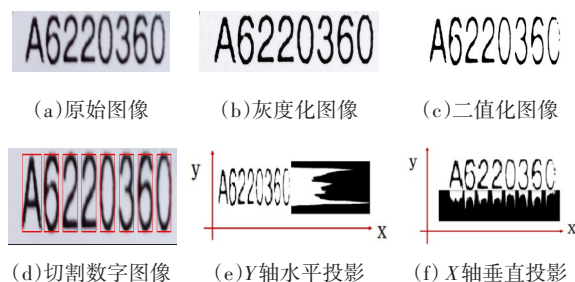


图2 图书条码预处理的效果

2.3 数字字符单元分割

本文采用OpenCV技术实现对采集的图书条形码图像进行分割处理,每一个数字之间都有差不多大小的空白区间,通过区间来分割图书条形码图像中的七位数字,如图2(d)所示。

依据投影法,利用透视原理将投影信息展示到直方图上,水平投影方式是沿着Y轴方向累计计算每列的白色像素点个数,以直方图来呈现各行所累计的白色像素点有多少,其中投影出来的大小为输入图像的高度部分,如图2(e)所示,以黑色为背景,累计在黑色背景上的白色像素点。

垂直投影方式是沿着X轴方向累计计算每行的白色像素点个数,如图2(f)所示,以直方图来呈现各列所累计的白色像素点有多少,其中投影出来的大小随输入图像的宽度而改变,以黑色为背景,累计在黑色背景上的白色像素点。

3 CNN 图像识别模型构建

3.1 构集自制数据集

本文所用数据抽取自2000多册图书的条形码。图书条形码都是统一的标准印刷体,数字识别只会受到不同的光照、倾斜角度、遮挡程度等影响,采集的图书条形码需要进行归一化处理等预处理,采用投影法将其分割成数字字符,依次按归类属性为0~9保存至文件夹中,其每个文件夹中的数字字符皆通过缩放至 28×28

大小。字符数量均衡有利于提高识别率，并且数据集越大越有利于训练，将每一个字符的采样数目统一为500个，然后将其归类属性为0~9，总共5000个样本，从每一个字符中随机抽取3750个图像构成训练集，剩下的1250张图像则构成一个测试集。

3.2 CNN模型仿真实验

本文5000个样本作为数据集使用，以随机抽样方式，将样本分成75%训练集及25%测试集，进行5次实验，实验5次后训练集及测试集准确率结果整理见表1。

表1 CNN模型5次随机抽样实验结果

	训练集		测试集
	准确率/%	准确率/%	识别时间(秒/张)
实验1	99.73	99.58	2.514×10^{-4}
实验2	99.63	99.34	2.324×10^{-4}
实验3	99.65	99.58	2.454×10^{-4}
实验4	99.74	99.55	2.351×10^{-4}
实验5	99.73	99.46	2.251×10^{-4}
平均值	99.70	99.50	2.378×10^{-4}

由表1得知，CNN模型5次随机抽样实验结果如下：训练集最低准确率为99.63%，最高为99.74%，平均准确率为99.70%；测试集最低准确率为99.34%，最高为99.58%，平均准确率为99.50%；在每张识别时间上，5次实验结果最低每张辨识时间为 2.514×10^{-4} 秒，最高每张辨识时间为 2.251×10^{-4} 秒，平均耗费时间为 2.378×10^{-4} 秒。

其中，最高准确率为99.74%的实验4所整理混淆矩阵，显示分类正确及错误的总和数，见表2。

由表2得知，正确识别数字数为1246，错误数为4个，分别为实际值为0，预测值为9的有1个；实际值为5，预测值为8的有1个；实际值为9，预测值为3的有2个。总体而言，经过训练和测试，该模型可以精确地识别出每一个数字，并且用户可以通过APP扫描获取相应的图书信息。

表2 测试集的混淆矩阵

实际值	预测值										总和
	0	1	2	3	4	5	6	7	8	9	
0	120	0	0	0	0	0	0	0	0	1	121
1	0	138	0	0	0	0	0	0	0	0	138
2	0	0	125	0	0	0	0	0	0	0	125
3	0	0	0	131	0	0	0	0	0	0	131
4	0	0	0	0	106	0	0	0	0	0	106
5	0	0	0	0	0	122	0	0	1	0	123
6	0	0	0	0	0	0	131	0	0	0	131
7	0	0	0	0	0	0	0	120	0	0	120
8	0	0	0	0	0	0	0	0	131	0	131
9	0	0	0	2	0	0	0	0	0	122	124
总和	120	138	125	133	106	122	131	120	132	123	1250

4 图书剔旧量化

图书剔旧是具有弹性而非随意性的工作，需实现原则性和灵活性的统一，各馆可根据其规模、特点、任务不同，以总体原则为指导，制定图书剔旧工作方案，并形成各图书馆剔旧工作的规则与方法，以达到图书剔旧的标准化、规范化^[3]。

通过文献分析法和调研国内高校图书馆，本文采用因素分析法和层次分析法^[4]相结合的方式构建剔旧指标(book weeding index, BWI)，以作为剔旧基准。

4.1 选取图书剔旧指标

图书剔旧指标的选取应充分考虑数据可得性和指标的量化难度^[3]，并尽可能地利用图书

管理系统中数据库本身信息,以确保指标的可测度。本文运用因素分析法对影响图书剔除的因素进行了分析,并将主要影响因素作为评测指标,运用这种科学而实用的方法构建一套评价指标体系,包括5个一级指标和8个二级指标。

如表3所示,剔旧指标体系主要由5项一级指标组成,每项指标又可细分为1~2个二级指标,共8项二级指标。利用主观赋权法确定各个指标的权重,综合多位专家的赋权经审议后确定最终权重。

表3 图书剔旧指标权重及参数设置

总指标	一级指标	二级指标	权重 $Q_i/\%$	参数设定值	描述说明
图书剔旧指标体系	时间因素 A_1	出版年代 A_{11}	20	剔旧当代年限设定值为0,上一年为1,以此类推	反映图书老旧程度
		版本 A_{12}	10	最新版本为设定值0,前一版为1,以此类推	
	著名因素 A_2	出版社 A_{21}	15	根据国家出版社排名分类为三个等级,第一等级设定值为0,第二等级设定值为1,其他设定值为2	反映图书品质
		作者 A_{22}	5	国家一级作家设定值为0,其他设定值为1	
	流通因素 A_3	复本量 A_{31}	20	1本复本量设定值为1,以此类推	反映图书受欢迎程度
		借阅次数 A_{32}	20	借阅次数为20次以上的设定值为0,每减少5次增加1,以此类推	
	价格因素 A_4	定价 A_{41}	5	价格500元以上的图书设定值为0,300~500元设定值为1,50~300元设定值为,50元以下设定值为3	反映图书资产价值
	替代因素 A_5	电子图书 A_{51}	5	图书系统中有电子图书设定值为1,无电子图书设定值为0	反映阅读的新渠道

4.2 生成剔旧指数

图书剔旧指数是各指标设定值与权重的加权和算法的综合得分,计算公式为 $BWI = \sum_{i=1}^2 A_{1i}q_i + \sum_{i=1}^2 A_{2i}q_i + \sum_{i=1}^2 A_{3i}q_i + A_{41}q + A_{51}q$, 其中 A_i 为二级指标设定值, q 为二级指标权重,各因素权重值能够体现各指标相对重要性。通过计算得到某本图书的剔旧指数,以此评判该书是否需要剔旧, BWI 越大说明该书所需剔旧程度越高。

5 结语

本研究在深度学习和数字图像识别技术的基础上,采用B/S架构服务模式,开发实现了一套图书剔旧系统。该系统通过APP前端获取数据源,利用自制数据集和卷积神经网络模型实现了对图像数字的快速、准确识别,进而实现了图书条形码的匹配和图书剔旧量化计算的自动化。基于此技术应用,可以极大地减轻图书

剔旧人员的工作强度,并能有效地提高图书剔旧数据的准确率和即时性。该系统的研究和实现,对于推动图书管理的智能化、自动化和数字化具有积极意义,也为相关领域的研究提供了新的思路和方向。

参考文献:

- [1] 尚文娟. 浅谈高校图书馆馆藏图书剔除[J]. 黑龙江畜牧兽医, 2016(24): 262-264.
- [2] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1997, 9(1): 62-66.
- [3] 信丹丹. 公共图书馆图书剔旧指数指标体系构建研究[J]. 图书馆论坛, 2013, 33(3): 122-125.
- [4] 鲍计国, 栾美晨. 高校图书馆剔旧工作创新实践: 以辽宁对外经贸学院图书馆为例[J]. 图书馆杂志, 2017, 36(2): 74-78.

(下转第49页)

实践与经验

文章编号: 1007-1423(2023)10-0034-09

DOI: 10.3969/j.issn.1007-1423.2023.10.007

基于主观逻辑和拍卖算法的公证人组跨链交易机制

廖政*, 魏博众, 罗健炜

(广东工业大学自动化学院, 广州 510006)

摘要: 作为区块链领域研究的热点, 跨链技术一直备受关注。公证人机制作为一种典型的跨链方法, 会出现“公证人节点职能集中化”“节点信誉度不高”“中间人费用过高”等问题, 对此提出了一种基于主观逻辑的公证人组跨链交易机制。通过引入主观逻辑的方法对单链上节点间的相互评估以及担任正式公证人后的交易信息进行评估, 筛选出高信誉度的公证人组, 使得公证人机制更为安全可靠。同时为了合理化公证人充当中间人时的费用, 提出了一种基于反向维克里拍卖算法的定价机制, 降低了公证人节点充当中间人时的价格。实验结果表明, 相对于传统的公证人机制, 该跨链交易的机制具有更好的安全性和更低的中间人费用。

关键词: 公证人机制; 区块链; 跨链; 主观逻辑; 信誉值排序; 拍卖算法

0 引言

区块链技术以其去中心化、可追溯、不可篡改等优点被应用到越来越多的领域^[1], 但也存在吞吐量低、链间数据孤岛、无法有效监管等问题。其中, 区块链数据的“孤岛问题”最为严重, 因为通常每条区块链都是相互独立的, 链与链之间不存在信息交换^[2]。因此, 如果需要其他链的信息, 便可使用跨链技术来满足这种需求^[3]。

本文在主观逻辑算法、拍卖算法的基础上, 提出了一种基于主观逻辑和拍卖算法的公证人组跨链交易机制。该机制主要由两个部分组成:

第一部分是根据信息变化实时地对公证人进行信誉值的排序, 筛选出参与下一笔跨链交易的正式公证人组。信誉值的排序需要两个信息, 一个是节点在两条链上和其他节点的交互信息, 另一个是节点作为正式公证人参与交易时的交易信息。以这两个信息作为判断依据, 调用主观逻辑算法得出各个节点的信誉值分数,

动态生成节点信誉值排序, 剔除排名靠后的节点生成候选公证人组。再随机选取部分节点构成正式公证人组进行跨链交互。通过这种方式一方面解决了公证人职能过于集中的问题, 另一方面也保证了参与跨链交易的公证人组有较高的信誉值, 大大降低了遭遇女巫攻击的风险。

第二部分是基于反向维克里拍卖算法的定价机制。我们知道一次跨链交易由交易的发起者(原链和目标链上的节点)、交易的响应者(公证人组)两部分组成。因为位于不同区块链的交易双方互相不信任, 所以由信誉值高的公证人担任中间人角色参与交易。但是交易的发起者和交易的响应者都是自私的, 都希望能够以更低的成本获取更大的利益, 因此为了最大化双方的利益, 便可采用该机制确认交易响应者执行服务任务时的价格。在公证人通过信誉值初筛的基础上, 我们进一步地对参与跨链交易的公证人进行选择。我们将公证人节点的选择抽象为服务提供商的竞标过程。通过基于反向维克里拍卖算法, 选择其中服务费用最低的一

收稿日期: 2023-01-05 修稿日期: 2023-02-27

作者简介: *通信作者: 廖政(1997—), 男, 江西吉安人, 硕士研究生, 研究方向为区块链、区块链跨链、联邦学习等, E-mail: lz4135@qq.com; 魏博众(1999—), 男, 河南驻马店人, 硕士研究生, 研究方向为区块链、联邦学习; 罗健炜(1998—), 男, 广东广州人, 硕士研究生, 研究方向为数字孪生、自动驾驶仿真测试

组公证人节点作为该笔交易的中间人来提供交换服务。跨链交易完成后，赢家公证人组将获得下一个最低出价作为奖励。通过这种方式，交易请求者向公证人节点支付的交易费用最小化。

区别于目前的研究成果，本文的贡献点如下：一是对公证人的信誉值排序采用了主观逻辑算法，该算法可以明确地将不确定性和来源信任考虑在内，通过评估节点在单链上和其他节点的交互信息以及节点作为正式公证人参与交易时的交易信息来得出公证人节点的信誉值；二是提出了一种拍卖算法，该算法可以通过比较不同公证人的出价，选出公证人节点中出价最为合理的公证人节点并通过智能合约自动完成对中间人奖励的发放。

1 相关研究

1.1 公证人机制概述

公证人机制是指互不信任的双方处在不同区块链上，为保证交易双方信息的同步性和合法性，需要一名双方均信任的公证人^[4]。然而，这又会产生“中心化问题”^[5]。

我们知道公证人机制能够有效地帮助实现跨链交互^[6-8]。但是传统的公证人机制仍存在以下三个主要问题：第一，公证人节点信用不足，存在可靠性隐患的问题；第二，公证人机制存在中心化问题，如果公证人中存在恶意节点，则跨链交易的安全性将不能够得到保证；第三，公证人激励机制的设计不够合理，无法充分调动公证人参与跨链交易的积极性导致跨链交易效率低下^[9-10]。因此希望通过设计一种新的机制，一方面弱化公证人节点的职能，同时增加节点的信誉度、提高公证人参与跨链交易的积极性，从而增大交易的吞吐量，提高跨链交易的效率。

1.2 主观逻辑算法概述

主观逻辑是将意见作为输入和输出变量的概率逻辑，意见明确地表达了对概率值的不确定性，并且可以表达对某个主题(例如信任)的无知程度^[11-12]。

本文利用主观逻辑算法，以公证人节点以往交易信息作为判断依据，得出各个公证人节

点的信誉值，根据信誉值对公证人节点进行排序，剔除信誉值排名靠后的节点来保证公证人组整体的可信度。节点以往交易信息包括两部分，一部分是节点之前作为正式公证人时经历交易后获得的价值信息；另一部分是节点在源链和目标链上的单链以往交易评估得到价值信息(若公证人节点之前未担任过正式公证人，则只考虑该部分价值信息)。

1.3 拍卖理论概述

反向维克里拍卖模型是反向定价机制和维克里拍卖模型的结合^[13-15]。目前的服务定价方法主要分为张贴定价和反向定价。与平台(或雇主)直接给出明确价格的定价机制不同，反向定价机制通过服务提供商的投标来确定服务任务的价格。服务平台需要根据各种因素选择最终的赢家来执行服务任务，并计算服务价格。这种定价机制不同于正常的拍卖^[16]。与反向拍卖的中标结果类似，它通常由出价较低的人中标。

为了防止投标人在实际场景中夸大真实成本，与恶意低价竞争，维克里拍卖机制也被称为第二高价密封拍卖，已被广泛引入定价过程^[17]。维克里拍卖首先由 William Vickrey 提出，它提供了一种机制，在这种机制中，如实报告其服务价值是每个竞标者的主导和最佳策略。在拍卖竞价过程中，每个参与者在不知道他人竞价的情况下秘密竞价。在投标结束时，所有当前报价都会被计算在内，最高的投标人获胜，但只支付第二高的报价。投标人没有理由误报其报价，因为他不知道对手的报价，而且他的报价不会影响最终价格。

2 基于主观逻辑的公证人节点信用评级模型

2.1 公证人节点信誉值排序流程

本文构建了一个动态的公证人节点信用评级模型，以保证公证人机制的安全性。本模型以单链上节点之间的交互信息和节点作为正式公证人时节点之间的交易信息，以此为判断依据，调用主观逻辑算法获取各节点的信誉值分数并动态生成排名，排名靠前的节点作为候选公证人组。最后，从候选公证人组中随机抽取

一定数量的公证人，这些公证人组成正式公证人组担任跨链交易的“中间人”。流程如图1所示。

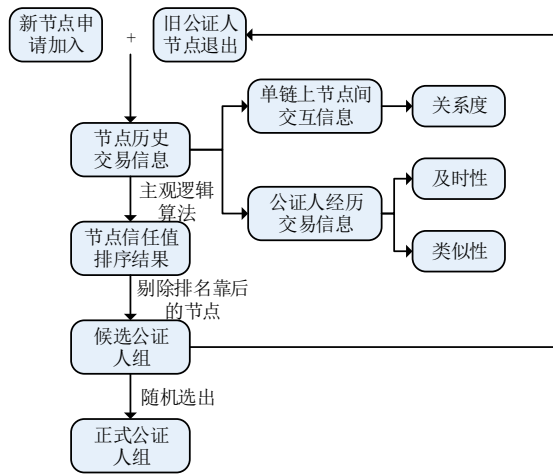


图1 公证人节点信誉值排序流程

2.2 多重主观意见的融合计算

如果主观逻辑模型只考虑局部意见，则意见的评价过于主观和不完整。因此，本文使用多重主观逻辑模型来强化关键声誉意见的作用，同时避免被有缺陷的声誉意见所误导。

从上文我们知道，评价者对评价者的先验知识越多，评价意见的可信度就越高。另外，评价者对评价者在一定时期内的评价意见更具参考性。另外，如果评价者和评价者在某些方面非常相似，评价意见的准确性也会大大提高。因此，本文主要考虑关系度、及时性和相似性三个权重来形成区域意见。

关系度：关系度是用于衡量单链上节点*i*（评估节点）对节点*j*（被评估节点）的关系程度的权重。当节点之间有较高的交易频率时，说明节点*i*对节点*j*的信誉评价更为准确可靠。节点之间的关系度是节点*i*和节点*j*在一定的时间窗口内的交易频率以及节点*j*与其他节点的平均交易频率的比率，即：

$$I_{i,j} = \frac{r_{i,j}}{\frac{1}{|N|} \sum_{i \in N} r_{i,j}} \quad (1)$$

其中， $r_{i,j}$ 表示在一段时间内，节点*i*对节点*j*的交易频率，假设节点*i*在该时间段内与*N*个节点交互。

及时性：及时性是用于衡量节点之间信誉

意见新鲜度的权重。一个节点的信誉值并不是保持不变的，而是随着时间的变化不断上下波动的。因此，在区块链上，节点之间的信誉意见并不总是可信可靠的。由经验可知，时间越近的交易事件比时间越远的交易事件有更强的影响。所以，若交易是最近发生的事件，则节点*i*对节点*j*的局部信誉意见就会产生较大的影响；若交易是过去的事件，则节点*i*对节点*j*的局部信誉意见的影响就会较小。本文设定节点*i*对节点*j*的本地信誉意见的及时性为

$$V_i^j = \alpha_1 (t - t_i')^{-\alpha_2} \quad (2)$$

其中， α_1 和 α_2 是预设的参数，用于调整及时性的作用。 t 是当前的时间， t_i' 是意见 $o_{i,j}$ 的形成时间。

类似性：类似性是通过观察者的表现状态与被观察者的表现状态的类似程度来衡量观察者观察条件优劣。即当观察者与被观察者的类似性越高时，观察者对被观察者的信誉意见的判断更为准确。节点*i*和节点*j*之间的类似性大小是由两个节点一起充当公证人节点时的交易情况来判断的。若双方均是交易完成率高的节点，那么两者的类似性则更高。这里，类似性的计算是基于两个节点的交易完成数量。本文用 Z_i^j 表示节点*i*和*j*之间的类似性，即：

$$Z_i^j = \frac{1}{1 + |\ln n_i^\alpha - n_j^\alpha|} \quad (3)$$

其中， n_i^α 和 n_j^α 表示节点*i*和*j*完成跨链交易的数量。

合成信誉意见计算：上文分别从关系度、及时性和类似性三个方面来对公证人节点的信誉度进行了分析和考虑。下一步，本文便以关系度、及时性和类似性作为权重进行合成信誉意见的计算。当以关系度、及时性和类似性作为权重，节点*i*对节点*j*的本地意见 $o_{i,j}$ 的总体权重为

$$w_i^j = \gamma_1 I_i^j + \gamma_2 V_i^j + \gamma_3 Z_i^j \quad (4)$$

其中 $\gamma_1, \gamma_2, \gamma_3$ 是三个预设定好的权重因子，并且满足 $\gamma_1 + \gamma_2 + \gamma_3 = 1$ 的条件。本文用 $o_{i,j}^{syn} = (b_{x_j}^{syn}, d_{x_j}^{syn}, u_{x_j}^{syn}, a_{x_j}^{syn})$ 来表示公证人节点*i*的本地意见与来自公证人节点*j*的其他相邻节点的信誉意见的加权合成信誉意见。故 $o_{i,j}^{syn}$ 的参数计算如下：

$$\left\{ \begin{aligned} b_{x,j}^{\text{syn}} &= \frac{1}{\sum_{x \neq j} v_x^j} \sum_{x \neq j} w_x^j b_{x,j} \\ d_{x,j}^{\text{syn}} &= \frac{1}{\sum_{x \neq j} v_x^j} \sum_{x \neq j} w_x^j d_{x,j} \\ u_{x,j}^{\text{syn}} &= \frac{1}{\sum_{x \neq j} v_x^j} \sum_{x \neq j} w_x^j u_{x,j} \end{aligned} \right. \quad (5)$$

其中, $b_{x,j}^{\text{syn}}, d_{x,j}^{\text{syn}}, u_{x,j}^{\text{syn}}$ 都满足 $b_{x,j}^{\text{syn}} + d_{x,j}^{\text{syn}} + u_{x,j}^{\text{syn}} = 1$ 。

最终信誉意见计算: 对于最终信誉意见的计算, 本文采用异或的操作 $o_{i,j}^{\text{final}} = o_{i,j} \oplus o_{i,j}^{\text{syn}}$, 计算如下:

$$\left\{ \begin{aligned} b_{i,j}^{\text{final}} &= \frac{b_{i,j} u_{x,j}^{\text{syn}} + b_{x,j}^{\text{syn}} u_{i,j}}{u_{x,j}^{\text{syn}} + u_{i,j} - u_{x,j}^{\text{syn}} u_{i,j}} \\ d_{i,j}^{\text{final}} &= \frac{d_{i,j} u_{x,j}^{\text{syn}} + d_{x,j}^{\text{syn}} u_{i,j}}{u_{x,j}^{\text{syn}} + u_{i,j} - u_{x,j}^{\text{syn}} u_{i,j}} \\ u_{i,j}^{\text{final}} &= \frac{u_{i,j}^{\text{syn}} u_{i,j}}{u_{x,j}^{\text{syn}} + u_{i,j} - u_{x,j}^{\text{syn}} u_{i,j}} \end{aligned} \right. \quad (6)$$

基于 $o_{i,j}^{\text{final}} = (b_{i,j}^{\text{final}}, d_{i,j}^{\text{final}}, u_{i,j}^{\text{final}}, a_{i,j}^{\text{final}})$, 节点 i 对节点 j 产生最终的最终信誉值 $g_{i,j}^{\text{final}}$ 可以这样计算:

$$g_{i,j}^{\text{final}} = b_{i,j}^{\text{final}} + d_{i,j}^{\text{final}} + u_{i,j}^{\text{final}} \quad (7)$$

3 基于反向维克里拍卖算法的定价机制

由上节的内容可知, 我们得到了信誉值排名靠前的候选公证人节点, 这些节点是相对可信可靠的。但是, 针对每一笔具体的跨链交易, 我们还需要一套适合的机制, 它能够选择出能够最大化交易发起者和交易响应者双方利益的公证人。因此, 本文提出了一种基于反向维克里拍卖算法的定价机制。

3.1 构建反向维克里拍卖模型

我们知道, 一次跨链交易由交易的发起者(原链和目标链上的节点)、交易的响应者(公证人组)两部分组成。因为位于不同区块链的交易双方互相不信任, 所以由信誉值高的公证人担任中间人角色参与交易。但是交易的发起者和交易的响应者都是自私的, 都希望能够以更低的成本获取更大的利益, 因此为了最大化双方的利益, 本文采用了基于反向维克里拍卖算法的定价机制来确认交易响应者执行服务任务时的价格^[18]。

在候选公证人通过信誉值初筛的基础上, 我们进一步地对参与跨链交易的正式公证人进行选择。我们将正式公证人节点的选择抽象为服务提供商的竞标过程。通过基于反向维克里拍卖算法, 选择其中服务费用最低的一组公证人节点作为该笔交易的中间人提供交换服务。跨链交易完成后, 赢家公证人组将获得下一个最低出价作为奖励。通过这种方式, 交易请求者向公证人节点支付的交易费用最小化。

在一次跨链交易过程中, 假设有 n 个公证人节点, 所有这些公证人节点均满足信誉值要求, 并且可以帮助完成这笔跨链交易。交易响应者的候选集为 $W = \{w_1, w_2, \dots, w_n\}$, 最终担任这笔交易正式公证人的赢家集为 $W_s (W_s \subseteq W)$ 。让 v_i 作为公证人节点的出价, γ_i 作为公证人节点真正想得到的实际报酬, $\min(v_{-w_i})$ 是除去 i 外所有投标人的最低投标价, 那么公证人节点 w_i 的收入 η_i 表示为

$$\eta_i = \begin{cases} \min(v_{-w_i}) - \gamma_i, & \text{if } v_i < \min(v_{-w_i}) \\ 0, & \text{if } v_i > \min(v_{-w_i}) \end{cases} \quad (8)$$

如果 $v_i > \gamma_i$, 则说明投标人获得负收入; 如果 $v_i < \gamma_i$, 当 $v_i < \min(v_{-w_i})$ 时, 则投标人的收入为 0; 如果 $v_i = \gamma_i$, 则说明投标人达到最优策略。

根据以上分析, 在投标过程中, 直接给出自己的实际价格是每个公证人节点的最佳策略。因此, 每次服务平台的选择都是从当前所有节点的实际报价中选择费用最低的公证人节点作为跨链交易的中间人。假设一次跨链交易需要的公证人节点数量为 k 个, 而满足信誉值要求的赢家集为 $W = \{w_1, w_2, \dots, w_n\}$ 。那么交易请求者需要花费的中间成本总计为

$$P_{\text{total}} = \sum_{r=1}^k v_{w_r} = \sum_{r=1}^k \min\{v_1, v_2, \dots, v_n\}^r \quad (9)$$

其中: r 表示的是对正式公证人节点的选择; P_{total} 是一次跨链交易中交易请求者需要向正式公证人组支付的最小交易中间费用。

3.2 正式公证人节点选择流程

反向维克里拍卖算法具有的特性可以很好地处理正式公证人节点选择的问题。在本小节中, 我们利用反向维克里拍卖算法机制来解决跨链交易公证人的选择问题并对该流程进行说

明。假设现在有一笔链A到链B的跨链交易需要完成，已知公证人节点 $W = \{w_1, w_2, \dots, w_n\}$ 均为信誉值排名靠前的节点。

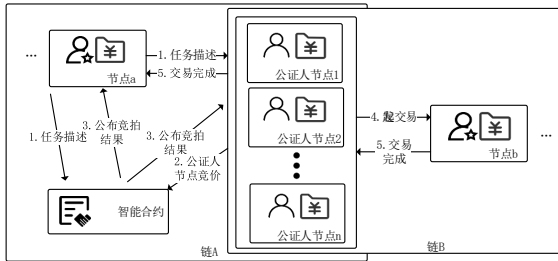


图2 正式公证人节点选择流程

正式公证人节点选择流程可以分为如下几步：

(1) 链A上的节点a发出一笔跨链交易请求，并建立了一个智能合约SC。其中，跨链交易请求包含交易量 $coin$ 、目的区块链B和响应截至日期 t ，合约用于执行基于反向维克里拍卖算法的正式公证人节点选择过程。链B上满足节点a跨链请求的节点b响应请求，并向智能合约SC提交服务费 v_i ；

(2) 按照智能合约SC的要求， n 个候选公证人节点开始进行竞价拍卖；

(3) 在候选公证人节点中，智能合约SC根据报价消除不符合条件的节点，并对候选公证人节点的服务报价进行排序，在完成投标排序之后，将公布交易价格 v_w 和中标的公证人节点 w ；

(4) 竞标成功的公证人节点 w 将作为节点a本次跨链交易的中间人帮助节点a完成这笔跨链交易，并提前提交下一次低价竞标的奖励；

(5) 在跨链交易完成后，智能合约将自动向公证人节点 w 支付报酬。

算法1描述了智能合约SC在公证人节点选择过程中的功能。

算法1：正式公证人节点选择流程

输入：交易金额： $coin$ ，响应截止时间： t ，候选公证人集 W ，目的区块链： B ，选择的公证人节点数量： k 。

输出：赢得竞标的公证人节点 w 和其交易费用 v_w ，赢得竞标的公证人节点组 Win 。

(1) 收集出标者们的竞价

1: 初始化: $Response\ Deadline = t, Bidding = \{0, \dots,$

$0\}, BiddingId = \{0, \dots, 0\}, Win = \{0, \dots, 0\};$

2: $send\ task(coin, t, W, B)$ to SC;

3: while $datetime.now()$ in t do

4: $getResponse(v_i, w_i);$

5: $Bidding[i] = v_i;$

6: $BiddingID[i] = w_i;$

7: end while

8: Stop Bidding Process

(2) 决定竞标成功者

9: Set $NT = getNotaryTable$ from WeCross

10: for $bidderID$ in $BidderID$ do

11: if $(bidderID \neq NT)$ then

12: $Bidding.remove(bidder_{th});$

13: $BidderID.remove(bidderID);$

14: end if

15: end for

16: for i in k do

17: $firstPrice = \min(Bidding);$

18: $Winner\ w' = BidderID[(Bidding_{th}(winPrice))];$

19: $Bidding.remove(winPrice);$

20: $finalPrice\ v_{w'} = \min(Bidding);$

21: $Win[i] = w'$

22: end for

23: return $Win;$

3.3 跨链支付方案流程

跨链支付方案的流程如图3所示。

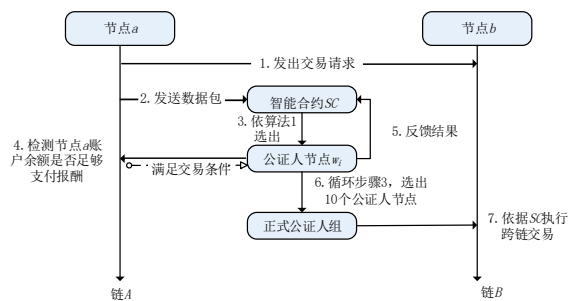


图3 跨链支付方案流程

(1) 链A上的节点a向链B上的节点b发出了交易请求。通过智能合约SC，节点a发送了一个数据包，其中包含节点b的账户地址、交易的金额、数据的哈希值和过期时间。SC按照3.2节中的步骤执行算法1，选择 w_1 作为正式公证人组的第一个公证人节点。公证人节点的中间费用由SC持有，交易完成后智能合约自动向

公证人节点支付报酬。

(2) 公证人节点 w_i 检查节点 a 账户中的金额是否足以支付交易。如果足够, w_i 将减少 a 节点账户中的余额, 否则交易将被拒绝。 w_i 将根据许可证构建并使用智能合约 SC 。

(3) 智能合约 SC 使用算法 1 从候选公证人集中选择下一个公证人节点, 重复步骤(1)。当选择的公证人节点数量达到 10 个时, 这 10 个公证人节点将组成正式公证人组担任该笔跨链交易的中间人。公证人组的节点数量越多, 安全性越高, 但交易的速度也会越低。

(4) 跨链交易完成后, 正式公证人组会把消息传递给节点 a , 当检测到节点 a 的确认后, 所有公证人节点的奖励将由智能合约自动支付。

4 实验及分析

4.1 主观逻辑算法的仿真实验设计

本实验首先部署了百名公证人节点, 从 0 至 99 连续整数位置编号, 收集节点间的关系度评价表, 即在时间段 t 内一个节点和另一个节点在单链上是否有过交易。10 个公证人节点之间的关系度示意图如图 4 所示。

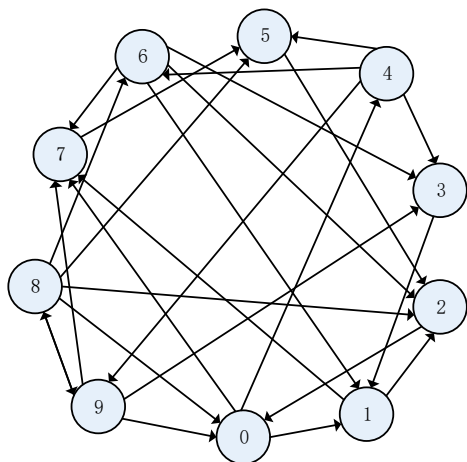


图 4 公证人节点关系度示意图

为保证实验进行, 需要收集部署的百名公证人节点在两条链上参与的每笔交易是否成功的信息、节点在充当公证人交易中的以往交易信息。另外, 在充当正式公证人的交易信息中, 交易成功用状态值 1 表示, 交易失败用状态值 0 表示。当状态值为 0 时, 处理时间和反馈评价用

“-”表示。节点 0 在单链上的交易信息如表 1 所示, 在充当公证人中的交易信息如表 2 所示。

表 1 节点 0 在单链上以往交易信息

交易编号	状态值
0	1
1	0
2	1
3	1
4	1
5	0
6	1

表 2 节点 0 充当公证人以往交易信息

交易编号	交易状态	处理时间	反馈评价
0	1	60	4
1	1	96	1
2	0	—	—
3	1	37	3
4	1	51	5
5	0	—	—
6	1	59	2

本实验基于收集到的节点关系度关系图、公证人节点在单链上以及充当正式公证人时的以往交易信息进行仿真测试, 基于主观逻辑算法的仿真参数设计如下: 对于某一个意见, 关系度, 及时性以及类似性的权重因子分别为 $\gamma_1 = 0.4$, $\gamma_2 = 0.3$, $\gamma_3 = 0.3$ 。关于及时性预定义的参数 $\alpha_1 = 10$, $\alpha_2 = 1.05$ 。为实现对比, 本实验采用了线性信誉值方案, 线性信誉方案的权重设置为 0.5, 则线性信誉值计算的函数可表示为: $x_{i,j} = 0.5x_{i,j}^{ave} + 0.5x_{i,j}^{las}$, 其中 $x_{i,j}^{ave} = b_{i,j}^{ave} + 0.5 * u_{i,j}^{ave}$ 且 $x_{i,j}^{las} = b_{i,j}^{las} + 0.5 * u_{i,j}^{las}$ 。 $b_{i,j}^{ave}$ 和 $u_{i,j}^{ave}$ 是 $b_{i,j}$ 和 $u_{i,j}$ 的平均值。

接下来, 通过实验比较多重主观逻辑(multi-weighted subjective logic)方案和传统线性信誉(traditional linear reputation)方案的性能。本实验设置了观察公证人节点行为的时间间隔为 10 分钟, 每分钟更新一次公证人节点的信誉值。异常的节点起初伪装成正常节点, 一段时间过

后开始进行攻击。

如图5所示，在观察的时间段内，随机选择10个不同的异常公证人节点进行信誉值展示。坐标轴的横轴是10个不同的异常公证人节点，纵轴则是公证人节点的信誉值。从图5我们可以清晰地看到在10分钟后，MSL方案的信誉值是普遍低于TLR方案的信誉值；在20分钟后，MSL方案信誉值的下降幅度也是远大于TLR方案的。这归功于本文提出的多重主观逻辑方案，相较于传统线性信誉方案考虑了先验知识对公证人节点信誉意见的影响，分别从关系度、及时性和类似性出发来对公证人节点进行信誉评价。正是因为加入了三个重要意见的信誉评价，避免了次要信誉评价的误导，才使得MSL方案的信誉评价具有更高的准确度和可靠性。另外，本文提出的MSL方案比起TLR方案而言，对异常公证人节点信誉值的变化更加敏感，这说明当公证人节点出现异常行为时，MSL方案能够有着更短的响应时间、更快的响应速度。所以，MSL方案对于公证人节点在异常行为方面的检测能力也更为出色。

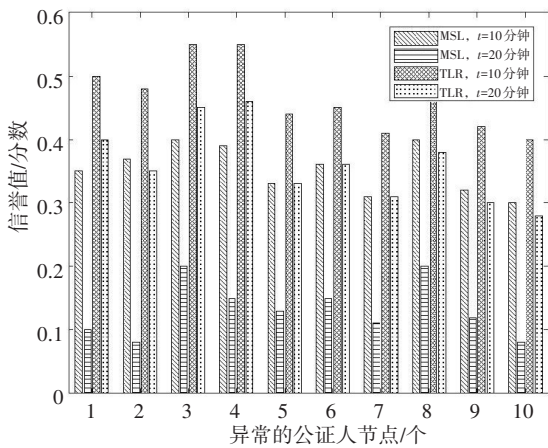


图5 10个异常公证人节点在两种方案下的信誉值比较

4.2 反向维克里拍卖算法的仿真实验设计

本实验在Ubuntu20.04操作系统上进行，主机CPU为AMD Ryzen 7 6800H with Radeon Graphics，时钟频率为3.20 GHz，使用的开源框架为WeCross跨链管理平台。本实验首先部署了100个候选公证人节点，为了评估本文的拍卖方案和随机选择公证人节点在中间人费用上的差异，

实验分别模拟了基于反向维克里拍卖选择和随机选择公证人节点两种方案，仿真结果如图6所示。

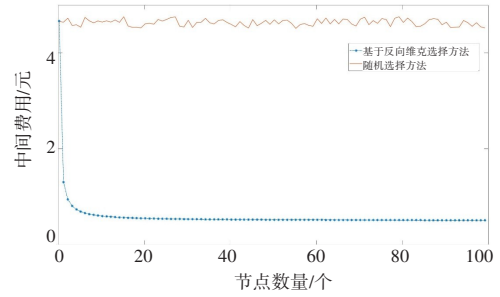


图6 两种方案对消耗中间费用的影响

由图6可知，随着参与节点数量的增加，随机选择公证人节点的方案始终将跨链中间费用维持在较高水平。在本文的方案中，随着节点数量的增加，中间费用的成本不断降低，最终保持在一个较低的水平。这表明本文的方案可以有效降低正式公证人的中间费用。

此外，考虑到加入拍卖机制后会消耗部分时间用于执行拍卖流程，跨链交易的交易执行过程会出现时延问题。为了清楚地展示基于拍卖机制的公证人节点选择方案对跨链交易时延的影响，本文测试了拍卖机制的附加时延。由于参与拍卖算法的候选公证人节点均是信誉值排名靠前的节点，所以我们可以对参与拍卖节点的数量设置限制(一般不会太多)，本实验设置参与拍卖算法节点的数量为100个。

从图7可以看到，随着参与拍卖的公证人节点数量的增加，执行反向拍卖所需的时间也在持续增加。但是当有100个候选公证人节点参与时，整个拍卖制造成的延迟仍然只有55 ms。因此，拍卖算法并不会导致跨链交易过程出现较大的时延问题。

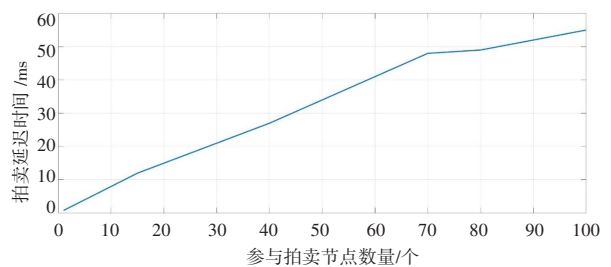


图7 反向拍卖机制的时延

5 结语

针对传统公证人机制中存在的“公证人节点职能集中化”“节点信誉度不高”“中间人费用过高”等问题, 本文首先提出了一种基于主观逻辑的公证人信用排序模型。通过引入主观逻辑的方法对单链上节点间的交互信息以及担任正式公证人后的交易信息进行评估, 筛选出信誉值排名靠前的节点作为正式公证人组的候选公证人节点, 这一方面解决了公证人职能过于集中的问题, 另一方面也保证了参与跨链交易的公证人组有较高的信誉值, 大大降低了遭遇女巫攻击的风险, 使得公证人机制更为安全可靠。

此外, 提出了一种基于反向维克里拍卖算法的定价机制, 实验结果表明该方案可以有效降低公证人的中间费用。相对于随机选择公证人节点参与交易, 该机制可以最大化双方的利益, 保证公证人执行服务任务时的价格合理, 使得交易请求者向公证人节点支付的交易费用最小化。综上所述, 本文提出的跨链交易机制具有更好的安全性和更低的中间费用。

参考文献:

- [1] NAKAMOTO S. Bitcoin: a peer-to-peer electronic cash system[M]. Manubot, Tech. Rep., 2008.
- [2] WOOD G. Polkadot: vision for a heterogeneous multi-chain framework [EB/OL]. <https://github.com/polkadot-io/polkadotpaper/raw/master/PolkaDotPaper.pdf>.
- [3] 潘晨, 刘志强, 刘振, 等. 区块链可扩展性研究[J]. 计算机研究与发展, 2018, 55(10): 2099-2110.
- [4] POON J, BUTERIN V. Plasma: scalable autonomous smart contracts[EB/OL], 2017, 22: 1-47. <https://www.plasma.io/plasma=deprecated.pdf>.
- [5] 孙国梓, 王纪涛, 谷宇. 区块链技术安全威胁分析[J]. 南京邮电大学学报(自然科学版), 2019, 39(5): 48-62.
- [6] 赵涛, 张凌浩, 赵其刚, 等. 基于聚类簇中心的共识跨链交换模型[J]. 计算机科学, 2019, 46(S2): 557-561, 566.
- [7] 刘桂华. 基于公证人组的区块链跨链机制[D]. 重庆: 重庆邮电大学, 2020.
- [8] 戴柄荣, 姜胜明, 李顿伟, 等. 基于改进PageRank算法的跨链公证人机制评价模型[J]. 计算机工程, 2021, 47(2): 26-31.
- [9] HE Y C, ZHU X Y, XU F F, et al. A novel cross-chain mechanism for blockchains[C]//Proceedings of the 2018 International Conference on Smart Blockchain, Berlin: Germany, 2018: 139-148.
- [10] MERKLE R. Protocols for public key cryptosystems [C]//Proceedings of the 1980 IEEE Symposium on Security and Privacy, Oakland, CA, USA, 1980: 122-122.
- [11] DENG L, CHEN H, JING Z, et al. Research on cross-chain technology based on sidechain and hash-locking[C]//Proceedings of the 2018 IEEE International Conference on Edge Computing. Piscataway, NJ: IEEE, 2018: 144-151.
- [12] 王思明. 基于区块链的车载边缘计算资源优化研究[D]. 广州: 广东工业大学, 2019: 25-35.
- [13] MCAFEE R P, MCMILLAN J. Auctions and bidding [J]. Journal of Economic Literature, 1987, 25(2): 699-738.
- [14] HUANG H Y, YEN D C. A secure reverse Vickrey auction scheme with bid privacy[J]. Information Sciences, 2006: 550-564.
- [15] AI Z, LIU Y, WANG X. ABC: an auction-based blockchain consensus-incentive mechanism[C]//Proceedings of the 2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS), Hong Kong, 2020: 609-616.
- [16] VICKREY W. Counterspeculation, auctions, and competitive sealed tenders[J]. The Journal of finance, 1961, 16(1): 8-37.
- [17] JOSANG A. A logic for uncertain probabilities [J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2001, 9(3): 279-311.
- [18] LIU W, WU H, MENG T, et al. AucSwap: a vickrey auction modeled decentralized cross-blockchain asset transfer protocol[J]. Journal of Systems Architecture, 2021, 117: 102102.

(下转第80页)

文章编号: 1007-1423(2023)10-0042-08

DOI: 10.3969/j.issn.1007-1423.2023.10.008

多关系型数据库向图数据库转换与同步方法

郭良银*

(西南交通大学制造业产业链协同与信息化技术支撑四川省重点实验室, 成都 611756)

摘要: 关系型数据库对于高关联性数据处理性能较低, 对查询而言就会产生大量的连接操作, 并且不适用于挖掘信息之间潜在的关联关系。针对以上的问题, 提出一种基于消息队列的多关系型数据库向图数据库转换与同步方法, 并进行设计与实现。首先定义数据的转换与同步方法, 依据先节点后边的原则对数据进行转换, 结合映射表以及触发器实现数据的同步, 使用JSON完成数据网络传输, 图查询语言接口操作图数据库。然后采用Java语言与RabbitMQ进行实现与测试。

关键词: 关系型数据库; 图数据库; 数据转换与同步; 消息队列; Cypher

0 引言

我们身处于一个互联的世界中, 互联网技术高速发展, 数据日益激增。关系型数据库在数据存储领域有着举足轻重的地位^[1], 然而在处理联系上做得并不好, 违背了其设计初衷——处理表格化结构。其在处理关系上主要通过主/外键连接的方式完成, 关系连通性的增强会造成连接操作的增加, 阻碍性能。我们经常需要对实体间的联系进行语义的区分, 并为其定义权重, 关系型数据库难以响应需求。而联系在图数据模型中是“一等公民”。在某社交网络^[2]查找“朋友的朋友的朋友”时, 关系型数据库表现出很严重的延迟, 而图数据库由于建模的不同, 不需要全表扫描, 采用免索引邻接的方式, 提升了响应的速度。因此, 将海量数据转存到图数据库中, 在图数据库上对关联数据进行实时查询、实时分析、挖掘潜在的关联关系是很有必要的^[3-6]。

目前主流的图数据库主要有Neo4j、GraphDB、BrightstarDB等^[7]。以数学中的图论作为理论基础, 主要的图模型有属性图和RDF(resource

description framework)两种, 属性图模型以Neo4j为代表, 由节点、边以及属性构成, 节点和关系均可以带有属性, 每个节点可以具有多个标签, 可以为边定义额外的语义。可以用四元组 $G = \langle V, E, P, T \rangle$ 进行表示^[8], 其中 G 代表属性图, V 表示节点, E 表示边, P 表示属性, T 表示标签。当前对数据库转换与同步技术已有不少研究, 鲁佳文等^[9]提出了一种对象关系型数据库到RDF的映射方法, 还充分考虑数据库的约束, 实现了一个转换的映射工具。Xu等^[10]提出了一种在关系型数据库模式和OWL本体之间创建映射的实用方法。Neo4j官方提供了将CSV文件导入图数据库中的方法, 但是对于多数据源的情况并不适用, 可能导致数据库中节点ID冲突的问题。当前对数据库同步技术的研究主要面向关系型数据库, Wang等^[11]利用中间件实现分布式数据库同步, 使得系统可以在本地处理冲突, 提升系统的性能。Wang等^[12]通过中心数据库的方式实现异构数据库集成方法。Xu等^[13]提出了一种基于阴影表法的异构数据库同步机制。Domingos等^[14]研究了移动设备的同步

收稿日期: 2023-01-04 修稿日期: 2023-01-23

作者简介: *通信作者: 郭良银(1997—), 男, 四川泸州人, 硕士, 研究方向为价值网与云服务平台技术, E-mail: 2941342224@qq.com

模型，提出了基于消息摘要的同步模型，尽量使得同步数据量最小化。但是均未针对图数据库。

本文利用触发器、同步服务与消息队列实现多关系型数据库向图数据库转换与同步方法，使得图数据库端服务与多关系型数据库服务之间的应用解耦，通过转换与同步方法的实现，可以在图数据库上对高关联性数据实现实时查询、实时分析、挖掘潜在的关联信息。

1 转换与同步方法

1.1 数据库转换

关系型数据库到图数据库的转换主要有两类问题：关系表与节点的转换，主/外键与边的转换。其中，数据类型的映射关系是需要解决的问题。本文采用JSON(JavaScript object notation)格式进行数据交换，需要将关系型数据库数据类型映射为JSON类型，然后解析JSON并应用于图数据库上的操作。JSON独立于语言，使用JavaScript语法来描述数据对象，其数据类型主要有数字、字符串、布尔值、数组、对象和空数据类型。例如，PostgreSQL对应的数据类型映射关系如表1所示。

表1 数据类型映射

PostgreSQL数据类型	JSON数据类型	图数据库数据类型 (Neo4j)
(none)	null	Null
boolean	true/false	Boolean
smallint, int, bigint	number	Integer
numeric/decimal	number	Float
char, varchar, text	string	String
array	array	List

当前图数据库的组成架构由接口层、计算层、存储层构成。计算层和存储层负责查询计算和数据存储，接口层作为架构顶层对外提供服务，提供了友好的操作方式，如API、SDK及可视化组件等，还提供了如Cypher、Gremlin、GSQL等主流的图查询语言接口^[15]。本文利用Cypher查询接口完成数据的转换与同步操作，依据先节点后边的转换原则，先将多关系型数据库表中的元组建立为节点，并保证其在图数

据库中的唯一性，然后将表之间的主/外键关系建立为边。

1.1.1 节点转换方法

关系型数据库中表数据一般对应于一个实体，由多行元组构成，其中每个元组一般都有唯一的主键标识。然而，同一个数据源不同的表之间可能存在主键相同的情况，在多个数据源的情况下节点唯一性问题更加明显，所以我们采用新建映射表的方式解决该问题，即 $T = \{old_key, new_key\}$ ，其中T表示映射表，old_key是数据库“表名+主键/复合主键”，new_key为分布式全局唯一ID。目前分布式ID生成解决方案主要有以下方式：

(1) 数据库主键自增。即统一在某个机器或者数据中心上，利用MySQL、PostgreSQL、Oracle等关系型数据库或者Redis、MongoDB等生成的满足唯一性要求的ID。

(2) UUID(universally unique identifier)。顾名思义，UUID表示通用唯一标识符，其包含了32个16进制数字，Java中提供了UUID.randomUUID()方法用于生成具有全局唯一性的ID。但是其采用字符串存储，空间消耗大。

(3) 雪花算法(Snowflake)。雪花算法是Twitter开源的分布式ID生成算法。结果是一个64位long型的ID。其核心思想是：使用41位作为毫秒数，10位作为机器的ID，其中5位表示数据中心，另5位表示机器ID，12位作为毫秒内自增流水号，意味着每个机器在每毫秒内可以产生4096个ID，最后还有一个符号位，永远是0。雪花算法生成ID速度比较快且有序递增，使用比较灵活。

本文采用雪花算法为图数据库中的节点生成唯一ID标识。同理，对于没有指定主键的表，可以根据表的结构特性选择具有唯一性约束的字段进行映射表的生成。统一生成了主键映射表以后，对节点进行转换操作。对于关系表中存在对其他表或当前表有引用关系的字段需要做额外处理，即根据主键映射表将其值设置为真正的引用ID。

1.1.2 边转换方法

关系型数据库中没有边的概念，我们根据数据表之间的主/外键约束为图数据库建立边。关系表之间的关系主要有一对一、一对多、多对多以及表的自关联几种情况。根据先节点后边的转换规则，我们已经对节点进行了正确的转换，所以仅需要指定必要的字段便可通过 Cypher 语句完成关系的构建。具体字段配置如下：

```

{
  "labelOfStartTable": "边起始标签",
  "labelOfEndTable": "边结束标签",
  "startTableAttr": "边起始属性名",
  "endTableAttr": "边结束属性名",
  "startTableAttrValue": "边起始属性值",
  "endTableAttrValue": "边结束属性值",
  "relationship": "关系类型",
  "relationshipAttr": {}
}

```

其中，边起始标签、结束标签以及关系类型与属性需要我们自定义完成。例如，电影从属于某个种类，我们可以使用“belong_to”定义关系类型。

1.2 数据库同步

目前对于关系型数据库的同步研究相对较多，而对于关系型数据库与图数据库的同步研究相对较少。数据同步主要针对数据的变更，当前对于增量数据的捕获方式主要有触发器法、日志法、API法等，各具优劣。并且不同的关系型数据库产品存在差异，相同的方法并非适用于所有的数据库产品。例如，对于MySQL而言，使用 canal 监听数据库 binlog 日志能够很好地完成增量数据的同步，但是其对于 oracle 并不是很适用。世面上也出现了不少的第三方数据库同步软件，例如，GoldenGate 商业软件与开源软件 Kettle 等。但是均没有针对图数据库。然而，当今主流的数据库均支持触发器机制。所以本文利用触发器机制生成同步表，如表 2 所示，然后结合消息队列实现多关系型数据库向图数据库同步。

表 2 数据库记录变更同步

字段名	字段类型	字段描述
id	int	同步表自增ID
table_name	varchar(30)	变更数据表名
table_fk	varchar(50)	变更数据标识
is_sync	boolean	是否完成同步
change_type	int	变更类型(0插入,1更新,2删除)

具体的，数据的同步分为对节点的同步和对关系的同步。需要同步的数据变更类型有插入、更改与删除三种。插入事件发生的时候，如果变更数据存在外键关联，需要根据“先节点后关系”的规则进行数据的同步，如果消息队列存在多个消费者，需要保证消息队列消费的顺序性。保证顺序性的方法有很多，内存队列就是其中的一种。对于消息队列的使用，还需要保证数据的可靠性，保证数据正确传输，防止数据的重复消费等，以 RabbitMQ 为例，具体如图 1 所示。删除事件发生的时候，如果删除的数据具有关联关系，那么根据“先关系后节点”的删除规则完成数据的同步。同理，对于数据更新可以拆分为删除与插入操作进行数据同步。

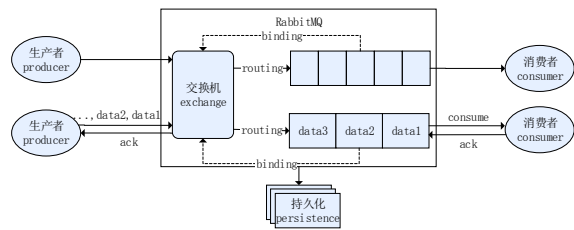


图 1 RabbitMQ 示意

2 设计与实现

基于以上对转换与同步方法的说明，接下来进行具体的实现与验证。对于数据的转换方法，始终依据“先节点后边”的规则进行，由于图数据库中每个节点需要唯一标识，这与关系型数据库中多表可以存在相同的主键产生了冲突，且当前面临的是多个关系型数据库，冲突更加严重。本文采用雪花算法 (SnowFlake) 为节点生成唯一的 ID 表示。对于边的转换方法，我们定义了具体的配置方法，图数据库端应用依据配置以及接收到的数据完成转换操作。对

于数据同步，我们采用触发器与消息队列的方式完成，无需更改原数据库端应用代码，直接在数据库底层对数据变更进行捕获。当图数据库应用服务升级优化时，转换与同步服务无需改动，实现了转换同步服务与图数据库服务之间的应用解耦。具体实现架构如图 2 所示。

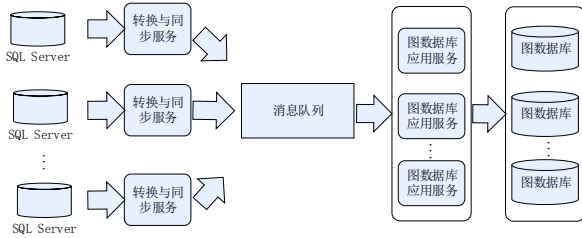


图 2 转换与同步架构

本文面向多关系型数据库，每个关系型数据库只与消息队列存在耦合关系，从而可以动态地加入与退出。同样地，图数据库应用服务仅作为消息队列消费者监听消息队列，进而完成数据的转换与同步。对于数据库转换而言，分为了节点的转换与边的转换。首先我们为数据库中的每条记录生成唯一的 ID 标识，用于图数据库中，并且记录下来，用作同步使用。首先遍历数据库中的数据记录，使用雪花算法对每条数据记录生成全局唯一的 ID 标识，然后存入主键映射表中。

对于节点的转换，需要遍历数据库中每张实体表，并根据主键映射表设置其全局 ID，如果存在外键关系，还需要根据映射表设置相应的属性，最后将数据发送到消息队列，从而完成节点数据的转换，具体流程如图 3 所示。

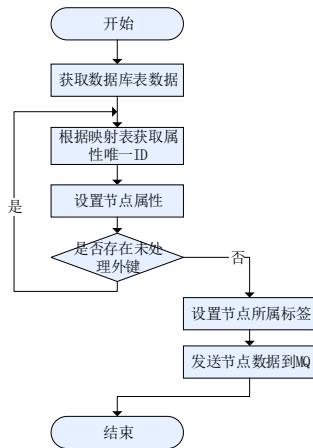


图 3 节点转换流程

对于关系的转换，需要明确指定关系的起始节点与终止节点，所以需要指定具体的节点标签、节点属性等配置信息，同样地，对数据库中存在的有关系都用相同的方式进行处理，最后作为消息队列的生产者，完成数据的发送，具体流程如图 4 所示。

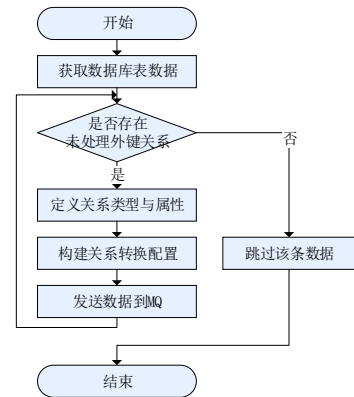


图 4 关系转换流程

对于数据库的同步，触发器与同步表的方式是目前主流关系型数据库均支持的一种形式，且本文利用消息队列完成解耦设计，只要同步服务能够作为生产者按照配置要求发送数据到消息队列，图数据库应用服务就能根据消息完成数据的同步操作。需要同步的数据变更类型有数据的插入、更新与删除。我们为数据库表编写对应的触发器。例如，在 PostgreSQL 中对 IMDB 数据集中的电影名称表(title)编写如下的触发器：

```

CREATE SEQUENCE "sync_table_seq"
INCREMENT 1
START 1
NO MINVALUE
NO MAXVALUE
CACHE 1;
--插入触发器
CREATE OR REPLACE FUNCTION title_insert_
trigger_fun()
RETURNS TRIGGER AS $movie$
BEGIN
INSERT INTO sync_table VALUES (nextval
('sync_table_seq'),'title', NEW.id, 0, false);
RETURN NULL;
END;

```

```

$movie$ LANGUAGE plpgsql;
CREATE TRIGGER title_insert_trigger
AFTER INSERT ON title
FOR EACH ROW EXECUTE PROCEDURE title_in-
sert_trigger_fun();
--更新触发器
CREATE OR REPLACE FUNCTION update_
title_trigger_fun()
RETURNS TRIGGER AS $movie$
BEGIN
INSERT INTO sync_table VALUES (nextval
('sync_table_seq'), 'title', OLD.id, 1, false);
RETURN NULL;
END;
$movie$ LANGUAGE plpgsql;
CREATE TRIGGER update_title_trigger
AFTER UPDATE ON title
FOR EACH ROW EXECUTE PROCEDURE up-
date_title_trigger_fun();
--删除触发器
CREATE OR REPLACE FUNCTION delete_title_trig-
ger_fun()
RETURNS TRIGGER AS $movie$
BEGIN
INSERT INTO sync_table VALUES (nextval
('sync_table_seq'), 'title', OLD.id, 2, false);
RETURN NULL;
END;
$movie$ LANGUAGE plpgsql;
CREATE OR REPLACE TRIGGER delete_title_trigger
AFTER DELETE ON title
FOR EACH ROW EXECUTE PROCEDURE de-
lete_title_trigger_fun();

```

触发器建立好了以后，当数据库表数据产生了变更，就会在同步表中插入对应的数据，转换与同步服务根据同步表中的数据完成数据的同步，具体流程如图 5 所示。

需要注意的是，数据同步对于节点和关系的变更存在顺序性要求。例如，对于数据库数据的插入，如果存在对关系的插入，我们需要依据“先节点后边”的插入规则完成同步；对于删除数据，如果同时涉及到节点和关系的变更，我们需要依据“先关系后节点”的删除规则完成数据同步。所以需要保证消息的顺序性消费，对于 RabbitMQ 而言，我们可以采用单队

列模式或者将同步消息发送到一个队列的方式实现，对于 Kafka 消息队列而言，可以将数据发送到同一个 partition，然后内部保证顺序消费。在图数据库应用服务端，作为消息队列的消费者，获取发送者发送的数据，然后通过图数据库查询语言完成相应的转换与同步操作。

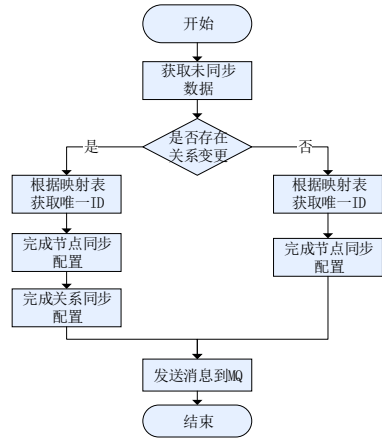


图 5 数据同步流程

3 结果验证与分析

本实验基于 Windows 64 位操作系统，512 GB SSD，机带 RAM 16.0GB、处理器 Intel(R) Core (TM) i5-10300H CPU @2.50GHz。软件环境为 Java 11.0.17、neo4j 4.4.16、RabbitMQ 3.9.11、PostgreSQL 15.1。数据集采用互联网电影资料库 (Internet Movie Database, 简称 IMDb)，获取其数据并将其转储到 PostgreSQL 中，进行转换的数据 E-R 图如图 6 所示。

分别对上面的表数据进行转换操作，均能完成转换，成功转换后形成 284263 个节点、279653 条边。转换后执行语句 “match p=(:movie_info_idx)-[*1..4]-(:movie)-[*1..4]-(:kind_type) return p limit 50”，查询得到结果示意图如图 7 所示。

对 PostgreSQL 执行语句 “insert into title values (3000000, 'The Quiet Mutiny', null, 7, null, null, 'Q3535', 1597940, null, null, null, 'dc8jljojojohoe')”，通过消息队列能够成功地完成数据插入同步，在图数据库端查看结果如图 8 所示。

对 PostgreSQL 执行语句 “update title set kind_id = 2 where id = 3000000”，在图数据库端查看更新数据同步结果如图 9 所示。

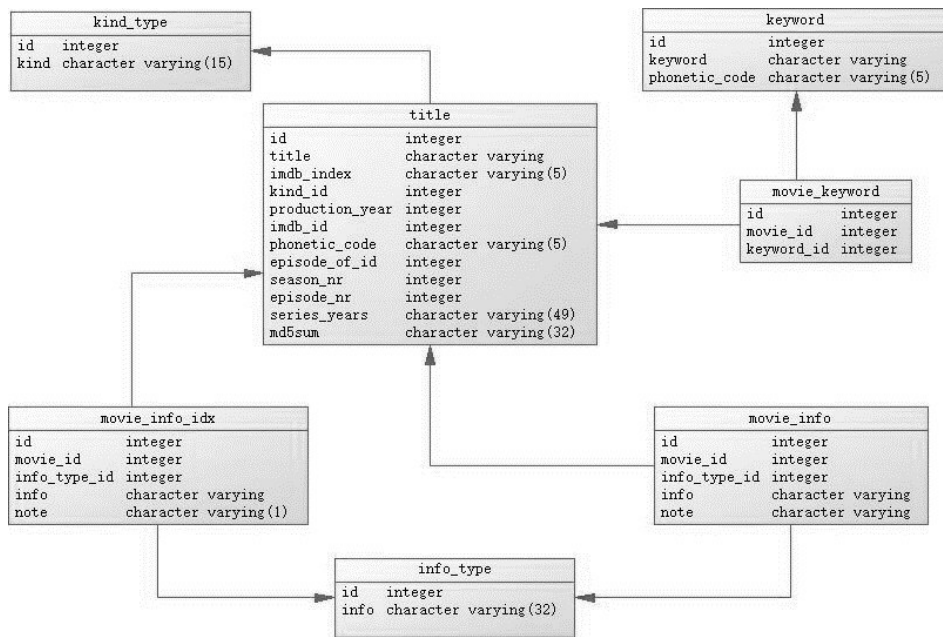


图6 IMDb数据集部分E-R图

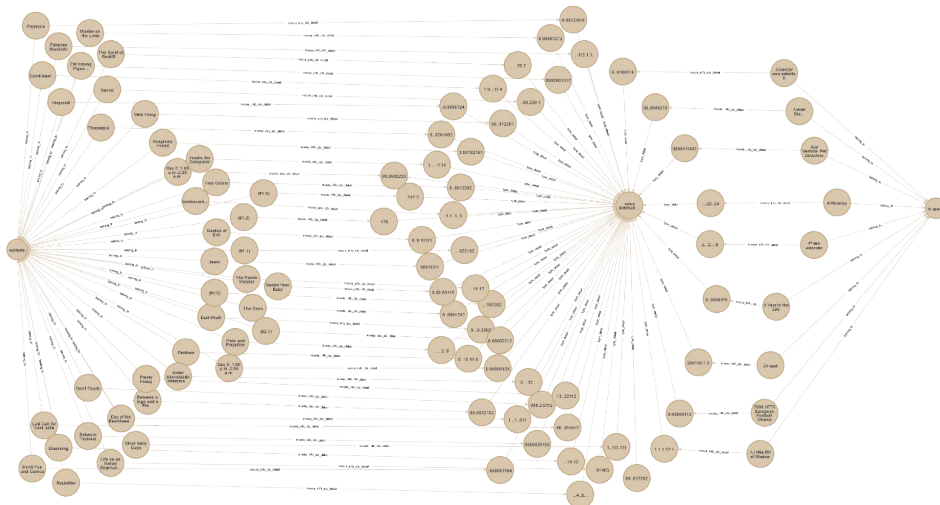


图7 部分转换结果示意

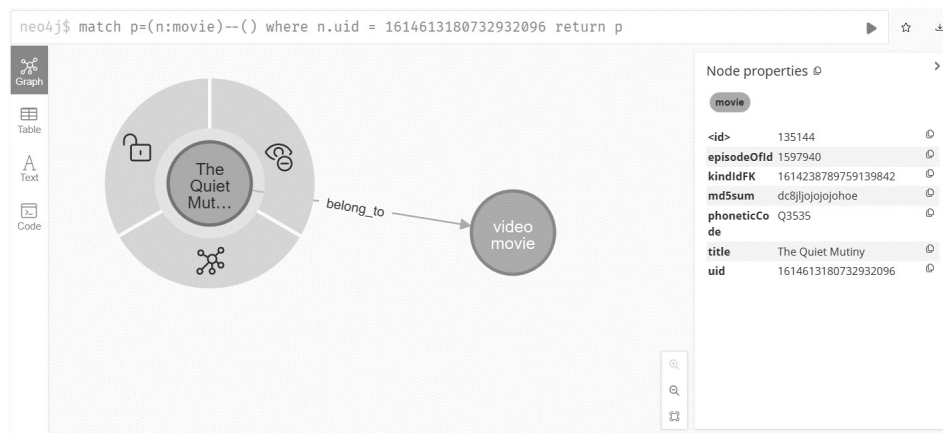


图8 插入数据同步结果示意

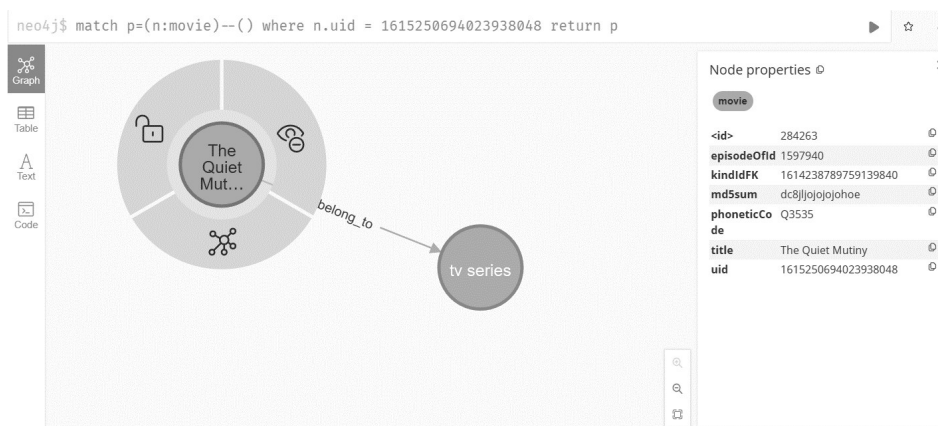


图9 更新数据同步结果示意

同理，对于删除操作均能够正常地完成数据同步，我们可以在图数据库上进行后续的关联查询与近实时分析。因此，对于关系型数据库向图数据库的转换与同步研究具有一定的实际意义。

4 结语

本文设计并实现了多关系型数据库向图数据库的转换与同步方法，利用消息队列实现应用解耦，实现了关系表与图数据库中节点与关系的转换与同步，能够更加方便地对关联性数据进行查询和分析，挖掘潜在的关联信息。具有一定的实际意义。但本文方法具有一定的局限性，对于实时性要求比较高的应用场景可能使得系统压力变大，仍需进一步改进。

参考文献:

- [1] 向海华. 数据库技术发展综述[J]. 现代情报, 2003(12):31-33.
- [2] 李桃陶, 周斌, 王忠振. 基于社交网络的图数据挖掘应用研究[J]. 计算机技术与发展, 2014, 24(10): 6-11.
- [3] 陆鹏. 基于Neo4j的大数据组织检索研究与应用[D]. 南京: 东南大学, 2015.
- [4] 王鑫, 邹磊, 王朝坤, 等. 知识图谱数据管理研究综述[J]. 软件学报, 2019, 30(7): 2139-2174.
- [5] ZHAO P, SHOU L D, CHEN K, et al. Storage and query model for localized search on temporal graph data[J]. Computer Science, 2019, 46(10): 186-194.
- [6] PAUL S, MITRA A, KONER C. A Review on Graph Database and its representation[C]//Proceeding of the 2019 International Conference on Recent Advances in Energy-Efficient Computing and Communication(ICRAECC). IEEE, 2019: 1-5.
- [7] DRAKOPOULOS G, GOURGARIS P, KANAVOS A. Graph communities in Neo4j: four algorithms at work[J]. Evolving Systems, 2020, 11: 397-407.
- [8] 李俊逸, 王卓, 马鹏玮. 图数据库技术发展趋势研究[J]. 信息通信技术与政策, 2021, 47(5): 67-72.
- [9] 鲁佳文, 严丽. 对象关系型数据库到RDF(S)的映射方法[J]. 计算机科学, 2021, 48(10): 145-151.
- [10] XU Z, ZHANG S, DONG Y. Mapping between relational database schema and OWL ontology for deep annotation[C]//Proceeding of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings) (WI'06). IEEE, 2006: 548-552.
- [11] WANG J, ZHANG D S. Research and design of distributed database synchronization system based on middleware[C]//Proceeding of the 2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA). IEEE, 2015: 685-688.
- [12] WANG Y, WANG X, RAO X. An heterogeneous database synchronization update base on web services [C]//Proceeding of the 2015 6th IEEE International Conference on Software Engineering and Service Science(ICSESS). IEEE, 2015: 166-169.
- [13] XU F, MA L, WEN F, et al. Design and implementation of a heterogeneous relational database synchronization mechanism based on tree distribution architecture [C]//Proceeding of the 2019 6th International Conference on Information Science and Control En-

- gineering(ICISCE). IEEE,2019:234-240.
- [14] DOMINGOS J, SIMÕES N, PEREIRA P, et al. Database synchronization model for mobile devices [C]//Proceeding of the 2014 9th Iberian Conference on Information Systems and Technologies (CISTI). IEEE,2014:1-7.
- [15] ANGLES R, ARENAS M, BARCELÓ P, et al. Foundations of modern query languages for graph databases [J]. ACM Computing Surveys (CSUR), 2017,50(5):1-40.

Conversion and synchronization method from multi-relational database to graph database

Guo Liangyin*

(Manufacturing Industry Chains Collaboration and Information Support Technology Key Laboratory of Sichuan Province, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: Relational databases have low processing performance for high-relationship data, and will generate a large number of connection operations for queries, and are not suitable for mining potential correlations between information. Aiming at the above problems, a message queue-based multi-relational database to graph database conversion and synchronization method is proposed, and the design and implementation are carried out. First define the data conversion and synchronization method, convert the data according to the principle of "node first, then edge", combine the mapping table and trigger to realize data synchronization, use JSON to complete the data network transmission, and use the graph query language interface to operate the graph database. Then use Java language and RabbitMQ to implement and test.

Keywords: graph database; relational database; data conversion and synchronization; message queue; Cypher

(上接第33页)

Research on application of intelligent and real-time weeding system for books in the construction of smart library

Luo Ke*

(Library of Shaoyang University, Shaoyang 422000, China)

Abstract: In the management of smart libraries, the process of weeding is a crucial component. To enhance the efficiency of weeding and alleviate the workload of librarians, a real-time intelligent weeding system based on OpenCV technology and Convolutional Neural Networks (CNN) is proposed in this study. The system utilizes OpenCV technology to segment barcode images of books and utilizes CNN to recognize the barcode images, comparing them with a database and using the Analytic Hierarchy Process to calculate the weeding weight values of books. Finally, the results are fed back to the frontend based on a predefined threshold. The design and implementation of the system effectively improves the efficiency of book weeding and provides an innovative idea for the intelligent service of university libraries.

Keywords: smart library; book weeding; hierarchical analysis; CNN

文章编号: 1007-1423(2023)10-0050-07

DOI: 10.3969/j.issn.1007-1423.2023.10.009

水声网络路由协议实现与湖试

张振儒¹, 杜秀娟^{1,2,3*}

(1. 青海师范大学计算机学院, 西宁 810000; 2. 青海省物联网重点实验室, 西宁 810000;

3. 藏语智能信息处理及应用国家重点实验室, 西宁 810000)

摘要: 水声网络(underwater acoustic networks, UANs)具有拓扑结构易变、低带宽、高延时、高误码率和能量衰减快等特性。基于层级的水声网络路由协议(Level Based Adaptive Geo Routing for underwater sensor network, LB-AGR)倾向于选取剩余能量较高的节点作为下一跳, 能够适应、甚至在某种程度上优化动态拓扑, 均衡网络的能耗, 从而延长整个网络的寿命。对水声网络路由协议进行研究, 在实现LB-AGR路由协议的基础上, 在青海湖搭建试验床, 通过多组水下源节点发送图片数据, 图片数据经过多跳LB-AGR路由到Sink节点, 之后通过互联网发送到远程服务器的方式对LB-AGR路由协议性能进行测试, 结果表明: LB-AGR协议实现了图片数据实时、可靠的传输, 平均交付率不低于92%。

关键词: 水声网络; 平均交付率; 路由协议; 湖试

0 引言

水声网络在海洋环境保护、水下救援、海洋灾难预防、海洋资源勘测等领域有重要支撑作用^[1-6]。根据UANs具有拓扑易变、高能耗、能量衰减快、低带宽、高延时等固有特性, 国内外高校以及科研机构相继提出了以数据传输路径、网络的拓扑结构、是否需要地理位置信息、有无层级等类别的路由协议^[7-9]。

由于搭建真实的水下网络平台并且进行试验困难较大, 因此大多数UANs路由协议在开发过程中都是基于网络仿真平台进行性能测试, 而复杂的水下通信系统和动态的水下环境并不能完全通过仿真软件模拟^[10-11]。我们在编程实现基于层级的水声网络路由协议(Level Based Adaptive Geo Routing for underwater sensor network, LB-AGR)的基础上, 搭建试验床, 对LB-AGR协议的可靠性进行验证。

1 LB-AGR协议机制

1.1 LB-AGR协议概述

LB-AGR路由协议中提出了转发因子的概念, 节点通过邻居表中的候选节点信息(剩余能量、节点密度和位置等)计算转发因子, 从而确定最佳下一跳。

在LB-AGR协议中, Sink节点会周期性发送泛洪控制报文, 格式见表1, 其中包括层级、发送节点ID、位置信息和包类型等信息。

在上行路由过程中, 节点会根据收到包头部的层级字段, 在其邻居表中搜索具有转发资格的下一跳节点, 并根据剩余能量和节点密度为每个具有资格的转发节点计算转发因子, 从而确定最佳的下一跳节点。

寻找最佳下一跳阶段需要的邻居节点的信息, 剩余能量、节点状态等信息通过扫描邻居表获取, 每个传感器节点维护一张邻居表。图1为邻居表结构体的定义。

收稿日期: 2023-01-09 修稿日期: 2023-02-28

基金项目: 国家自然科学基金(61962052);青海省物联网重点实验室(2022-ZJ-Y21)

作者简介: 张振儒(1998—), 男, 甘肃张掖人, 硕士, 研究方向为无线网络; *通信作者: 杜秀娟(1970—), 女, 河北石家庄人, 博士, 教授、博士生导师, 从事领域为无线网络与安全、物联网技术, E-mail: dxj@qhnu.edu.cn

表 1 控制报文格式

Bit位	8	8	8	2	...	1	8	1	16	4	16	8	变长	16
字段	层级	发送节点 ID	目的节点 ID	类型=0 (0:DATA 1: ACK 2: RTS 3:CTS)	...	流向 0:下行 1:上行	sink 节点 ID	(源节点/目的) 0:位置 1:节点ID	位置或 ID (源/目的) 全“1”为 广播	应用优先级 (应用协议 类型)	包 ID	负载 长度	数据	校验
	上一跳节点		下一跳 节点					源或目的节点						
位置	头部												负载	尾部

```
typedef struct Neighbor
{
    unit8_t id;          // 节点id
    unit8_t level;      // 节点层级
    unit8_t ap;         // 节点剩余能量
    POS    pos;         // 节点位置
    unit8_t status;     // 节点状态
}Neighbor_attribute_(aligend(1));
```

图 1 邻居表结构体定义

1.2 LB-AGR 协议初始化机制

LB-AGR 协议通过 Sink 节点周期性的泛洪控制报文，完成网络初始化机制。此过程中所有节点通过从泛洪报文获取自身层级以及邻居节点等信息的方式完成网络初始化。Sink 节点泛洪的控制报文头部信息包括：发送节点 ID、发送节点层级、发送节点的剩余能量等信息。基于层级的定向泛洪机制保证了泛洪控制报文能够沿着从 Sink 节点到远处传感器节点的下行路径传输，使得网络中的节点都可以接收到泛洪控制报，而不会像传统泛洪机制那样造成过多的开销。LB-AGR 路由协议网络初始化算法伪代码如算法 1。

算法 1 LB-AGR 路由协议网络初始化算法

```
1: L_sink =0;          //Sink 节点层级设置为 0
2: L_node =255;      //普通传感器节点层级设置为 255
3: upon 普通节点接收到报文 do
4:   if (报文==Hello 报文)      //判断报文类型
5:     if(L_node ==255)
```

//普通节点第一次收到 Hello 报文

```
6:   L_node =L_send+1;
7:   节点在更新报文后继续泛洪此报文;
8:   else          //普通节点以获取自身层级
9:     if (age_time>0&&L_node<L_send)
10:      节点信息不做更新，丢弃此 Hello 包;
11:    else        //节点移动后层级信息发生变化
12:      L_node=L_send+1;
13:      节点在更新报文后继续泛洪此报文;
14:    end if
15:  end if
16: end if
17: end upon
```

1.3 LB-AGR 协议上行路由机制

LB-AGR 协议主要侧重于向上行路由传感器节点收集发送的数据。在上行路由转发过程中，为了使节点能量均衡，最大限度地提高网络的寿命，会根据式(1)所示的路由算法，将能量、节点密度等参数引入算法，确定最佳下一跳。

$$\alpha_{desira}^i = \alpha_1 \cdot \frac{Density_i}{\sum_i Density_i} + \alpha_2 \frac{AP_i}{AP_{init}} \quad (1)$$

其中： α_1, α_2 的取值范围在0到1之间、 $\alpha_1 + \alpha_2 = 1$ ， $\alpha_2 = \frac{\max_i (AP_i) - \min_i (AP_i)}{AP_{init}}$ ；公式中 α_{desira}^i 表示每个候选节点引入的综合转发因子， $Density_i$ 表示节点 i 的上行邻居节点数、 i 表示候选节点 $node_i$ 、 AP_i 表示节点 $node_i$ 的剩余能量、 AP_{init} 表示节点的初始能量^[12]，最后，通过 $\max\{\alpha_{desira}^i\}$ 确定唯一的最佳下一跳节点。

图2为上行数据传输流程，当节点收到上行

数据包时，首先会根据数据包头部信息更新本节点邻居表，其次根据数据包头接收节点字段判断自身是否为接收节点。如果接收节点是本节点，则会基于邻居表中层级、剩余能量、节点密度等信息计算每个候选节点的综合转发因子，从而确定具有转发资格的最佳下一跳节点。若接收节点不是本节点，不再进行任何操作。

在确定最佳下一跳的过程中，可能会存在多个候选下一跳节点的现象，若这些候选节点都参与转发同一个分组，将会导致大量的冲突与重传。因此，为了减少冲突和能耗，LB-AGR从候选节点 $node_i$ 中找出最佳下一跳节点。

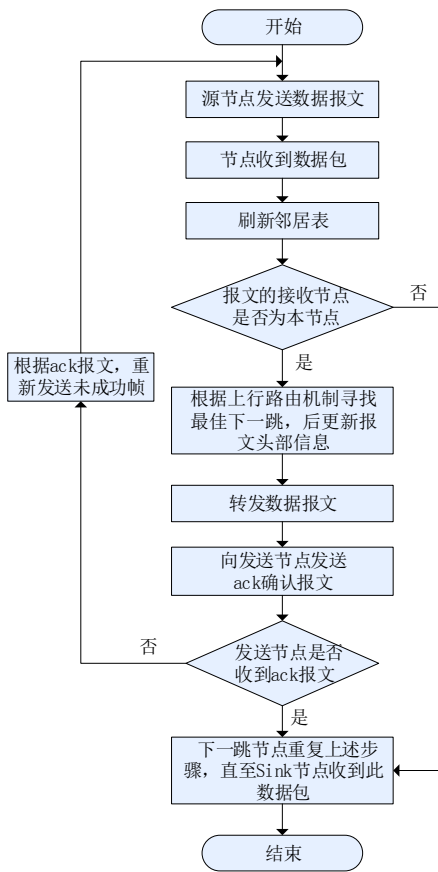


图2 上行数据传输流程

2 试验前期准备工作

2.1 试验设备介绍

青海湖试验床使用树莓派嵌入式开发板、OFDM水声Modem、电源系统、4G转WIFI工业路由器、远程服务器多套设备搭建，表2为主要

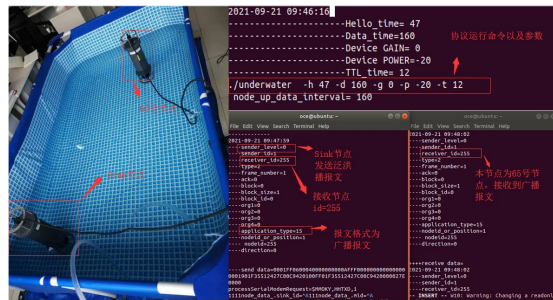
设备清单。

表2 设备清单

设备名称	设备型号	设备功能
AquaSeNTOFDM	AMN-13	水声信号调制解调功能
树莓派开发板	3B	路由协议运行载体
工业路由器	LTE 3G/4G	将Sink节点接入到互联网中
远程服务器	Ubuntu 16.06.1	存储并处理UANs获取的数据信息

2.2 前期协议调试与运行

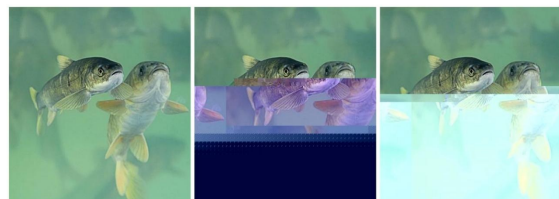
为保证青海湖试验床的实验正常有序进行，团队在学校物联网实验室内通过部署的两个水下传感器节点对协议和硬件设备进行了测试。图3为实验室调试节点，在试验床内进行测试实验，实验过程中通过改变广播包周期、增益、发送功率等控制信息进行测试。



(a) 节点入水测试 (b) 网络运行实时情况

图3 实验室调试节点

在调试过程中，当服务器端收到来自Sink节点的数据包后，可以将数据恢复为图片，恢复图片大小比初始图片大100 byte(原数据为3900 byte)，偶尔还会出现恢复数据为发送数据两倍的情况，导致数据恢复后图片打不开、出现错位、颜色失真等现象。图4为图片发送前与接收恢复后的对比图。



(a) 原数据图 (b) 失真例图1 (c) 失真例图2

图4 图片发送前与接收恢复后对比

针对出现的错误，经过多次数据传输实验与分析后总结如下：

(1) 实验过程中保存到开发板的数据标记时间不准确，可能会导致后期的湖试数据无法使用。

登录到开发板后发现，开发板的系统时间在每次登录后都会默认为上次关机时间，解决方法为：设置开发板系统时间模式为开机后从网络自动获取北京时间。在接下来的多次测试中，所有的数据标记时间都可以正确标记。

(2) 在分析实验数据时发现，当源数据节点发送第二个数据块后，Sink节点又再次接收到了第一个数据块的最后一帧数据。

出现此现象原因为：实验室中试验床水深以及大小有限，两个Modem通信距离较近会出现回声效应或多径效应，导致水声信号紊乱(后期在青海湖试验床内进行试验时，未出现类似问题)。

在实验室水池调试成功后，接下来将所有实验设备密封后搬运至青海湖区域进行进一步实验，验证路由协议的功能，评测协议的性能。

3 湖试与结果

3.1 试验床搭建

根据实地环境，本次试验床布置共使用四个传感器节点、一台远程服务器主机、两艘快艇构成。每套传感器节点包括树莓派控制板、水下声学调制解调器、水上浮标、锂电池和缆线组成。

为了在实验中得到准确实验数据，在所有设备布置在水下前，我们根据情况定义了Sink节点和源节点，且源节点中数据为设置的特定数据(一张水下图片)。图5为试验床部署场景。



(a) 岸上节点控制图 (b) 母船节点控制图

图5 试验床部署场景

3.2 网络拓扑结构以及实验参数

本次实验使用的水下传感器节点编号分别

为1号、65号、5号、36号，其中1号设置为Sink节点。实验参数见表3。

表3 实验参数

参数	参数值
Sink节点广播周期/s	47
增益/dB	0
数据包间隔/s	60
发送功率/dB	-20
深度/m	2
通信半径/m	1000

根据青海湖水下环境以及天气情况，经过多次尝试最终确立了如图6所示的实验拓扑结构。实验过程中，65号源节点发送数据包后，5号节点或者36号节点转发，最后由1号Sink节点接收并且通过互联网发送到远程服务器对数据进行存储。

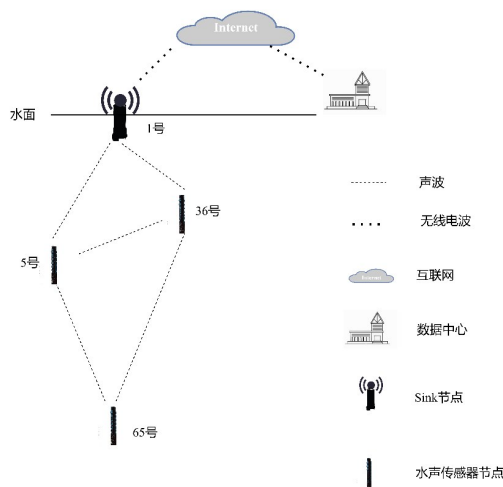


图6 试验床拓扑结构

3.3 实验过程中出现的问题以及解决过程

湖试过程中遇到了以下问题，下面进行分析。

首先，实验过程中码头上大型邮轮的轮机对水下声道有所影响。根据现场情况，所有实验组改变节点的发送功率，尽可能降低影响。

其次，由于青海湖水下暗流较多，两组节点在预先设定的实验点部署一段时间就会发生位移，使实验组随着暗流漂移到水深较浅的岸边。为此，经过多次验证我们选择了合适的水面环境，即使节点发生位移也会在不触碰湖底的情况下完成数据发送，充分说明LB-AGR协

议在 UANs 网络拓扑结构发生改变的情况下可以完成数据传输工作。图 7 为湖试期间发送与恢复数据对比图，图片未存在明显变化。



(a) 原数据图 (b) 接收恢复后数据图

图 7 湖试期间发送与恢复数据对比

3.4 试验结果分析

实验过程中，根据预先设定的数据参数进行了多个周期的多组实验，为了保证后期对实验数据的分析，实验过程中所有的输入信息与输出信息都会实时保存到开发板根目录的 data 文件夹下。图 8 为屏幕打印输出信息。

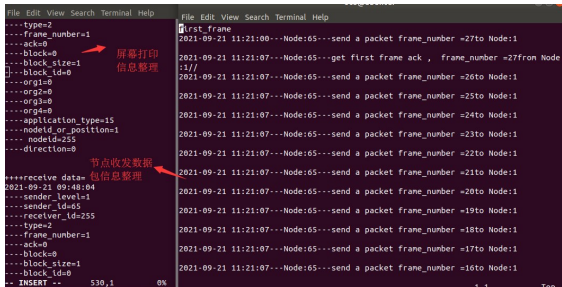


图 8 试验屏幕打印输出信息

本文对 LB-AGR 协议进行性能评估指标的定义如下。

(1) 包交付率(packet deliver ratio, PDR). 是指一个数据发送周期内，Sink 节点成功收到的数据包数与源节点发送的总数据包数之比，计算公式如下：

$$PDR = \frac{\sum_{i=1}^n P_r}{\sum_{i=1}^n P_s} \% \quad (2)$$

式中： P_r 为 Sink 节点成功收到数据包的数量， P_s 为源节点发送的总数据包数量， n 为数据发送周期的次数。

(2) 平均端到端延时(average end-to-end delay, AEED)。是指一个数据发送周期内，Sink 节点成功收到数据包的时刻与源节点发送

数据包的时刻的差值的总和与发送数据发送周期次数的比值。计算公式如下：

$$AEED = \frac{\sum_{i=1}^n (T_{Sink} - T_{source})}{n} \quad (3)$$

式中： T_{Sink} 为 Sink 节点成功收到数据包的时刻， T_{source} 为源节点发送数据包的时刻， n 为数据发送周期。

试验完成后，根据试验数据得到了发送周期与平均端到端延时、包交付率的关系，数据关系如图 9、图 10 所示。

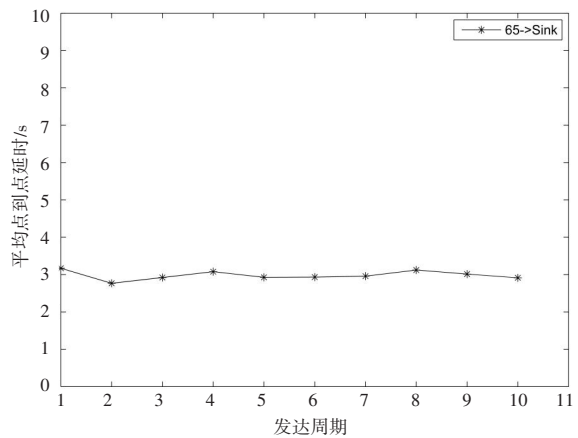


图 9 发送周期与平均端到端延时

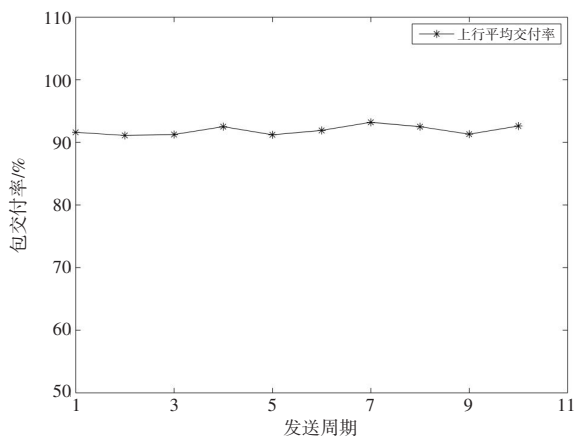


图 10 发送周期与包交付率

图 9 中，在不同的发送周期内，LB-AGR 路由协议的平均端到端延时都稳定在 3 s 左右。图 10 中，不同发送周期内，LB-AGR 路由协议的交付率都在 92% 以上。Peng 等^[13]在水下环境较为良好的定制水箱内，通过布置的四个水声节点进行基于矢量的转发路由协议(VBF)测试试验，

得到数据包的平均端到端延时为 13.4 s，包交付率为 90%。对比在水箱内进行试验，本次在水下环境较为复杂的青海湖进行的 LB-AGR 路由协议试验较为成功，源节点可以将图片数据及时可靠地发送到 Sink 节点。

对第一个数据发送周期内的情况进行进一步分析。如图 11 所示，此数据发送周期内上行流量交付率都在 65% 以上，第 8 组到第 13 组的交付率波动较大。这是由于实验过程中，UANs 试验床范围内的湖浪较大，使得节点随着湖浪的流向而移动，导致数据传输链路不稳定，且实验过程中经过游轮的轮机在水下产生噪声，增加了水声信号在传输过程中丢失或发生错误的概率。

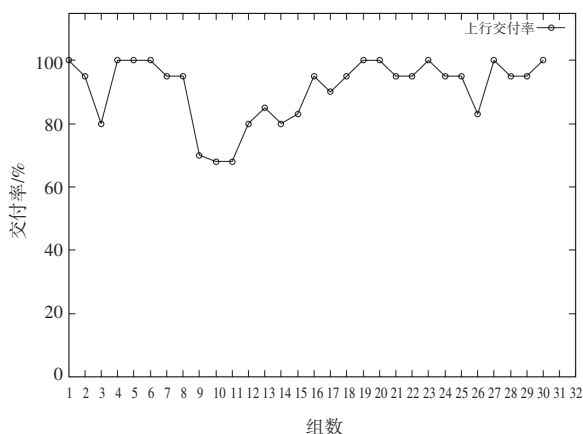


图 11 交付率随组数的变化

第一个数据发送周期内的端到端延时变化如图 12 所示，当 65 号节点通过多跳发送数据到 Sink 节点的平均端到端延时为 3.17 s，最大端到端延时为 4.87 s，最小为 2.61 s。数据多跳传输过程中，中间节点通过存储转发的方式进行路由，由于 UANs 拓扑结构易变、高误码率等特性导致多跳传输数据过程中每一跳的丢包和重传发生概率较高，因此端到端交付率存在 2.62 s 的起伏。

本次 UANs 试验网络中，未安装节点能耗记录设备，无法评估本次试验的能耗情况。后期会根据此情况，对 UANs 各节点配置能耗监控设备从而解决本次试验中无法精确获取 UANs 系统能耗情况的问题。

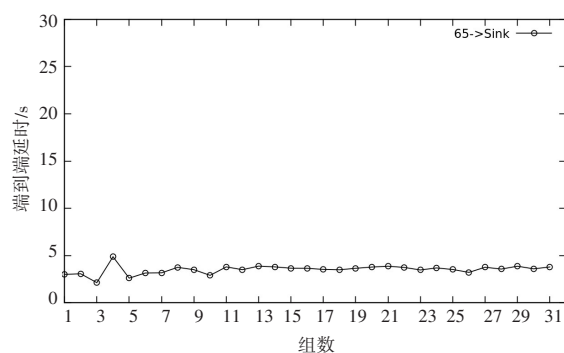


图 12 端到端延时随组数的变化

4 结语

本文基于 LB-AGR 协议，对此协议进行了编程实现以及在青海湖进行水声网络试验。通过分析实验数据，本次基于 LB-AGR 协议青海湖水声网络通信试验完成度较高。LB-AGR 协议平均端到端延时为 3.17 s，平均交付率为 92.1%。根据试验环境、试验设备等方面综合分析，本次青海湖 UANs 试验中，LB-AGR 协议实现了图片数据实时、可靠的传输，平均交付率不低于 92%。在接下来的研究中，我们会结合本次实践所得的数据与经验，针对本次湖试出现的问题逐步进行解决，提出更加优越的路由方案。

参考文献：

- [1] 夏娜, 闻心怡, 杜华争, 等. 基于水平集的围猎式水下传感器节点布置研究[J]. 计算机学报, 2021, 44(5): 1024-1036.
- [2] 陈惟一. 水声网络通信环境仿真与软件实现[D]. 哈尔滨: 哈尔滨工程大学, 2020.
- [3] DEEPAK S P, MUKESHKRISHNAN M B. Secured route selection using E-ACO in underwater wireless sensor networks [J]. Intelligent Automation & Soft Computing, 2022, 32(2): 963-978.
- [4] KHASAWNEH A M, KAIWARTYA O, KHALIFEH A, et al. Green computing in underwater wireless sensor networks pressure centric energy modeling [J]. IEEE Systems Journal, 2020, 14(4): 4735-4745.
- [5] CERQUEIRA L S, VIEIRA A B, VIEIRA L F M, et al. A cooperative protocol for pervasive underwater acoustic networks[J]. Wireless Networks, 2021, 27(3): 1941-1963.
- [6] 上义, 杜秀娟. 水声网络 MAC 协议综述[J]. 现代计

- 算机, 2021, 27(30):11-16.
- [7] 吴名星, 康松林, 陶志勇, 等. 采用转发优先级的水下传感网络的机会路由[J]. 传感技术学报, 2019, 32(1):133-138.
- [8] KUMARI S, MISHRA P K, ANAND V. Fault resilient routing based on moth flame optimization scheme for underwater wireless sensor networks [J]. *Wireless Networks*, 2020, 26: 1417-1431.
- [9] 杜秀娟, 苏毅珊. 水下传感器网络研究[M]. 北京, 科学出版社, 2016.
- [10] ANITHA D, KARTHIKA R A. DEQLFER: A Deep Extreme Q-Learning Firefly Energy Efficient and high performance routing protocol for underwater communication [J]. *Computer Communications*, 2021, 174:143-153.
- [11] ANWAR K, ATIQU R, MAHD Z. Modem design for underwater acoustic networks: taxonomy, capabilities, challenges, applications and future trends [J]. *Journal of Intelligence & Fuzzy Systems*, 2020, 39(6): 8161-8171.
- [12] 杜秀娟, 苏毅珊. 水下传感器网络研究[M]. 北京, 科学出版社, 2016.
- [13] PENG Z, CUI J H, WANG B, et al. An underwater network testbed: design, implementation and measurement [C] // *Proceedings of the Second Workshop on Underwater Networks*, New York, NY, USA, 2007: 65-72.

Realization and test in lake of routing protocol for underwater acoustic networks

Zhang Zhenru¹, Du Xiujuan^{1,2,3*}

(1. Computer Department, Qinghai Normal University, Xining 810000, China;

2. Qinghai Provincial Key Laboratory of IoT, Qinghai Normal University, Xining 810000, China;

3. The State Key Laboratory of Tibetan Intelligent Information Processing and Application, Xining 810000, China)

Abstract: Underwater Acoustic Networks (UANs) are characterized by variable topology, low bandwidth, high latency, high bit error rate and fast energy decay. Level Based Adaptive Geo-Routing for underwater sensor network (LB-AGR) selects the node with higher residual energy as the next hop, which can adapt to, or even optimize to some extent, the dynamic topology and balance the energy consumption of the network, thus extending the entire lifetime of the network. The routing protocol for UANs is studied, and realizes the LB-AGR routing protocol. Further, we build a test bed in Qinghai Lake, in which picture data are sent from multiple underwater source nodes, and the data are routed to a sink node through multi-hop with LB-AGR protocol, then the data are sent to a remote server through the Internet. The experiment results show that the LB-AGR protocol achieves real-time and reliable transmission of image data with an average delivery rate of no less than 92%.

Keywords: UANs; average delivery rate; routing protocol; test in lake

文章编号: 1007-1423(2023)10-0057-06

DOI: 10.3969/j.issn.1007-1423.2023.10.010

数据元治理框架及建议

程红云, 方亮, 陈美*

(上海宝信软件股份有限公司, 上海 201203)

摘要: 数据元作为数据标准, 既是数字化建设的基础工作, 同时又是构建高质量数据资产的关键。当前数据元建设过程中, 如数据元重复、质量低、建设不规范、工作效率低等问题突出。通过研究国内外数据元研究成果, 提出了数据元治理框架及建议, 旨在为业界数据元建设提供参考。

关键词: 数据元; 数据元治理; 治理框架; 数据元建设

0 引言

随着大数据提升到国家战略高度, 同时继土地、劳动力、资本、技术之后, 数据被列为第五大生产要素, 整个社会正从信息化时代迈入数字化时代, 目前大数据已经融入到社会的方方面面。从信息化到数字化, 需要实现数据互联、互通、互融, 与此同时对数据标准的需求和重视也日益增强。

数据元, 又叫数据元素, 在文献[1]中, 数据元被定义为“由一组属性规定其定义、标识、表示和允许值的数据单元”。数据元作为数据标准, 在逻辑模型设计中用于统一定义和规范数据模型的属性, 在物理模型设计中基于逻辑模型建设数据表或者直接建表引用数据元, 实现数据语言统一。

目前在建设数据元过程中, 普遍出现数据元重复、质量低、建设不规范、工作效率低等问题, 为数据元建设带来一定的挑战, 为此本文提出了数据元治理框架及建议。

1 数据元工作进展

1.1 研究现状

1.1.1 数据元理论研究

数据元的基础理论是来源 ISO/IEC JTC 1/SC

32 数据管理与交换分技术委员会制定的元数据注册系统标准系列, 数据元作为管理数据的基本单位, 从数据元的基本模型、概念模型、管理模型、命名规范、注册等^[2-7], 给出了比较全面的理论框架。对于我国, 全国信息技术标准化技术委员会(SAC/TC28)基于 ISO/IEC 成果, 等同制定了元数据注册系统标准系列^[8-13]。SC 32 的数据元成果已经成为全球共识的理论基础。

1.1.2 数据元方法研究

在制定方法方面, 文献[14-19]基于 SC 32 的数据元成果, 从数据元素的概述^[14]、命名、定义和表示^[15]、基本属性^[16]、提取^[17]、标准化阶段^[18]、管理工具^[19]等, 给出了比较全面和系统的方法; 文献[20]给出了面向政务领域的元数据属性规范及制定过程; 文献[21]给出了面向全国组织、干部、人事管理信息的数据元属性规范及制定过程等。

1.2 建设过程的问题及分析

尽管数据元已经具有一定的理论和方法基础, 但是在实践建设过程中, 仍然遇到很多挑战, 本文通过分析大量实践案例, 归纳主要问题如下:

(1) 组织方面。大部分组织对数据元建设不够重视, 未建立公司级数据元组织, 或组织架构

收稿日期: 2023-01-10 修稿日期: 2023-04-23

作者简介: 程红云(1972—), 女, 江西抚州人, 学士, 高级工程师, 主要研究方向为工业大数据应用与治理; 方亮(1986—), 男, 安徽安庆人, 研究生, 工程师, 主要研究方向为工业大数据应用与治理; *通信作者: 陈美(1981—), 男, 安徽六安人, 博士研究生, 工程师, 主要研究方向为数据治理和数据资产, E-mail: chenmei_195724@baosight.com

构不合理，出现无人管理、管不了、管理效率低等。

(2) 建设方面。缺少顶层设计指导，出现如各自为政建设、散乱建设、对全量数据制定标准等现象，造成投入大、建设慢、阻力大等。

(3) 重复问题。对相同业务场景、数据上传下达等情况，会有大类重复数据，导致数据元重复严重。

(4) 质量问题。未重视数据元质量，数据元内容的不完整、不正确、不规范等比较突出。

总结目前数据元建设过程中的问题，主要原因在于缺少顶层的组织、体系、方法等指导，为此本文开展数据元治理研究。

2 数据元治理框架

2.1 治理框架

本文针对数据元建设过程的问题，参考了国际数据管理协会(DAMA)^[22]、IBM数据治理委员会(IBM DG Council)^[23]、国际信息系统审计和控制协会(ISACA)^[24]、Gartner^[25]、全国信息技术标准化技术委员会^[26]等治理成果，结合数据元的特征，提出了数据元的治理框架，如图1所示。

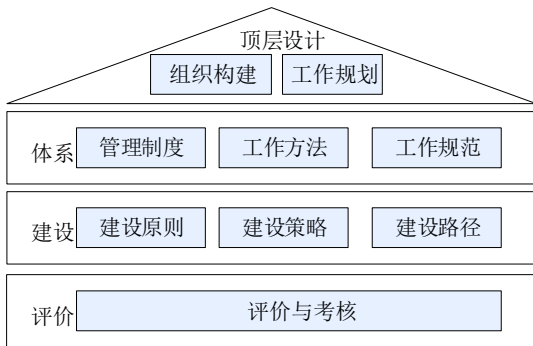


图1 数据元治理框架

数据元治理核心要素包括顶层设计、体系、建设及评价，具体内涵如下：

(1) 顶层设计。通过组织构建、工作规划，从顶层指导全局管理和建设。

(2) 体系。通过制定管理制度、工作方法和工作规范，保障数据元建设可控、可管和可落地。

(3) 建设。通过建设原则、建设策略、建设路径，保障有序和高效建设。

(4) 评价。通过评价检验建设效果，通过考核驱动建设可持续。

数据元治理要素的详细内容将在第3节中进行阐述。

2.2 治理机制

在数据元治理框架中，参考“PDCA”原则，建立各要素之间的工作机制，如图2所示。

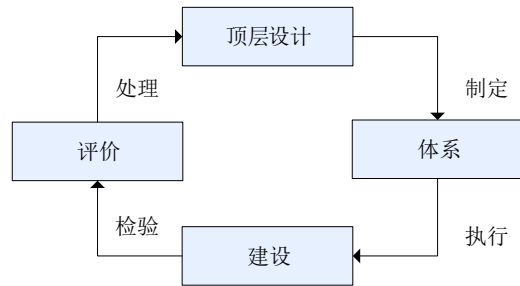


图2 数据元治理机制

数据元治理机制内涵：从顶层设计数据元的管理组织和规划，按照顶层设计研究和制定体系，基于体系建设数据元，通过评价检验数据元建设效果，最后通过处理完善顶层设计，从而形成了动态的数据元治理闭环。

3 数据元治理内容

3.1 顶层设计

顶层设计主要包括组织构建和工作规划，从顶层为数据元建设提供指导。

3.1.1 组织构建

组织是数据元建设和管理等工作的执行者，数据元作为标准之一，通常归属于公司数据标准管理委员会领导，属于下属专题组。

数据元组织架构的设计，要求架构合理和管理高效，同时还要能够解决数据元建设过程中的各类管理问题，如重复现象问题、建设不规范问题等；此外，数据元属于业务层面定义，因此必须要由业务方主导。

结合分析，本文提出了以“业务板块”为中心的组织架构设计，如图3所示。

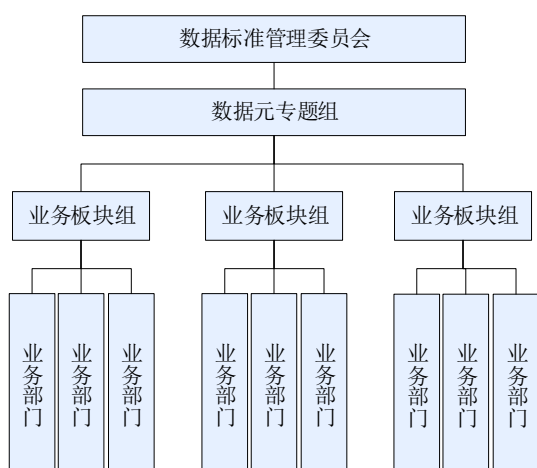


图3 数据元组织架构

数据元组织架构建议由数据元专题组、业务板块组、业务部门三层架构组成，各层职责如下：

(1) 数据元专题组。是数据元管理部门，负责数据元的体系建设、工作规划、发布以及资源协调、监督考核等工作。

(2) 业务板块组。是数据元建设的执行部门，要求相同业务领域属于同一个业务板块组，负责统筹本业务领域的数据元设计、建设、评审和维护等工作。

(3) 业务部门。是数据元建设的具体制定部门，服从业务板块组的安排，牵头或参与本领域的数据元制定。

3.1.2 工作规划

对于一个大型组织机构，数据元的数量通常很大，需要巨大的投入，并且是一项长期的建设工程，因此需要对数据元建设进行科学的规划与设计。

首先，结合公司的实际情况，建议制定3年或5年等数据元长期工作计划，要按部就班、有序地推进建设；

其次，制定各年度的工作目标、工作范围、相关责任方、工作要求等，确保过程可控、可监督；

最后，制定考核方法，驱动各部门按照计划按时推进工作。

3.2 体系建设

为了落实顶层设计落地，既要制定制度进行管控，同时还要制定方法和规范指导建设。

本文参考了业界的相关体系经验以及结合数据元的自身特点，提出了“管理办法-工作方法-工作规范”三维体系框架，如图4所示。

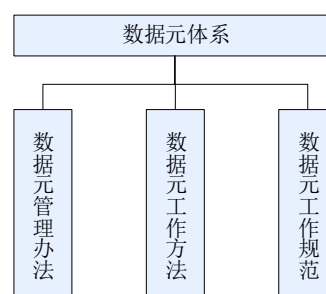


图4 数据元体系框架

管理办法是通过制度对数据元的建设和管理进行管控，保障数据元相关工作能够得到落实和执行。管理办法的主要内容包括任务、职责、流程、机制等内容。

工作方法是制定统一的工作方法，指导相关部门开展数据元建设和管理工作。数据元建设主要由业务人员负责实施，因此数据元工作方法，要基于业界通用数据元方法，站在业务人员视角，进行内容简化，确保业务人员能够快速理解和掌握。工作方法主要内容应包括基础理论、制定方法、管理方法、平台操作指南等相关内容。

工作规范是通过制定统一的工作规范，保障数据元的“名型值”统一。在实践中发现，即使方法相同，不同部门或不同人员，制定的数据元“名型值”千差万别，因此需对数据元的业务、技术和管理关键属性制定统一规范，如业务属性中的中英文名，要规范语法和语义，包括有效字符、非法字符、字符组成规范、长度等；如技术属性中的数据类型，由于相同字段可能涉及不同数据库，需要统一数据类型、转换规则等；如管理属性中的组织代码，要制定统一的组织代码长度、字符等规范。

3.3 建设方法

在实践中发现,对于大型组织,其场景通常很复杂,既有相同业务场景,同时又有数据流转或上传下达等情况,体系只能解决数据元如何制定,无法解决数据元怎么建设,为此本文提出“建设原则、建设策略、建设路径”,配套体系共同指导数据元建设。

3.3.1 建设原则

针对数据元建设过程的重复现象、管理难等相关问题,本文总结提出了4条数据元建设原则。

原则一:共建共享。对相同业务场景,按照共建共享原则,由业务板块组牵头,指派牵头业务部门和参与业务部门,共建共享。

原则二:源头权威。对交叉业务现象和上传下达数据流转等场景,对于相同数据元,由业务板块组牵头,明确数据产生源头,该数据在数据流中首次出现的位置即为源头,由数据源头部门负责制定标准,相关方引用即可。

原则三:对象管理。为了保证数据元管理的统一和对齐,数据元要以对象为单位进行建设、管理和应用。

原则四:治理准则。对相同业务场景,治理准则是按照共建共享的原则和相同“工作方法和工作规范”,如果系统和数据库相同,规范定义数据元属性;如果系统和数据库不同,先统一数据元命名,然后再规范定义数据元属性。

3.3.2 建设策略

对于大型组织通常会拥有很多系统以及大量的表和字段,原则上要对所有数据制定数据元,但是现实中要兼顾成本、价值和意义,需要制定数据元建设策略。

数据元作为数据标准,主要目标是统一数据语言,建议重点针对两类数据:一类是常用的数据,如流通共享、加工分析、交叉业务、上传下达等数据,通过数据元确保数据定义和理解一致,减少沟通成本;另一类是共性业务场景,通过共建共享数据元,可以避免重复建设,同时还能支撑统一管理、监督、对标分析等价值。

建设策略能够帮助组织更好掌握数据元的

工作重点、范围和边界,更加精准指导工作,提高建设效率。

3.3.3 建设路径

数据元建设路径是保障数据元有序和高效的建设,本文参考了文献[17, 20-21],提出数据元的建设路径:

第一步:数据元规划。由数据元专题组统筹,负责顶层设计,包括构建组织和工作规划。在工作规划中,要从公司级层面划分业务板块,并明确各业务板块组及相关参与方,尤其相同业务场景要归属同一个业务板块组。

第二步:数据元建设。由业务板块组牵头,按照“建设原则和建设策略”,负责本领域的数据元规划、设计和任务分解,并为每项数据元建设任务指派牵头业务部门和参与业务部门。

第三步:数据元制定。由牵头业务部门,按照数据元“建设原则和建设策略以及工作方法和工作规范”,梳理、定义本领域的数据元,参与业务部门负责补充和完善内容,最后提交到业务板块组进行评审。

第四步:数据元评审。由业务板块组负责组织相关专家对数据元进行评审,评审不通过退回业务部门进行完善,评审通过则提交数据元专题组进行发布。

第五步:数据元发布。由数据元专题组负责对数据元进行最后确认,然后在公司级层面发布。

第六步:数据元应用。各部门在数据建模的时候,要引用数据元定义数据模型属性,然后再基于数据模型建设数据表,从而推动数据元落地,实现数据语言统一。

3.4 评价方法

为了检验数据元建设效果,需要对数据元进行评价。

数据元作为标准,同时本身也是数据,兼具数据的一些共性质量问题。针对数据元质量低的问题,通过实践研究发现,即使工作方法和工作规范相同,也无法避免人为、工具等因素,产生数据元质量问题。

本文参考了文献[27-28],结合数据元的数据特点,提出了“工具校验+专家评审”的数据

元评价机制，数据元质量评价维度如下：

(1) 完整性。按照“工作规范”要求，数据元被赋予数值的程度，评价方式为“工具校验”，在数据元制定阶段执行，由业务人员负责。

(2) 规范性。按照“工作规范”要求，数据元数据符合规范的程度，评价方式为“工具校验”，在数据元制定阶段执行，由业务人员负责。

(3) 准确性。按照业务专家的专业知识，数据元准确表示其所描述的真实值的程度，评价方式为“专家评审”，在数据元评审阶段执行，由业务板块组及相关专家负责。

对于“工具校验”，原理是将“工作规范”要求和规则固化到工具中，借助工具实现自动化评价。

数据元质量的评价计算方法可参考文献[28]。

4 数据元治理工作建议

数据元作为数据标准，可看作是数字化建设的基础性工作，数据元的建设十分重要，结合实践总结四条工作建议：

(1) 争取高层支持。数据元建设通常涉及跨部门，需要多部门协同，如果没有高层的领导支持，工作很难开展。

(2) 业务人员制定。数据元本质上归属业务元数据，属于业务层面的定义，因此必须由业务人员负责制定，才能将数据元梳理清楚、定义准确。

(3) 重视标准权威。严格按照数据元的“制定、评审、发布和应用”管理流程，保障数据元标准的权威性。

(4) 质量合理控制。数据元作为标准，其规范程度越高，建设成本越高和建设效率越慢，因此应立足现实、结合实际，对数据元的规范程度进行合理控制，可对关键属性进行严控，对非关键属性可以放宽，后续可以不断地完善和迭代。

5 结语

本文针对当前数据元建设过程中的问题和挑战，基于国内外数据元和数据治理的研究和建设成果，提出了数据元治理框架以及工作建

议，并详细阐述了治理框架、治理内容和工作建议。在数据元治理框架中，提出了治理框架和治理机制，明确数据元治理要素及各要素之间的工作机制；在数据元治理内容中，围绕顶层设计、体系、建设和评价，核心治理要素进行详细阐述；在数据元治理工作建议中，从建设角度，围绕组织保障、责任主体、标准权威性、质量，提出四点工作建议。

本文的贡献是提出了比较全面和系统的数据元治理框架，能够有效和直接指导数据元建设。对象作为数据元的核心管理要素，对保障数据元的“名型值”统一与规范至关重要，后续将开展对象的研究工作。

参考文献：

- [1] 全国信息技术标准化技术委员会. 数据管理能力成熟度评估模型:GB/T 36073—2018[S]. 北京:中国标准出版社,2018.
- [2] ISO/IEC. ISO/IEC 11179-1:2015 Information technology: metadata registries (MDR): part 1: framework[S],2015.
- [3] ISO/IEC. ISO/IEC 11179-2:2019 Information technology: metadata registries (MDR): part 2: classification[S],2019.
- [4] ISO/IEC. ISO/IEC 11179-3:2013 Information technology: metadata registries (MDR): part 3: registry metamodel and basic attributes[S],2013.
- [5] ISO/IEC. ISO/IEC 11179-4:2004 Information technology: metadata registries (MDR): part 4: formulation of data definitions[S],2004.
- [6] ISO/IEC. ISO/IEC 11179-5:2015 Information technology: metadata registries (MDR): part 5: naming principles[S],2015.
- [7] ISO/IEC. ISO/IEC 11179-6:2015 Information technology: metadata registries (MDR): part 6: registration[S],2015.
- [8] 全国信息技术标准化技术委员会. 信息技术元数据注册系统(MDR)第1部分:框架:GB/T 18391.1—2009[S]. 北京:中国标准出版社,2009.
- [9] 全国信息技术标准化技术委员会. 信息技术元数据注册系统(MDR)第2部分:分类:GB/T 18391.2—2009[S]. 北京:中国标准出版社,2009.
- [10] 全国信息技术标准化技术委员会. 信息技术元数据注册系统(MDR)第3部分:注册系统元模型与基本属性:GB/T 18391.3—2009[S]. 北京:中国标准出

- 出版社,2009.
- [11] 全国信息技术标准化技术委员会. 信息技术元数据注册系统(MDR)第4部分:数据定义的形成:GB/T 18391.4—2009[S]. 北京:中国标准出版社,2009.
- [12] 全国信息技术标准化技术委员会. 信息技术元数据注册系统(MDR)第5部分:命名和标识原则:GB/T 18391.5—2009[S]. 北京:中国标准出版社,2009.
- [13] 全国信息技术标准化技术委员会. 信息技术元数据注册系统(MDR)第6部分:注册:GB/T 18391.6—2009[S]. 北京:中国标准出版社,2009.
- [14] 徐冬梅,杨建军,张展新,等. 数据元素的标准化方法(1):概述[J]. 信息技术与标准化,2008(Z1):49-51.
- [15] 徐冬梅,杨建军,张展新,等. 数据元素的标准化方法(2):命名、定义和标识[J]. 信息技术与标准化,2008(3):52-54.
- [16] 徐冬梅,杨建军,张展新,等. 数据元素的标准化方法(3):数据元素的基本属性[J]. 信息技术与标准化,2008(4):55-56.
- [17] 杨建军,徐冬梅,张展新,等. 数据元素的标准化方法(4):数据元素的提取[J]. 信息技术与标准化,2008(5):59-62.
- [18] 杨建军,徐冬梅,张展新,等. 数据元素的标准化方法(5):数据元素标准化阶段[J]. 信息技术与标准化,2008(6):53-55.
- [19] 张展新,徐冬梅,杨建军. 数据元素的标准化方法(6):数据元素管理工具[J]. 信息技术与标准化,2008(8):51-53.
- [20] 全国信息技术标准化技术委员会. 电子政务数据元第1部分:设计和管理规范:GB/T 19488.1—2004[S]. 北京:中国标准出版社,2004.
- [21] 全国信息技术标准化技术委员会. 全国组织、全国组织、干部、人事管理信息第1部分:数据元:GB/T 14946.1—2019[S]. 北京:中国标准出版社,2019.
- [22] DAMA International. DAMA data management knowledge system guide[M]. Beijing:Tsinghua University Press,2012.
- [23] IBM Corporation. IBM data governance council maturity model: building a roadmap for effective data governance[R],2007.
- [24] ISACA. COBIT 5:a business framework for the governance and management of enterprise IT[S]. 2012.
- [25] Gartner Group. Big data [EB/OL]. <http://www.gartner.com/it-glossary/big-data>.
- [26] 全国信息技术标准化技术委员会. GB/T 34960.5—2018 信息技术服务 治理 第5部分:数据治理规范:GB/T 34960.5—2018[S]. 北京:中国标准出版社,2018.
- [27] 全国信息技术标准化技术委员会. 信息技术 数据质量评价指标:GB/T 36344—2018[S]. 北京:中国标准出版社,2018.
- [28] 方亮,程红云,陈美. 数据元质量评价方法研究[J]. 现代计算机,2022,28(1):60-63,69.

Governance framework and suggestions of data element

Cheng Hongyun, Fang Liang, Chen Mei*

(Shanghai Baosight Software Co., Ltd., Shang 201203, China)

Abstract: As data standard, data element is not only the basic work of digital construction, but is also the key to building high-quality data assets. Currently, in the construction process of data elements, the problems such as duplication, low quality, non-standard construction, low efficiency of work, etc., is prominent. By analyzing research results of domestic and foreign, this paper puts forward the governance framework and suggestions of data element, and the purpose is to provide reference for data element construction in the industry.

Keywords: data element; data element governance; governance framework; data element construction

文章编号: 1007-1423(2023)10-0063-04

DOI: 10.3969/j.issn.1007-1423.2023.10.011

基于规则引擎的 DXF 与空间矢量数据互转方法研究

李辉勤*, 龙 杰

(贵州省煤矿设计研究院有限公司瓦斯研究分院, 贵阳 550025)

摘要: 针对 DXF 文件对空间分析能力的支持性弱、空间坐标系参数缺失的问题进行研究, 设计自定义转换器, 实现 DXF 文件到空间矢量数据的映射转换。利用 netDxf 工具库对 DXF 文件进行解析, 使用 SuperMap iObjects 组件进行可视化平台构建, 结果表明, 自定义的转换器对于 DXF 文件的矢量信息、图形信息具有良好的映射效果, 转换效率高、一致性强, 转换结果在 SuperMap 平台的验证中呈现质量较好。

关键词: 数据转换; DXF; GIS; 瓦斯地质图; netDxf

0 引言

CAD 软件的使用为自然资源规划和设计工作带来了极大的便利, 其中以 .dwg 结尾的数据格式最为常用。随着自然资源规划和设计工作向更加智能、高效和科学的道路上前行, 空间分析的作用日益凸显。 .dwg 格式能对数据的矢量信息、图形符号进行存储, 但并不支持对地理信息的存储, 在使用 SuperMap、ArcGIS 等传统 GIS 软件时, 无法对 .dwg 数据进行直接的空间分析, 且 .dwg 数据格式的存储结构并不公开, 导致使用者难以对它的信息进行有效的提取和自定义转换, .dxf 格式作为对 .dwg 格式文件的开源替代, 在常用的 CAD 软件中皆提供相互转换的功能。

矿井瓦斯地质图是瓦斯地质工作成果的反映, 可以最大限度地集中瓦斯地质信息、展示瓦斯地质规律、计算瓦斯储量, 是进行瓦斯灾害预测及防治的重要图件^[1-3], 但是瓦斯地质图的制作常依赖于 AutoCAD 等制图软件, 通常不会以空间数据格式(shapefile、geotiff 等)进行输出, 因此在后续的空间分析方面存在诸多阻碍。研发 DXF 格式与 shp 格式的转换器, 构建基于 GIS 的瓦斯地质可视化管理系统, 可以有效地动

态管理矿井瓦斯地质信息, 分析瓦斯地质规律。本文以 SuperMap iObjects 插件作为基础研发工具, 对 .dxf 数据结构进行解析, 并实现自定义的 DXF 文件到空间矢量数据的转换。

1 GIS 数据与 DXF 文件组织结构

1.1 GIS 空间数据组织结构

GIS 空间数据是用于描述地理实体的空间、属性和时间特征的。空间特征指地理实体的空间位置、几何及其拓扑关系; 属性特征描述与地理实体相关联的属性信息; 时间特征指地理实体随时间的变化而不断变化, 是一种动态信息, 可用于记录地理实体的历史特征。根据地理实体的空间几何特征, 可将空间数据抽象为点、线、面及注记四种类型, 可用矢量和栅格两种数据组织形式进行表达^[4], 空间数据组织结构见图 1。

1.2 DXF 文件组织结构

DXF 格式本质上是由成对的组码和组值组成, 指定了数据元素的类型及含义。完整的 DXF 文件由标题、类、表、块、实体和对象等六个节组成, 节中的内容包含对象名称、句柄、

收稿日期: 2023-02-16 修稿日期: 2023-02-27

作者简介: *通信作者: 李辉勤(1992—), 男, 贵州榕江人, 本科, 助理工程师, 研究方向为地理信息系统应用开发, E-mail: 872469219@qq.com; 龙杰(1992—), 男, 贵州凯里人, 硕士, 助理工程师, 研究方向为地理信息系统应用开发

图层名、颜色值、线型名、线型比例、线宽、厚度、扩展数据等属性信息及对象的坐标信息，而这些信息均由组码和组值组成^[5]，DXF文件组织结构见图2。

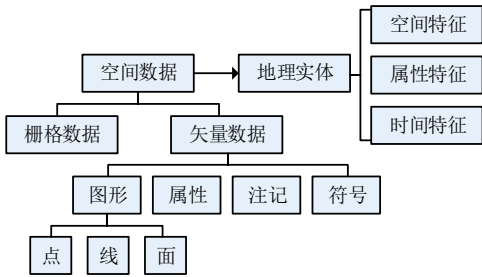


图1 空间数据组织结构

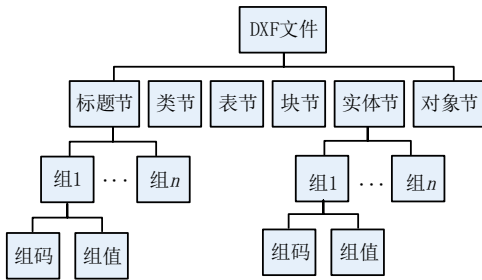


图2 DXF文件组织结构

2 转换思路

DXF格式和SHP格式属于不同的数据类型，它们之间存在转换差异，例如对于非矢量图形、文字注记等内容，不能够正常转换，进行DXF与GIS空间数据转换应该遵循以下原则：一是分层存储，根据要素特征和类型的不同，分图层组织数据，将同一类型、特征相似的要素存储到同一图层；二是属性统一，即DXF要素的标注信息及扩展数据与GIS空间数据的属性信息相对应；三是坐标系统一致，由于DXF文件通常采用笛卡尔坐标系，因此，对应的GIS数据不得使用地理坐标系，应该使用对应的投影坐标系。在转换过程中要充分掌握DXF数据类型与GIS数据类型的映射关系，明确DXF和GIS平台之间的符号库、字体、大小、线宽、颜色等信息的换算关系。①DXF文件导入到GIS平台的步骤为：读取DXF文件，分层获取每个图层的所有要素，遍历每个要素对象并转换为GIS几何对象，设置符号、颜色、字体、线宽等属性，存

储到矢量数据集。涉及到DXF转数据集DXF2Dataset类、解析及转换实体CoreTransformEntities类、符号库映射MappingHelper类三大类。SuperMap 10i除了包含GIS平台常见的点、线、面类型外，还提供了一种特殊的DXF图层类型，该图层可以同时存放点、线、面及文本对象，为DXF转GIS数据提供了兼容性完好的载体，可将转换结果存入该图层类型中。②地图文档转DXF文件的步骤为：获取地图文档的图层列表，遍历每个图层的空间要素及显示样式，转换为DXF要素对象，设置符号、颜色、字体、线宽等属性^[6]，保存到DXF文件。涉及到地图转DXF文件Map2DXF类、几何数据转换CoreTransformGeo类、符号库映射MappingHelper类三大类。

在转换过程中，通过调用SuperMap iObjects提供的接口实现对GIS数据的处理，调用netDxf内置算法来实现对DXF文件的解析、转换和存储操作。

3 关键技术

3.1 DXF转GIS空间数据的关键技术

DXF转为GIS空间数据的基本过程是解析DXF文件的数据信息，并将其转换为GIS几何对象，最后存储到GIS空间数据的过程，由于DXF文件的数据组织结构与GIS空间数据的组织结构不同，在转换过程中有一些关键技术必须解决。

3.1.1 数据类型映射

读取到DXF的实体信息EntityObject对象后，根据其Type属性获取EntityType类型，将其转换为SuperMap对应的Geometry对象，数据类型映射关系见图3。

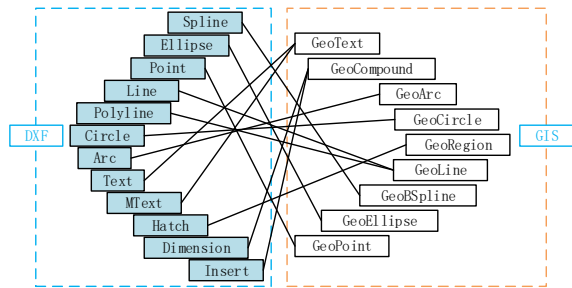


图3 数据类型映射关系

3.1.2 符号处理

为保证转换结果与原DXF显示效果一致，分析了瓦斯地质图常用的符号、线型和图案填充，在SuperMap平台定制相应的点线面符号库，并形成符号映射表，在转换过程中设置对应的符号，例如DXF实体使用GRAVEL图案，则SuperMap平台使用采空区20000001符号进行填充。

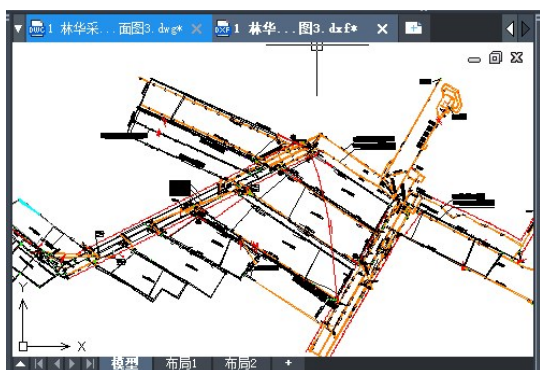
3.1.3 颜色转换

AutoCAD系统采用颜色索引ACI，取值范围为0~255，分别表示256种颜色。SuperMap采用RGB符号系统，RGB与ACI具有一定的换算关系，netDxf内置了AciColor.ToColor方法，可将ACI转换为RGB颜色。

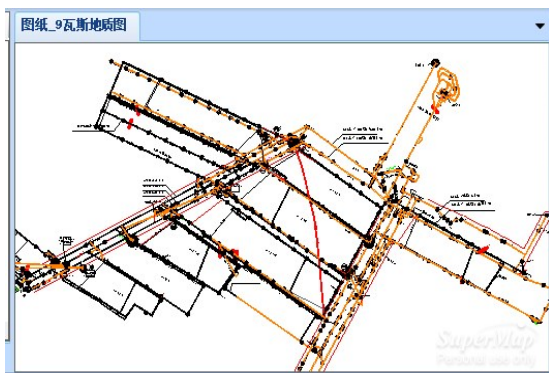
3.1.4 字体设置

DXF文件的Text和MText均表示文本对象，其Style属性中包含字体、字宽、字高、颜色、对齐方式等信息。在SuperMap构建GeoText对象，并按照DXF对象的Style设置其TextStyle属性。

DXF转GIS空间数据实例见图4。



(a) DXF文件预览



(b) DXF转GIS数据预览

图4 DXF转GIS空间数据实例

3.2 地图文档转DXF文件的关键技术

地图文档转DXF文件的基本过程是将地图的各个图层内的要素对象转换为DXF实体对象，最后输出成标准DXF文件的过程，将地图文档转为DXF文件的过程中要解决几个关键技术点。

3.2.1 从专题图提取样式

众所周知，GIS空间数据是由点、线、面及其属性信息构成，使用SuperMap的制图系统对图层进行符号化设置和专题图处理，形成完善的地图文档。在转换的过程中，需要读取专题图层的Theme信息，获取要素对象的GeoStyle，将其转换为DXF实体对象的Style。

3.2.2 控制显示状态

GIS系统提供了空间和属性过滤，可根据条件设置要素对象的显示状态。为保证输出的DXF文件与地图文档内容一致，在转换过程中要获取图层的DisplayFilter属性，根据其设置的条件提取相关要素进行输出。

3.2.3 标注转文本

GIS系统中，为了方便读图，经常会对图层设置动态标注。在转换过程中，将标注设置成标签专题图，读取标签专题图的ThemeLabel信息，并获取标注的文字内容，构建DXF的Text对象，设置文本内容和样式^[7]。

4 结语

本文对GIS空间数据组织特征、CAD数据模型及DXF文件组织结构进行深度分析，基于上文方法，构建了具有强逻辑关系的规则转换器，实现了DXF文件和GIS空间数据的互转映射。该转换程序已集成到自研的瓦斯地质可视化管理平台中，并在贵州林华煤矿得到检验，实践结果良好，在对林华煤矿的瓦斯地质图、防突预测图、采掘工程平面图等图纸进行DXF与GIS空间数据互转时，性能稳定、效率高、效果好，文字、图形及颜色保持完整，没有明显的的数据丢失和失真情况发生。

参考文献:

- [1] 李晓华,周炳秋,韩真理,等. 基于GIS的瓦斯涌出动态预测可视化系统[J]. 煤矿安全,2016,47(4):99-102.
- [2] 郭军. 基于瓦斯地质图的煤矿瓦斯预测预报系统设计与开发[J]. 煤矿开采,2007(4):4-6.
- [3] 庞敏,周荣福,卢红星,等. 基于组件式GIS的瓦斯地质图管理系统设计与开发[J]. 煤矿安全,2008(10):21-24.
- [4] 孟妮娜,周校东. GIS空间数据转换为DXF格式的方法研究[C]//2007'全国测绘科技信息交流会暨信息网成立30周年庆典论文集,2007:489-491.
- [5] 杨义辉,周谊,李明建. GIS空间数据转换为DXF格式的若干问题[J]. 测绘信息与工程,2008(5):43-44.
- [6] 魏域君. 基于FME的GIS数据到CAD数据转换技术研究[J]. 科学技术创新,2022(26):17-21.
- [7] 耿继军,刘荣,吴聪聪,等. 利用ArcMap和数据库改进GIS到CAD数据转换[J]. 测绘科学,2017,42(3):161-164.

Research on the method of converting DXF and space vector data based on rule engine

Li Huiqin*, Long Jie

(Gas Research branch, Guizhou Coal Mine Design Research Institute Co., Ltd., Guiyang 550025, China)

Abstract: In view of the weak support of DXF files for spatial analysis and the lack of spatial coordinate system parameters, a user-defined converter is designed to realize the mapping and conversion of DXF files to spatial vector data. The DXF file is parsed using the netDxf tool library, and the visualization platform is built using the SuperMap iObjects component. The results show that the customized converter has a good mapping relationship for the vector information and graphic information of the DXF file, and the conversion efficiency is high, and the consistency is strong. The conversion results are presented in the validation of the SuperMap platform with good quality.

Keywords: data conversion; DXF; GIS; gas geological map; netDxf

开发案例

文章编号: 1007-1423(2023)10-0067-07

DOI: 10.3969/j.issn.1007-1423.2023.10.012

低代码赋能电信运营商行业数字化转型

赵 龙, 冯强中, 汪伟伟, 许 磊, 解宇恩*

(科大国创云网科技有限公司, 合肥 230088)

摘要: 随着云计算、大数据、5G、物联网、人工智能等新IT技术的发展与应用, 如何最大化释放新IT技术带来的能量, 打破业务与开发间的壁垒, 提高业务应用研发生产力, 是各行业企业数字化转型的战略核心。低代码提供易用的定制化软件开发能力, 降低应用开发的准入门槛, 促进新IT技术深度融合, 助力软件开发降本增效提质, 敏捷响应快速变化的市场需求, 支撑海量数字化场景快速落地, 是赋能数字化转型的“加速器”。

关键词: 低代码平台; 电信运营商; 数字化转型

0 引言

基于低代码平台部分场景可以实现无代码开发, 同时对数据模型、表单、流程、规则、报表的可视化设计极大降低人员开发门槛, 业务老师也可以做场景开发, 促进了人才转型。

低代码平台促进了人才转型, 让更多的角色加入到场景的开发中, 提升了人才整体中的要素数量; 同时提升了各角色利用平台更充分地发挥出自身价值和能力, 提高了人才整体要素的配置和各要素在人才整体中的地位和作用。

支持极简模式和专家模式, 可以快速复用场景; 支持领域组件和模板的大量积累, 提升了低代码的复用能力; 支持场景复用和场景迁移, 可以实现快速业务变更。

简单场景可以实现无代码开发; 中等场景可以实现低代码开发(可视化开发+API代码开发); 高难场景可以实现高代码开发(可视化+二次开发); 随着组件、模板的领域化积累, 可以实现极速复用开发。

1 低代码平台架构

科大国创软件股份有限公司以中国科技大学为技术依托, 以擅长的基于网络的软件技术、数据库技术、云计算技术、移动互联技术、系统集成技术、智能控制技术、网络终端产品的研究与开发技术, 面向国内、国际提供优质的产品和服务。科大国创在现有的技术和产品的基础上进行低代码平台的深入研究、构建与应用。

1.1 总体架构

低代码平台是一个基于云原生架构体系打造的开放式生态系统, 是一个软件产品组合, 一个支持流程类业务场景快速开发的工具链。致力于形成业务与技术深度协作的终极敏捷形态, 总体架构如图1所示。

1.2 功能架构

支持流程类业务场景的可视化开发与全生命周期管理, 利用平台+组件/模板实现业务知识 with 业务经验的显性化沉淀, 助力流程类业务快速加载上线, 其功能架构如图2所示。

收稿日期: 2023-01-06 修稿日期: 2023-01-14

作者简介: 赵龙(1982—), 男, 安徽铜陵人, 硕士研究生, 副总裁, 研究方向为通信运营商IT咨询规划、软件系统设计、智慧社区、云计算和数据智能; 冯强中(1978—), 男, 安徽合肥人, 硕士研究生, 副总裁, 研究方向为数据挖掘和自然语言处理; 汪伟伟(1990—), 男, 安徽六安人, 硕士研究生, 部门经理, 研究方向为云计算和大数据分析; 许磊(1984—), 男, 安徽合肥人, 本科, 部门经理, 研究方向为通信运营商IT咨询规划、软件系统设计; *通信作者: 解宇恩(1998—), 男, 安徽合肥人, 本科, 部门高级软件工程师, 研究方向为软件系统设计、前端可视化, E-mail: xie.yuen@ustcinfo.com

1.3 技术架构

技术选择遵循主流、成熟、适度先进原则，

兼容电信及联通自研PaaS组件，符合运营商组件选择标准及上云规范，其技术架构如图3所示。

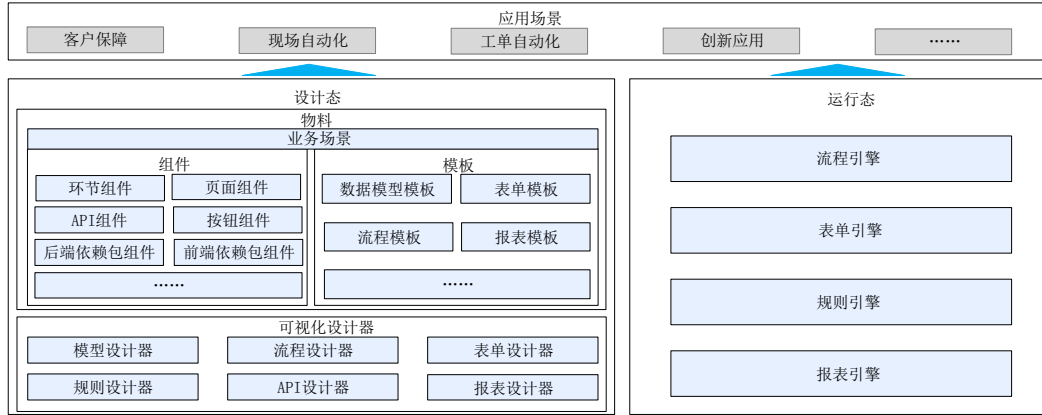


图 1 总体架构

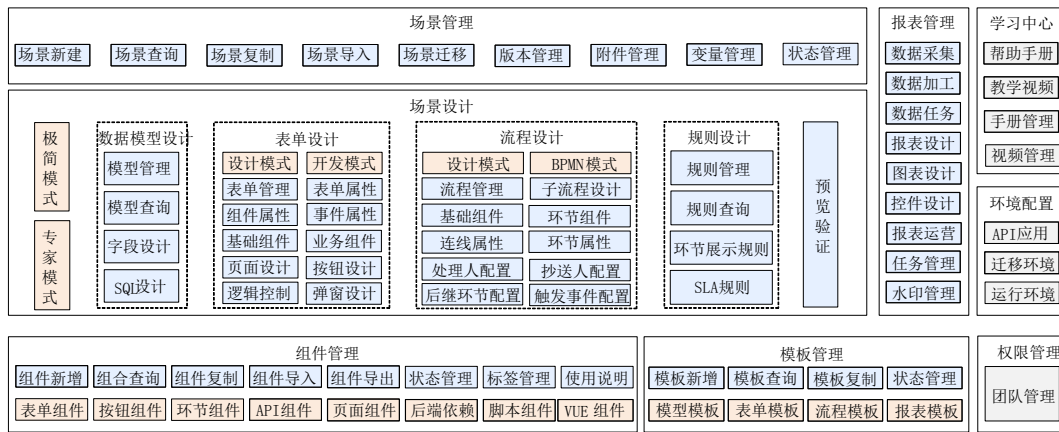


图 2 功能架构

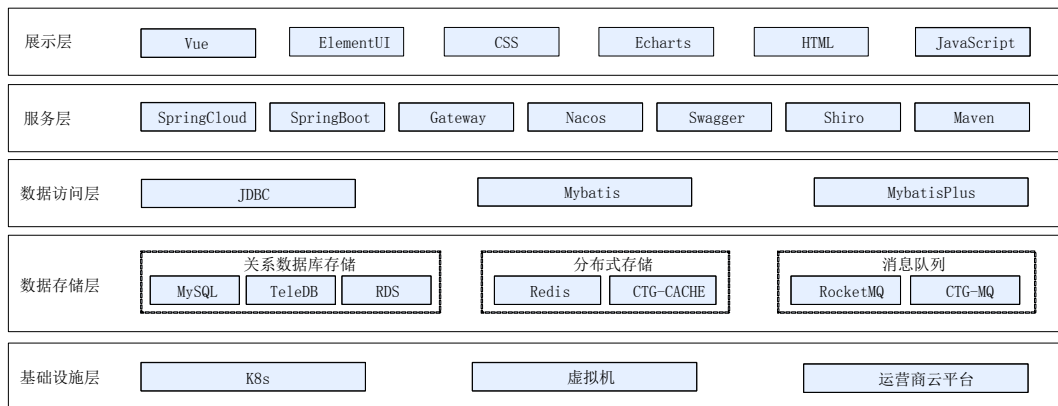


图 3 技术架构

1.4 部署架构

部署架构如图4所示。

1.5 架构演进

1.5.1 V1.0版本

- (1)发展阶段：导入期；
- (2)核心特征：工具型、功能视角(低代码模块以独立产品形式研发，以系统功能为维度组织菜单)；
- (3)集成模式：逻辑松耦合、物理紧耦合(流程、表单模块作为独立工具，与业务系统设计中心、业务单引擎逻辑耦合；使用数据同步等方式实现目录数据一致)；
- (4)开发模式：分散配置，业务场景的组装

通过硬编码实现。

1.5.2 V2.0版本

- (1)发展阶段：成长期；
- (2)核心特征：平台型、业务视角(添加“平台”层，用于牵引各模块功能及物料，以业务场景为维度组织用户界面与操作)；
- (3)集成模式：逻辑紧耦合、物理松耦合(底层模块、平台层、业务系统之间通过“业务场景”逻辑紧耦合，但开发层面通过API进行数据交互，不会冗余存储数据，更符合微服务思想规范)；
- (4)开发模式：一站式设计开发，业务场景的组装者是低代码平台。

设计态交互过程和运行态设计过程分别如图5和图6所示。

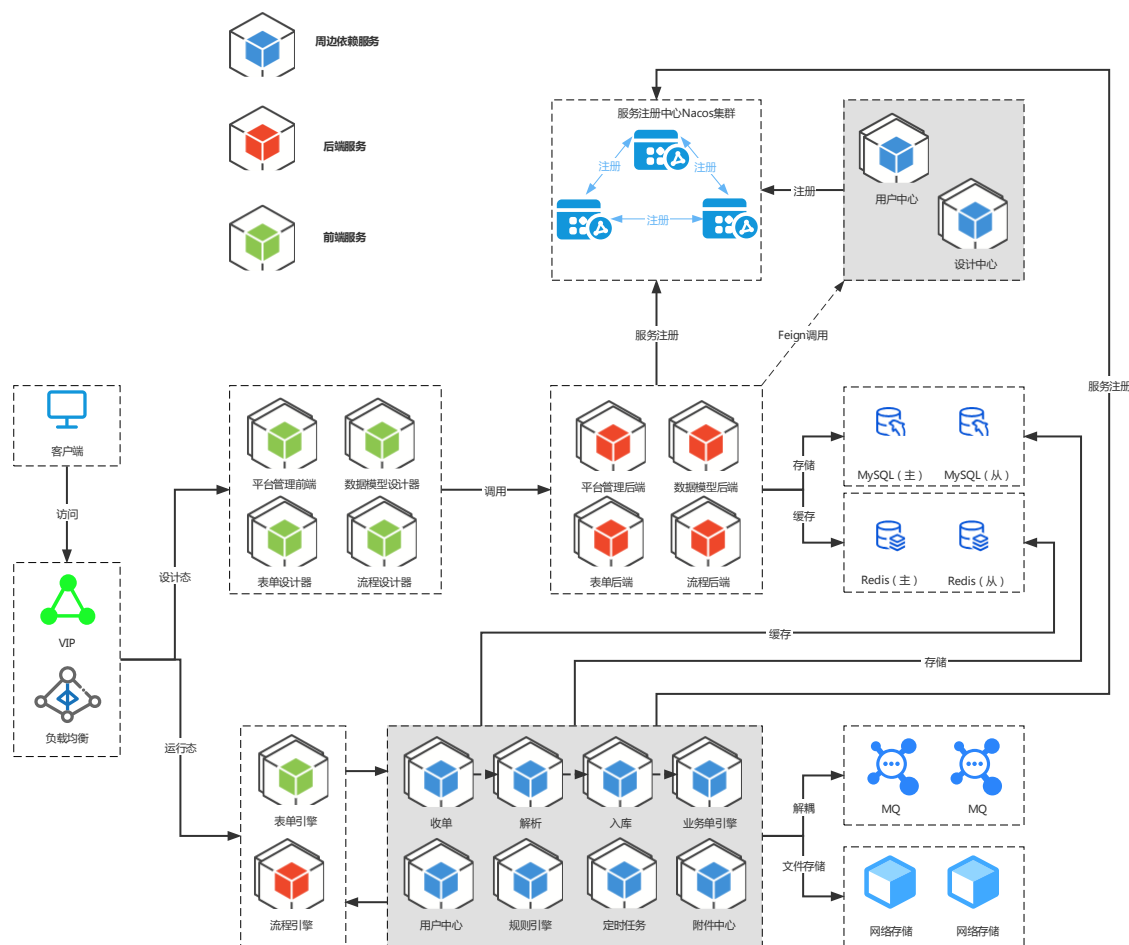


图 4 部署架构



图 5 设计态交互过程



图 6 运行态设计过程

2 低代码平台的核心特性

2.1 降本增效,快速上线,灵活变更

运用低代码技术将底层架构和基础设施等抽象为图形界面,以可视化设计及配置为主,少量代码为辅,提供快速搭建页面、设计数据模型、创建业务逻辑等能力,实现应用快速构建。

它同时是面向专业开发者的应用开发工具,通过图形界面或领域专用语言(DSL)简化应用开发过程,以标准化工具为基础,场景化工具及个性化工具为扩展,同时可支持常规在线编程工具对应用源码进行二次开发,有效提升应用开发效率、提升共性业务组件的复用率、降低工程维护成本^[1]。

2.2 两种模式,精准定位

极简模式和专家模式,支持业务专家与技术专家使用。极简模式是一种快速零代码搭建场景的方式。用户可以选择已发布的公开的场景作为模板使用,快速创建新的场景专家模式是一种经过表单、流程、规则等的可视化配置搭建场景的方式。专家模式是为了高级开发者准备的,用户可以选择从“数据模型设计”开始或从“表单设计”开始进行场景开发,即提供“模型驱动”和“表单驱动”两种场景开发模式。

2.3 一站式流程类业务开发

通过一站式、引导式的方式,将同一场景的所有模块管理和设计在统一步骤条下引导设

计,包括场景下的数据模型、表单、流程、规则的管理和设计器设计;同时在统一的步骤条下还可以直接搜索、预览、复用平台丰富的模板和组件。支持多种组件和模板的积累,通过积累的成果极大提升平台能力的复用性,降低重复工作量,从而极大地提高对应用场景的开发效率。

2.4 业务逻辑规则化配置

借助开发思想中的条件分支判断、“插槽”等方式,将表单元素行为控制、流程参与者匹配、流程后继环节匹配等业务逻辑固化成可视化规则配置方式的实现。

2.4.1 条件分支判断

利用规则思想、开发中的条件分支判断思想,抽象出通用判断模型,对判断的结构控件、条件元素进行自由组合

2.4.2 设置多个“插槽”

在表单设计、流程设计、规则设计、组件设计等模块设置业务逻辑的相关配置“插槽”。用户可在场景开发的全过程中进行配置以满足业务需求。

2.4.3 灵活的DSL语法转化

设计态通过用户在页面可视化配置,生成对应业务逻辑描述的DSL,在低代码中为json或xml。

运行态通过DSL语法解析器生成页面Vue代码和流程环节、参与者的规则,结合业务实例数据实时动态渲染表单页面,驱动流程环节前进或后退、派单给环节参与者。

2.5 业务组件组合复用

通过丰富的基础组件,并通过配置、复用、

上传代码等方式构建新的组件,组件库提供复用这些组件的能力,助力快速、便捷、灵活地进行业务开发。

2.6 代码增强,便捷测试

通过提供HTML、JS、CSS、XML在线编辑器和在线测试功能,解决需要特殊样式的表单和流程、需要绑定复杂事件的定制组件、后台有交互的组件的问题,达成灵活支撑个性二次开发的目标,助力用户无需打包编译即可快速二次开发与调试。

场景1:对于需要添加特殊样式可以自由编写html5+css代码;

场景2:复杂的数据处理、个性化的交互事件、特殊的页面元素与流程环节监听等场景,用户可以通过编JavaScript满足业务需要。

3 低代码赋能电信行业典型案例

3.1 中国电信总部实例

(1)级别:运营商集团级应用系统;

(2)定位:云网售中、售后服务与维护保障调度的核心能力,负责客户保障、云网保障、IT保障、生产指挥及综合化维护等业务活动的集约管理和调度,支撑云网生产智能化调度、云网一点服务敏捷化响应、云网管理一体化协同。

(3)业务需求设计上线的新老模式对比。

面对新业务场景的支撑或存量业务场景优化,基于低代码可视化设计能力,提供数据配置、环节配置、流程配置、规则配置及测试部署的可视化设计、在线测试及快速发布,实现新业务场景支撑或存量业务场景优化周期由月缩短至周或天。

业务系统迁移重构工作量对比见表1。



图7 使用低代码前后开发对比

表 1 业务系统工作量对比

维度		简单难度场景	中等难度场景	(极)高难度场景
占比		28%	31%	41%
使用低代码平台开发模式		100% 场景无代码实现	70% 场景无代码实现 30% 场景低代码实现	50% 场景无代码实现 50% 场景低代码实现
平均开发工作量(人日)	不使用低代码平台	2~4	8~13	15~20
	使用低代码平台	0.5~1.5	4~7	10~13
开发效率提升		提升约 70%	提升约 50%	提升约 30%

业务流程划分为3个难度级别，分别是简单、中等、(极)高，数量占比约为3 : 3 : 4。经实践和总结，对于简单和中等以及极高难度业务使用低代码平台开发均可以大幅度提高开发效率。

3.2 移动网络运维自动化驾驶项目

(1) 级别：运营商省级应用系统；

(2) 定位：通过可视化拖拽和简单配置操作，快速完成自动化巡检、故障处理业务流程构建和网络运维驾驶舱构建；

(3) 成效：巡检自动化率、故障及时率、故障单处理率显著提升，降低应用开发的技术门槛，通过低代码平台为 DICT 人才培养提供技能培训，促进人才转型。

3.3 电信分析优化项目

(1) 级别：运营商省级应用系统；

(2) 定位：通过可视化拖拽和简单配置操作，快速完成自动化巡检、故障处理业务流程构建和网络运维驾驶舱构建。该项目是基于海量数据结合 AI 及大数据技术能力进行智能分析，注智云网运营生产流程，实现事先化、主动化、智能化运维，并更加注重海量数据下的报表设计与分析；

(3) 成效：解决了该省份电信业务相关的数据分析与统计方面的问题，具体包括产能分析、指标考核、工单分析、工作量核算、智能网管等相关报表。功能方面，报表平台除常规报表制作的基础功能之外，还具备动态列、小计、数据控件、权限管控、图表设计、存储过程、预设样式等，同时报表对于导出方面做了很大的优化提升，支持百万级数据的清单导出，不管是性能消耗上还是数据量上都将是提升。

4 低代码发展展望

4.1 国家政策导向

国家政策和整体环境共同促进低代码持续发展。“十四五”规划大力推动企业数字化转型，低代码作为数字化转型的重要工具受到广泛关注。后疫情时代，国家高度重视数字化发展对经济快速恢复的重要作用，国务院关于印发扎实稳住经济一揽子政策措施的通知指出，“促进平台经济规范健康发展。充分发挥平台经济的稳就业作用，稳定平台企业及其共生中小微企业的发展预期，以平台企业发展带动中小微企业纾困。引导平台企业在疫情防控中做好防疫物资和重要民生商品保供‘最后一公里’的线上线下联动^[2]。”在政策扶持导向和整体发展趋势的影响下，低代码也将获得多重发展机遇。

4.2 低代码能力将会更加丰富和成熟

云计算、大数据、AI、RPA 机器人流程自动化等多种新技术蓬勃发展，不同技术的结合尝试往往能带来惊喜。比如 RPA 与 AI 技术的结合，将 OCR 图片文字识别、NLP 自然语言处理、NLG 自然语言生成等诸多主流人工智能技术与 RPA 相结合使用，使得非结构化数据业务流程也能实现自动化。低代码利用其高度封装的特点，介入后将复杂的应用配置过程简化，极大降低了开发难度，创建 AI 自动化业务流程。低代码与新技术的组合拳，是提高工作效率、提升产品价值的“开挂”工具。低代码、RPA、AI 等前沿科技深度融合应用到实际业务场景，将带来全新的开发体验。

4.3 低代码平台将会更加自动化和智能化

未来低代码平台将会覆盖更多应用场景，这对平台的易用性和灵活度提出了新的挑战。随着智能技术的普及，将会有更多的泛自动化、智能化能力赋能到企业应用中，由自动化向智能化发展，是对低代码产品更高的要求。数字化发展必然伴随着自动化能力的提升，先实现自动化，再发展智能化是必经之路^[3]。低代码平台的自动化引擎和智能化引擎要求深度结合企业业务场景，基于结果判断配置策略，对业务流程进行自动执行，实现工单流程的自动化处理。引入外围大数据/AI能力，通过直接引用或二次组合封装支撑调度业务场景或对外共享开放。纵横向处理过程可视、可管、全域工单关联，对工单或相关配置项的处理有重大意义的状态变化的关注及时通知，支持不同视角对工单端到端处理的评价。

4.4 低代码将会掀起上云热潮

Paas+低代码历经十余年的发展，云原生技术已经成为坚实的底座，各行业企业也稳步完成业务上云计划。低代码与PaaS相结合，将云架构、云部署、云资源等能力可视化管理，最大化发挥云上的敏捷开发能力，基于云化的基础设施和平台研发新应用，帮助企业完成上云计划，并持续提供云上可持续发展方案。除此之外，低代码将持续增强集成能力，促进形成平台生态体系，增强云资产复用能力和云管理

能力，提升企业上云的价值和意义。此外，云原生将能帮助低代码平台更快更好地对开发的应用进行一站式管理和发布，通过流水线构建、部署、打包、发布、监控等助力低代码应用安全快捷的运行，进一步提升应用的稳定性和安全性。

5 结语

低代码平台是一个基于云原生架构体系打造的开放式生态平台，是一个软件产品组合，提供了一系列的技术能力和工具集，方便面向企业构建差异化的应用。为应对更多企业需求，下一步，如何高效整合各环节资源，进行合理化分类、封装、重构是接下来低代码开发平台需要重点解决的问题。未来几年内，促进企业内部人才转型，优化人才结构以及降本增效仍是企业最大的需求。低代码平台需要做到的是基于云开发、云协同、云共享、云应用，实现软件研发全生命周期“云端”管理，通过全面协同能力，实现产品全链路在线贯通，打造全角色云上协同、高效研发新模式。

参考文献：

- [1] 冯喆,张强.低代码开发平台在企业数字化转型中的应用研究[J].互联网周刊,2022(10):13-15.
- [2] 不啻微芒,造炬成阳:中国低代码行业研究报告[R]//艾瑞咨询系列研究报告,2022(8):527-582.
- [3] 潘莉,戴路,吴姗.低代码标准建设探索及应用实践[C]//中国标准化年度优秀论文(2022)论文集,北京,2022:586-592.

Low-code enabling the digital transformation of telecom operators

Zhao Long, Feng Qiangzhong, Wang Weiwei, Xu Lei, Xie Yuen*

(Kedaguochuang Cloud Network Technology Co., Ltd., Hefei 230088, China)

Abstract: With the development and application of new IT technologies such as cloud computing, big data, 5G, the Internet of Things, and artificial intelligence, how to maximize the release of energy brought by new IT technologies, break the barriers between business and development, and improve the productivity of business application R&D is the strategic core of digital transformation of enterprises in all industries. Low code provides easy to use customized software development capabilities, lowers the threshold for application development, promotes the deep integration of new IT technologies, helps software development reduce costs, increase efficiency and improve quality, responds quickly to rapidly changing market demands, and supports the rapid implementation of massive digital scenes. It is an “accelerator” for enabling digital transformation.

Keywords: low-code platform; telecommunication operators; digital transformation

文章编号: 1007-1423(2023)10-0074-07

DOI: 10.3969/j.issn.1007-1423.2023.10.013

基于知识图谱的法理知识服务系统设计与实现

莫城为, 温怡烽*, 邓邦鹏

(中国人民解放军 75610 部队, 深圳 518000)

摘要: 为满足特定用户准确聚焦、有效关联、高效运用法理知识现实需求, 利用信息化技术拓展优化法理知识服务手段, 设计实现一种面向法理知识服务的信息系统。利用知识图谱、Vue 前端开发框架等相关技术, 选取部分香港法例作为试验数据, 设计并实现了一种基于知识图谱的法理知识服务系统。通过开发部署, 该系统设计初步实现了知识管理、语义搜索、智能推荐、关联分析四项核心功能, 可有效提高法理知识服务效益, 为知识图谱在特定垂直领域应用提供参考借鉴。

关键词: 知识图谱; 法理知识; 设计与实现

0 引言

随着国家法治建设与现代信息技术的不断发展, 法律信息化已成为普法用法的重要发展方向。与普通民众相比, 特定用户希望能根据自身职业特点, 按需定制更加聚焦垂直领域的法理知识服务, 以减少对常规法理知识服务手段的依赖(如人工查阅法律法规原文、咨询法律专业人士等), 进而实现以知识驱动提升业务工作质效。因此, 针对特定用户提出的法理知识服务现实需求, 选取部分法律法规作为试验数据, 结合知识图谱^[1]在知识聚合^[2]、语义关联^[3]、智能推荐^[4]等方面的应用成果, 构建相应的法理知识图谱。在此基础上, 设计实现具有知识管理、语义搜索、智能推荐、关联分析四项核心功能的法理知识服务系统, 为特定用户提供精准高效的法理知识服务。

1 法理知识图谱的构建

1.1 数据的采集与清洗

“电子版香港法例”是官方的香港法例资料

库, 其中提供的香港法规内容完备、分类清晰、更新及时, 是优质的法理知识原始数据获取来源。因此, 从中选取了《专利条例》等 93 部香港现行法律法规, 作为构建法理知识图谱的试验数据。

试验数据原始格式为 RTF, 清洗方法是将非结构化的 RTF 文件转换为半结构化的 HTML 文件, 通过分析 HTML 文件中的 CSS 属性差异, 利用 XPATH 属性定位提取 HTML 文件中的目标内容(如: 法例名称、法例描述等), 并存储为 JSON 文件, 实现数据清洗。数据采集清洗部分代码如图 1 所示。

```
# 获取第一个subrule出现的地方
sibling = getFirstSubrule(subRuleXPath)
# 将第一个subrule之前的内容作为rule的属性
jsonRuleList = getRule(sibling/following-sibling)
while sibling:
    subRule = getSubrule(sibling)
    jsonSubruleList.append(subRule)
    sibling = next(sibling)
```

图 1 数据采集清洗代码(部分)

收稿日期: 2023-02-03 修稿日期: 2023-02-23

作者简介: 莫城为(1990—), 男, 广东陆丰人, 硕士, 工程师, 研究方向为数据开放、隐私安全; *通信作者: 温怡烽(1999—), 男, 广东深圳人, 学士, 助理工程师, 研究方向为知识图谱、网络空间安全, E-mail: 582269292@qq.com; 邓邦鹏(1991—), 男, 广西资源人, 硕士, 助理工程师, 研究方向为知识图谱、指挥信息系统

1.2 法理知识图谱构建框架

Neo4j是一种高性能图数据库，可将数据存为节点和关系，适用于高度关联的数据关系建模以及变化较快的海量数据存储与管理^[5]。因此，本文选择Neo4j作为知识图谱存储与管理载体，将清洗后的试验数据导入Neo4j图数据库中，进行法理知识图谱构建，基本思路是：

(1) 根据数据清洗结果进行知识建模，对试验数据进行实体、属性、关系抽取。

(2) 通过实体对齐、知识推理、数据补齐等技术处理，实现对试验数据的知识融合。

(3) 将处理后的试验数据存入Neo4j图数据库中，完成知识图谱的构建，并依托Neo4j图数据库软件配套功能实现信息管理、语义搜索、关联分析等前端接口。整个构建过程大致可分为知识建模—知识抽取—知识融合—知识应用四个具体步骤，如图2所示。

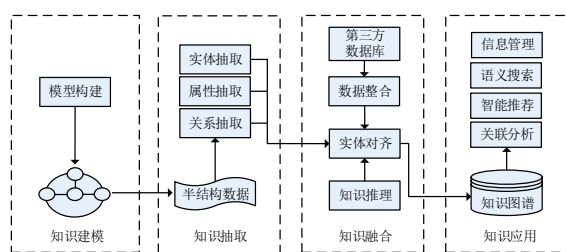


图2 法理知识图谱构建框架

1.3 法理知识图谱构建具体步骤

(1) 知识建模。知识建模是一种有效描述知识数据结构、关系的方法，能够将现实世界中存在的客观知识表述成计算机等电子设备可识别处理的对应内容。对于法理知识图谱构建，本文主要采用“自顶向下”的知识建模方法，通过对清洗结果进行梳理，共提炼出法例、子例、法条、关键词四类法律知识实体以及它们之间存在的包含、涉及、关联三种关系，知识建模结果如图3所示。

(2) 知识抽取。知识抽取主要是根据知识建模结果，有指向性地对试验数据中的实体、属性、关系实例进行提取，实现从半结构化JSON数据到<实体—关系—实体><实体—属性—属性值>等结构化三元组数据的有效转化。

(3) 知识融合。知识融合通过对知识抽取结

果进行去冗、消歧、整合等技术处理，实现对结构化数据的完善与补充。在此基础上，将知识融合结果按照三元组结构存入知识图谱，完成对知识图谱数据的构建。

(4) 知识应用。知识应用作为联结知识图谱与系统业务之间的“信息纽带”，其主要目的是在完成知识图谱数据构建的基础上，进一步实现对知识图谱功能的构建。本文主要利用Cypher查询语言^[6]、GDS图算法库^[7]、Vue前端开发框架^[8]等关键技术，以实现系统前端支持的信息管理、语义搜索、智能推荐、关联分析功能。

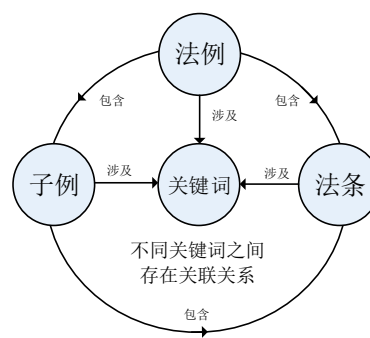


图3 法理知识图谱知识建模

2 法理知识服务系统需求分析

为满足特定用户对法理知识服务提出的现实需求，设计实现一套轻量化、易部署、交互界面友好的法理知识服务系统(以下简称系统)，通过区分不同维度的现实需求，从功能需求、性能需求、部署需求三个维度进行需求分析。

2.1 功能需求

系统功能需求主要包括：①信息管理需求，系统可全面掌握后端知识图谱综合情况，便于实现多视角的统计分析，以启发用户发现并聚焦目标知识；②信息检索需求，用户可通过系统检索准确获取符合用户意图的有效信息；③智能推荐需求，系统可基于数据关联及后端推荐算法，根据用户检索结果主动推送相应关联知识。

2.2 性能需求

系统性能需求主要包括：①综合呈现需求，

系统通过数据可视化方式，以达到呈现方式更加形象、蕴含信息更加丰富、数据分析更加有效，用户理解效率高、印象深的效果；②简易操作需求，系统可通过提供文字输入框、鼠标点击按钮等方式，以实现用户与系统之间及时可靠的人机交互操作；③高效响应需求，为了优化用户体验，提高知识服务效益，系统应当在用户检索、图形渲染等方面进行性能优化。

2.3 部署需求

系统部署需求主要包括：①离线(局域网)部署需求，特定用户希望系统可部署在封闭条件、局域网络和单机环境下，以适应特定用户工作环境及要求；②快速部署需求，由于系统部署在离线(局域网)环境下，不同用户可能需要自行部署、重复部署，因此系统部署方法应操作简单、易于上手，实现快速部署；③低配部署需求，系统部署对计算机软硬件要求适中，并能实现对计算机操作系统的上下兼容与整体迁移。

3 系统设计与实现

3.1 系统框架设计

系统以法理知识图谱作为底层支撑，以通用的“浏览器/用户”(B/S)模式实现人机交互，可对法理知识图谱数据进行信息管理；并针对用户提出的法律关注点、关注需求进行知识图谱科学分析，高效反馈对应的可视化语义关联子图及法理依据，可帮助用户有效获取法理依据辅助决策信息，突破人脑思维限制，及时作出优选决策。系统架构如图4所示。

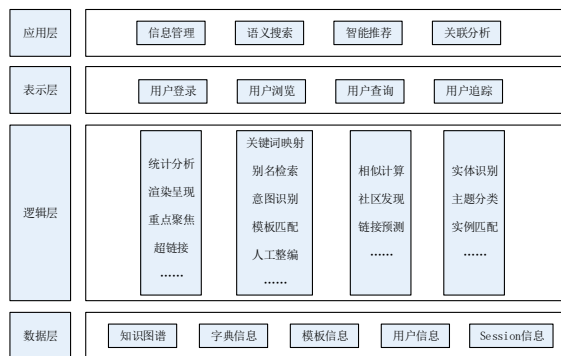


图4 法理知识服务系统框架

3.2 系统功能实现

3.2.1 信息管理

信息管理功能主要实现对法理知识图谱的多维度展示，具体包括四个功能模块：①法理知识图谱3D建模，该模块利用d3-force-graph开源包，形象绘制法理知识图谱中实体之间的语义关系；②实体、关系柱状图，该模块利用Echarts开源包，实现对法理知识图谱中实体、关系总数的实时统计与展示；③实体类型饼图，该模块利用Echarts开源包，实现对法理知识图谱中各实体类型及占比的实时统计与展示；④热点词云，该模块依托系统后端 Neo4j 图数据库配套的 PageRank^[9]等中心度算法，实时筛选并展示法理知识图谱中重要度优先的“关键实体”。

3.2.2 语义搜索

语义搜索功能主要实现对用户关注点及关注需求的有效检索及反馈，在基于字符串匹配的基础上，得益于底层知识图谱等关键技术的加持，使语义搜索范围不再局限于用户搜索内容字面本身，而是结合搜索内容在知识图谱中对应的语义关联，返回用户期望获取(符合用户真实意图)的搜索结果。目前，系统语义搜索功能主要支持两种搜索模式：

(1) 基于关键词的语义搜索。该模式主要通过建立关键词与法理知识图谱中实体之间的有效映射(识别输入的关键词后，将关键词嵌入 Cypher 查询语句，实现对法理知识图谱中对应实体的有效映射)，将对关键词的搜索任务转化为查询知识图谱中对应实体所涉及的语义关联子图及对应的语义信息，其具体操作流程如图5所示。

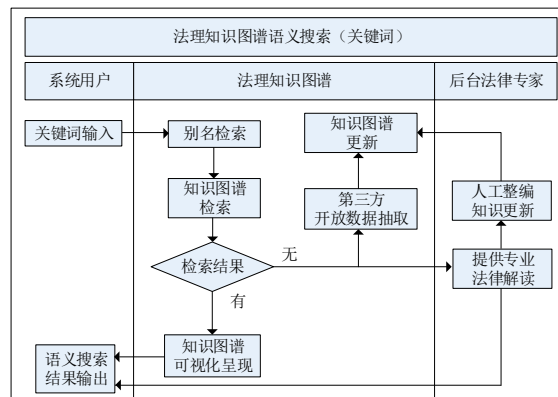


图5 基于关键词的语义搜索

(2) 基于自然语言问句的语义搜索。该模式主要通过“意图识别—模板查询”的检索方法，首先对用户输入的自然语言问句进行意图识别，目的是对用户输入的自然语言问句进行问句分类，并实现对应 Cypher 查询语句的模板匹配，从而将人类自然语言转换成知识图谱能够正确理解的 Cypher 查询语句，如表 1 所示。系统设计了部分自然语言问句结构对应的问句意图及匹配模板，例如当用户输入“专利涉及的法理依据”时，系统可通过意图识别判定问句意图为关系查询，匹配对应的查询模板为“match (n {name: ‘专利’}), p=(n)-[r: 涉及]-(m) return p”。通过意图识别的有效映射，实现法理知识图谱检索。最后，将知识图谱检索得到的语义关联子图及对应的语义信息作为语义搜索结果，其具体操作流程如图 6 所示。

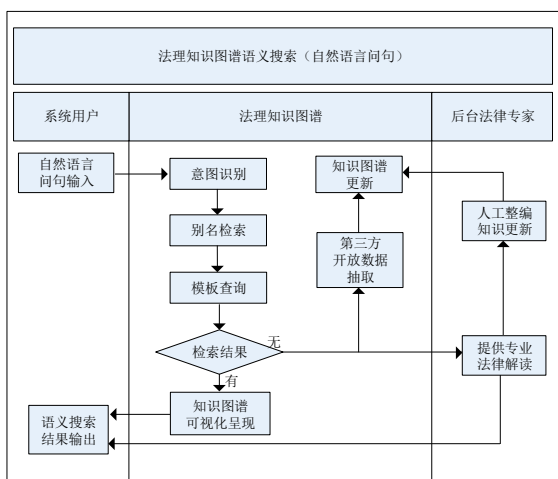


图 6 基于自然语言问句的语义搜索

3.2.3 智能推荐

智能推荐功能主要是在语义搜索结果基础上，对其中某一实体进行关联实体的挖掘与推荐，能够有效实现对同类法理依据的启发式关联。具体实现方法是通过 Neo4j 图数据库中的图数据科学(graph data science, GDS)算法库，利用相似计算、社区发现、链接预测等图算法，对知识图谱中的不同实体进行空间计算，并将计算结果作为智能推荐的基本依据，帮助用户深度挖掘有效信息，以信息流驱动决策流，加快特定用户业务活动“OODA^[10]”循环运转效率。

3.2.4 关联分析

关联分析功能模仿人类“阅读理解”的思维方式，依托法律专家设计的关联分析案例库，主要实现对具体涉法场景描述的主题理解以及法理逻辑关联。其中，主题识别通过有效提取场景描述中的关键词(句)，利用主题识别模型快速确定该场景描述所对应的主题标签；法理逻辑关联通过知识图谱语义搜索，将搜索反馈的实体及相关语义信息与关联分析案例库中的预设实例进行匹配，综合主题识别对应的知识图谱反馈结果以及案例库匹配实例，能够对用户输入的涉法问题描述进行科学关联分析并反馈相关法理知识依据。

3.3 系统功能演示

系统主界面展示的是法理知识图谱总体情况概览。如图 7 所示，左侧柱形图反映了法理知

表 1 意图识别模板(部分)

自然语义问句结构	问句意图	匹配模板
法理实体的“来源”“解释”……	属性查询	match (n{name: ‘法理实体’}) return n.来源 解释……
法理实体“涉及”“包含”“关联”	关系查询	match (n{name: ‘法理实体’}), p=(n)-[r: 涉及 包含 关联]-(m) return p
法理实体 A 与法理实体 B 之间的直接联系 关键联系	最短路径查询	match(a{name: ‘法理实体 A’}), (b{name: ‘法理实体 B’}), p=shortestPath((a)-[*]-(b)) return p
同时涉及法理实体 A、B 的法条有哪些	并列查询	match (a{name: ‘法理实体 A’}), (b{name: ‘法理实体 B’}), (m: article), p=(a)-[]-(m)-[]-(b) return p
……	……	……

识图谱中的实体、关系数量(试验数据目前共梳理 18710 项法理实体、281954 条法理关系); 饼图反映了法理知识图谱中的实体类型及数量分布; 热力词云反映了法理知识图谱中关联度优先的“关键”实体; 右侧主屏通过 3D 建模实现了对法理知识图谱原貌的部分展示, 侧面说明了对法理知识图谱原貌的部分展示, 侧面说明了如果不能通过合理方式对体量庞大的法理知识进行处理与利用, 不仅难以满足用户快速聚焦其所需的法理依据的需求, 反而会对用户认知造成干扰, 并引发信息过载等负面影响。通过系统主界面搜索框可调用语义搜索功能, 该功能可实现关键词与部分自然语言问句两种模式的语义搜索。当用户搜索“香港”“公司”等法理依据中涉及的实体关键词时, 可快速获取

对应的知识图谱语义关联子图及相关联的文字解读信息; 当用户输入“与公司相关的法理依据”等自然语言问句时, 系统可根据自然语言结构模板理解问句意图, 定位所匹配的知识图谱语义关联子图及相关联的文字解读信息, 为进一步的实体信息挖掘提供参考依据。

系统语义搜索界面如图 8 所示, 右侧为语义关联信息子图, 可点击子图中任意节点获取其在知识图谱中的具体语义信息(通过左侧的信息视窗同步展示)。此外, 信息视窗还提供智能推荐功能, 依托后端知识图谱模式层提供的图算法, 对数据层中的实体数据进行计算, 可实现语义关联子图中隐含信息的有效挖掘, 拓展了语义搜索的深度与广度。

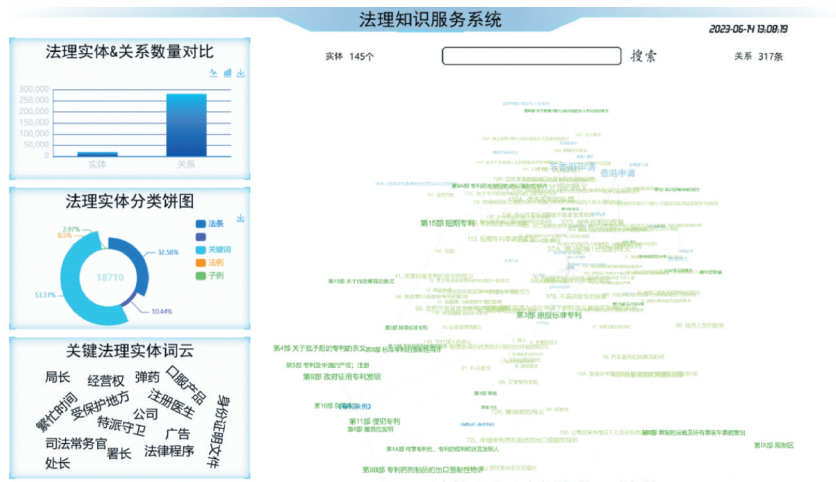


图 7 法理知识服务系统主界面



图 8 法理知识服务系统语义搜索界面

在系统关联分析功能界面中，将涉法场景描述输入关联分析框中，可实现对该场景描述的主题识别与法理逻辑关联。例如，当用户输入的场景描述对应主题识别结果为“可享专利”

时，可准确定位法理知识图谱中与“可享专利”相对应的法理依据，并匹配关联分析案例库中的对应实例，综合知识图谱语义信息及匹配实例，同步呈现关联分析结果，如图9所示。

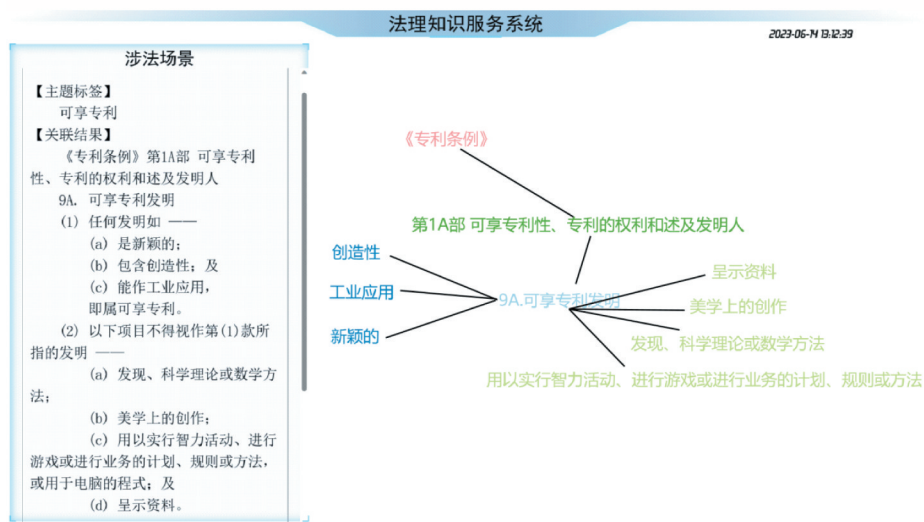


图9 法理知识服务系统关联分析界面

4 结语

本文设计实现了一种基于知识图谱的法理知识服务系统，是适应法律信息化发展的一次工程实践探索。该系统通过有效利用底层构建的法理知识图谱，初步实现了信息管理、语义搜索、智能推荐、关联分析四项核心功能，有助于满足特定用户对于法理知识服务的现实需求。当前，该系统正处于改进优化阶段，通过试运行发现，系统具有较好的复用性与迁移性，能够在底层知识图谱数据不断更新迭代的条件下，实现更加智能、高效的知识服务，并能够向金融、医疗、交通等其他垂直领域进行有效迁移，对于法理知识服务及知识图谱应用研究具有一定的理论参考和实践应用价值。

参考文献：

- [1] AMIT S. Introducing the knowledge graph [R]. America: Official Blog of Google, 2012.
- [2] 李涓子, 侯磊. 知识图谱研究综述[J]. 山西大学

报(自然科学版), 2017, 40(3): 454-459.

- [3] 王昊奋, 漆桂林, 陈华钧. 知识图谱方法、实践与应用[M]. 北京: 电子工业出版社, 2019: 1-2.
- [4] 王一鸣. 基于知识图谱的推荐技术研究及应用[D]. 成都: 电子科技大学, 2018.
- [5] 赵雪芹, 杨一凡, 于文静. 基于Neo4j图数据库的工程档案知识图谱构建及应用[J]. 档案与建设, 2022, (5): 48-51.
- [6] Neo4j. Cypher query language[EB/OL]. [2022-09-26]. <https://neo4j.com/developer/cypher>.
- [7] Neo4j. The Neo4j graph data science library manual v2.1 [EB/OL]. [2022-09-26]. https://neo4j.com/docs/graph_data_science/current.
- [8] 张建平. Vue.js 前端开发实战[M]. 武汉: 华中科技大学出版社, 2021: 58-127.
- [9] BRIN S, PAGE L. The anatomy of a large-scale hypertextual Web search engine [J]. Computer Network and ISDN System, 1998, 30(1-7): 107-117.
- [10] 黄东亮, 戴苏榕, 李辰. 人工智能在OODA循环中的应用[J]. 航空电子技术, 2020, 51(2): 55-59.

Design and implementation of the legal knowledge server system based on knowledge graph

Mo Chengwei, Wen Yifeng*, Deng Bangpeng

(Unit 75610 of PLA, Shenzhen 518000, China)

Abstract: In order to meet the practical needs of specific users for accurate focus, effective association and efficient application of legal knowledge, the information technology is used to expand and optimize the service means of legal knowledge, and implement an information system for legal knowledge service. Based on Knowledge Graph, Vue front-end development framework and other related technologies, and selecting some Hong Kong regulations as the test data, designed and implemented a legal knowledge service system based on Knowledge Graph. Through development and deployment, the system preliminarily realizes four core functions of knowledge management, semantic search, intelligent recommendation and association analysis, which can effectively improve the efficiency of legal knowledge service and provide reference for the application of Knowledge Graph in specific vertical fields.

Keywords: knowledge graph; legal knowledge; design and implementation

(上接第 41 页)

Cross-chain trading mechanism of notary group based on subjective logic and auction algorithm

Liao Zheng*, Wei Bozhong, Luo Jianwei

(School of Automation, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: As a hot spot in the field of blockchain research, cross-chain technology has always attracted much attention. As a typical cross-chain method, the notary mechanism will have problems such as “centralized functions of notary nodes”, “low node credibility”, and “excessive intermediary fees”, for which a cross-chain transaction mechanism of notary group based on subjective logic is proposed. By introducing the method of subjective logic, the mutual evaluation between nodes on the single chain and the transaction information after serving as a formal notary are evaluated, and the notary public group with high credibility is screened out, making the notary mechanism more secure and reliable. At the same time, in order to rationalize the cost of notaries acting as intermediaries, a pricing mechanism based on the reverse Vickrey auction algorithm is proposed, which reduces the price when notary nodes act as intermediaries. Experimental results show that compared with the traditional notary mechanism, the cross-chain transaction mechanism has better security and lower intermediary fees.

Keywords: notary mechanism; blockchain; cross-chain; subjective logic; reputation value sorting; auction algorithm

文章编号: 1007-1423(2023)10-0081-06

DOI: 10.3969/j.issn.1007-1423.2023.10.014

基于微信小程序的在线答题系统的设计与实现

岑俊慷, 刘卫*

(桂林康复疗养中心, 桂林 541000)

摘要: 系统贴合当下用人单位需求, 以在线答题为牵引, 建立能实现“无纸化”反复练习、强化用户学习记忆的微信小程序。以开发者工具为依托, 运用云开发环境下云函数、云存储、云数据库资源。重点围绕单选题+多选题+判断题模式, 通过编译程序代码, 实现页面跳转、数组传值、循环嵌套等模块, 搭建集选题、答题、统分、查阅等多功能于一体的平台。辅以.wxss样式美化, 使小程序更加简洁、完整。

关键词: 微信小程序; 在线答题; 跳转; 数组; 循环

0 引言

如今, 工作、生活的快节奏之下, 人们能支配的时间渐趋“碎片化”, 与传统的品书赏文相比, 关键搜索和直接获取将演变为人们获取知识的主流方式, “手机+微信”成为国人生活的标配, 微信正在主导这场学习方式的变革。大多数用人单位均会定期组织员工进行政治理论、业务素养考试。考试前, 单位会为员工提供题库下载渠道, 由员工自行下载打印背记。若想模拟考试, 还需将题库答案删除或隐藏, 再打印出来练习。但这种方法浪费纸张、浪费时间, 而且无法重复利用, 影响学习效率。根据不完全统计, 在常规题型中, “单选题+多选题+判断题”通常占有60%~100%的比重, 成为考试必出的“三剑客”。基于此, 可以运用微信小程序开发一套在线答题系统, 涵盖单选、多选、判断的所有题目, 可以反复练习加强记忆。答题过程中, 只需轻点一下, 即可随时查阅答案解析, 真正实现“无纸化”学习。该系统是从微信内部衍生出来的应用, 无需额外下载和安装, 只需用微信扫一扫或搜一下即可打开应用, 不占手机内存, 用完即走、无需卸载、方便快捷, 适用于绝大多数上班族群体。

1 系统需求分析

1.1 总体思路

在线答题系统的编程逻辑, 是一个实现集合数组的“循环+选择+比较+截取”, 通过索引自加或自减完成跳转, 以及组件属性条件切换的过程。用户首次进入小程序, 可以注册绑定个人信息。退出后, 再次进入即可直接登录。根据自身需求选择相应题库作答, 同时可切换为“随机”模式。每一题的答案、解析常规下处于隐藏状态, 但可随时调阅求证。每套题库作答完毕可提交统分, 并上传数据库, 方便用户自行比较。思路架构见图1。

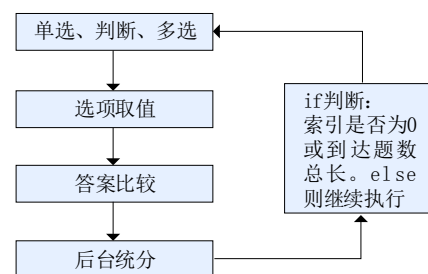


图1 思路架构

收稿日期: 2023-01-29 修稿日期: 2023-02-22

作者简介: 岑俊慷(1988—), 男, 广西桂林人, 助理工程师, 本科, 主要研究方向为计算机软件、编程开发; *通信作者: 刘卫(1984—), 男, 湖南永州人, 工程师, 本科, 主要研究方向为计算机网络、计算机软件

1.2 页面框架

该系统小程序设计至少需要6个页面。其中5个页面是固定的，分别是index、select、score、explain、about，依次代表“主页”“选择(题库)”“记录”“说明”“关于”。而select页面内又可包含 $1\sim+\infty$ 个list页面(以数字下标区分)，用于展示各题库(名称)，为用户选择提供指引。在成功调试list_1.js、list_1.wxml编译的基础上，随着将来题库数量的不断增加，后续的list页面只需套用模块程序即可。作为与用户交互的界面不仅要实现各种功能，同时需要良好的界面设计，方便用户操作^[1]。因此需辅以样式层wxss编译美化UI界面整体设计，提高用户使用体验，力求小程序更加完美。页面框架见图2。

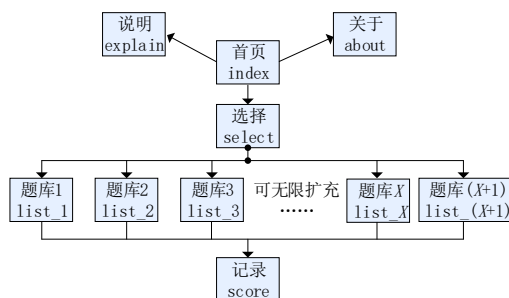


图2 页面框架

2 开发者工具相关应用

该系统主要依托微信开发者工具实现编码设计和调试，总体有40多处组件、函数、参数等相互关联(含事件和自定义组件)，其中使用较为频繁的有以下几项：

2.1 组件

<view>、<text>和<image>虽为该系统常用组件，但它们的用法属于最基础的内容，此处不再赘述。

2.1.1 导航navigator

主要用于index.wxml页面，实现页面与页面之间的相互跳转，以url=“页面路径”作为分页目标，组件内可包含img或icon用以触发跳转^[2]。

```
<navigator url="../../select/select">
<image src="../../img/icon_1.png" mode="widthFix">
</image>
</navigator>
```

2.1.2 视图容器Swiper

主要用于index.wxml页面，实现顶部多图左右滑动，宣扬主题和意境，提高设计内涵，避免模式呆板。同时，在swiper组件中可以放置多个swiper-item组件，然后由swiper组件控制这些swiper-item组件轮换显示。如：

```
<swiper class="swiper" indicator-dots="true" orientation
="horizontal" autoplay="true">
<block wx:for="{{swiperList}}" wx:key="key">
//各图片路径已放在.js的swiperList[]中
<swiper-item><image src="{{item}}" mode="widthFix"/>
</swiper-item>
</block>
</swiper>
```

2.1.3 表单radio-group和button

主要用于list.wxml系列页面。radio-group作为一个容器组件，本身只需以bindchange的事件绑定属性，当radio-group中的选中项发生变化时触发change事件，就会执行bindchange属性中绑定的事件处理函数，通过event.detail.value可得到选中radio组件的值；同理，button的用法也是通过bindtap触发事件处理函数，此处主要用于“上一题”和“下一题”的题目切换^[3]。

2.2 布局

2.2.1 APP配置

除修改主界面“window”样式外，更多是添加页面(分页)。在“pages”:[]内输入每个页面路径，分别以逗号(英文)隔开，如：

```
“pages”:[
“pages/index/index”,
“pages/select/select”,
“pages/select/list_1/list_1”,
“pages/select/list_2/list_2”,
“pages/select/list_3/list_3”,
“pages/score/score”,
“pages/explain/explain”,
“pages/about/about”
]
```

2.2.2 flex排列

弹性布局，实现“控件”(icon)的整齐分

布，为盒状模型提供最大的灵活性。此处主要用于图标的并列规整，使界面自带层次感。按

Flexbox Playground 参照标准^[4]于 .wxss 对应输入属性即可，见图3。

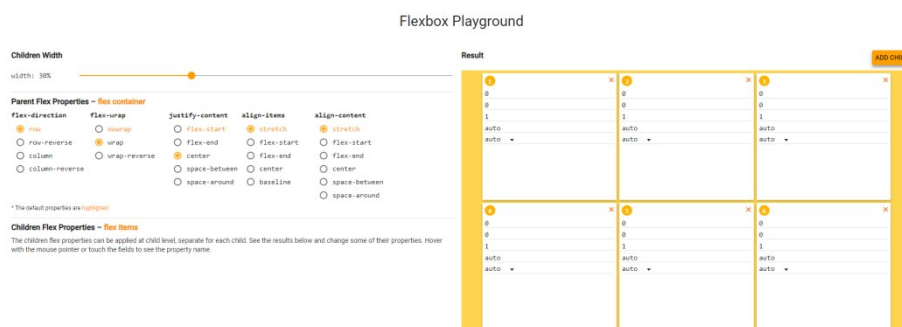


图3 Flexbox Playground标准

2.3 语法

2.3.1 变量初始化

为避免数据混淆和程序漏洞的产生，系统启用之初均要定义参数变量(类型)，进行初始化赋值。如：

```
data: {
  my_id: 0,
  my_tm: "",
  my_da: "",
  pageIndex: 1,
  my_opertion: [],
}
```

2.3.2 wx:for 循环

主要用于列表渲染(各选项的显示)，在组件上用 wx:for 绑定一个数组，数组当前项的变量名默认为 item，当面临多层循环，需要指定字段名时，则可使用 wx:for-item 来定义。如：

```
<label style="display: flex;" wx:for="{{my_opertion}}"
  wx:for-item="rs" wx:key="value">
  <view> <radio value="{{rs.value}}"/> </view>
  <view> {{rs.name}} </view>
</label>
```

2.3.3 数据绑定、传参赋值

在页面对应的 .js 文件中，data 定义数据后，把数据绑定到页面中渲染，使用 2 个 {{ 将变量包起来即可，如 <view> {{要绑定的数据名称}} </view>，主要用于变量的值不断产生变化的情

况，相当于动态绑定。触发事件后，会启动对应的事件处理函数，采用 this.setData({}) 方法进行数据的实时更新，以此来实现视图层对数据的实时更新^[5]，从而实现同一页面下不同题目的相互转换。

3 微信小程序端实现的功能及效果

答题小程序支持账号登录和游客身份进行答题，账号登录状态下可查阅历次作答的记录(含时间、题类、正确率、分数等信息)。用户可通过首页的标识或导航栏跳转到“题库选择”“记录查询”“软件介绍”“关于我们”等页面。

3.1 首页-index 文件

首页顶端搭配醒目标志“Destiny_答题小程序”。增加 1 个 swiper 组件，放置 3 张像素(长宽)相同的图片。组件属性设为“显示面板指示点”“水平方向”“自动切换 swiper-item”。往下依次放置 4 个集合(图片+文字)和 3 个 icon，以 flex 的方式排列，调整 margin、padding 的值和对齐方式。各集合和 icon 均要添加导航组件 navigator 以实现页面跳转。最终首页界面见图4(含 3 幅轮播画面)。

3.2 explain 文件和 about 文件

首页点击“软件介绍(说明)”或“关于”可进入。主要介绍在线答题系统的使用方法和功能，以及开发者个人信息，见图5。此页面主要是图文展示，无过多技术含量，重点在于。

wxss 样式层的不断调整 and 美化。

3.3 选择-select 文件

点击“我的题库”可进入。用于指引的“白手套”为.gif 格式的动态图片，与上方的“请选择题库类型”同样呈 flex 排列、居中。往下分别添加 button 组件(type= “primary”)用以点击跳转至各题库页面，随着将来题库类型的不断增加，可于此页面添加<scroll-view scroll-y= “true” >语句予以延长视图容器(扩容)。页面右上角为 switch 组件，其 checked 属性默认为“false”，即按顺序出题。选中后，checked 属性变为“true”，同时触发事件处理函数，出题顺序切换为“随机”模式，原理和 Math.floor(Math.random()*100)用法^[6]相似。最终选择(题库)界面见图6。

3.4 题库-list 系列

每套题库存放于不同页面，如《党史、新中国史、改革开放史、社会主义发展史》题库

放在 list_1 页面，题目之间的切换也仅在此页面进行，见图7。开头部分会标注该题的题型，帮助用户缩短反应时间。单选题和判断题的选项应有“互斥”特性，多选题则应具备“复选”功能。选项选中后系统会将“check”的值(A/B/C/D)与对应数组的“answer”的值(A/B/C/D)作比较，如果相同，则回答正确(分数“fs”的值+1)，否则回答错误(不加分)。最终分数将在最后一题回答完毕并点提交后显示。在答题过程中，可随时点击“查看答案”调阅该题答案，判断题还附带解析。多选题的统分判断，则是在单选题代码的基础上，添加“答案个数”判断以防止漏选，添加“答案值”逐一比较以允许“乱序”选中^[7](如BCA或CAB)。

3.5 记录-score 文件

此页面需要账号登录才会显示，需要调用数据库。主要是将用户历次答题的时间、题库名称、正确率、最终分数逐条显示出来，见图8。



图4 首页 index 界面



图5 软件介绍(说明)界面



图6 选择(题库)界面



图7 题库 list_1 界面

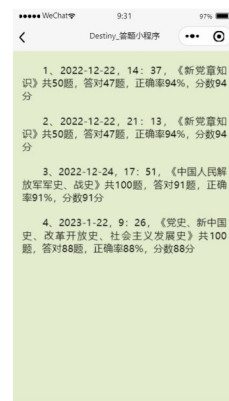


图8 记录 score 界面

4 关键代码

```

onLoad() {
    /* 对页面状态数据初始化,
    生命周期回调-监听页面加载 */
    this.setData({
        my_id: this.data.questions[0].id,
        /* 大数组索引为0,第1道题 */
        my_tm: this.data.questions[0].tm,
        /* 第1道题的题目内容 */
        my_ption: this.data.questions[0].option
        /* 第1道题的所有选项(小数组) */
    })
},
radioChange: function(e) {
    console.log("radio:", e.currentTarget.dataset.msg);
    console.log("radio:", e.detail.value);
    this.setData(
        /* 选中的选项所对应的值(A、B、C、D) */
        this.data.questions[e.currentTarget.dataset.msg].
        check=e.detail.value
    )
    console.log("this.data.questions:", this.data.
    questions);
},
submitUp: function() { /* 函数功能:上一题 */
    if (this.data.pageIndex>0) {
        this.data.pageIndex=this.data.pageIndex-1;
        /* 当前页码>0时,页码数自减1 */
    }
    this.setData({ /* 传参,新值替换旧值 */
        my_id: this.data.questions[this.data.pageIndex].id,
        my_tm: this.data.questions[this.data.pageIndex].tm,
        my_ption: this.data.questions[this.data.page-
        Index].option
    });
},
submitNext: function() { /* 函数功能:下一题 */
    console.log("this.data.questions.length:", this.data.
    questions.length);
    if (this.data.pageIndex<this.data.questions.length-1)
    {
        this.data.pageIndex++; /* 当前页码<最大页码数
        (题目总数)时,页码数自加1 */
    }
    console.log("this.data.pageIndex:", this.data.page-
    Index);
    this.setData({ /* 传参,新值替换旧值 */
        my_id: this.data.questions[this.data.pageIndex].id,

```

```

my_tm: this.data.questions[this.data.pageIndex].tm,
my_ption: this.data.questions[this.data.page-
Index].option
});
}

```

5 使用云开发

微信不仅提供了完整的微信开发者工具开发环境,并可进行云开发,使用云函数直接把数据存入云数据库,为前端开发人员节省了环境搭建的工作量^[8]。

5.1 云函数

在云端运行的代码,微信私有协议天然鉴权,开发者只需编写自身业务逻辑代码^[3]。

“cloudfunctions | 当前环境”目录下新建云函数文件夹“demo”(含index.js),在index.js中输入以下代码:

```

Const cloud = require('wx-server-sdk')
cloud.init({env:cloud.DYNAMIC_CURRENT_ENV})
exports.main=async(event,context)=>{
    const wxContext=cloud.getWXContext()
    return {}
}

```

然后,在“index.js”右键,点击“云函数增量上传:更新文件”,至函数上传成功即可。

5.2 云存储和云数据库

云端文件存储,自带CDN加速,支持在前端直接上传/下载,可在云开发控制台可视化理,见图9(不能使用中文命名)。

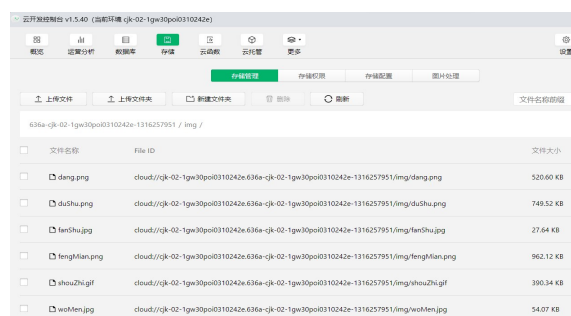


图9 云存储的图片

文档型数据库,稳定可靠;支持在小程序端和云函数中调用,见图10。



图 10 云数据库的list集合(数组)

开发在线答题系统所用到的图片、集合、数组，均可在云存储和云数据库上直接调用，大大节省了搭建后端服务器的时间，这也是云开发功能所提供的强大技术支撑。

6 结语

历经六个寒暑，微信开发者工具从0.12版本(内测版)到现在1.06版本，调试基础库更达到了2.30.0版本，不断更新换代、自我完善。自带封装的接口、组件、函数持续增加，功能越来越强大。工具增添的“热重载”“云开发”等实用选项，为用户提供了更加友好的开发环境，让编程更显“亲民”。上述的答题小程序仅是初级的开发产物。如今，微信小程序已广泛应用于政务、医疗、教育、服务等众多领域，正如微信之父张小龙所说“它实现了应用‘触手可

及’的梦想”，运用微信小程序开发出满足单位或个人需求的系统平台，是将来每名信息技术从业者均应必备的技能之一，是让单位真正实现“互联网+”办公模式的主流渠道。

参考文献：

- [1] 冯保壮, 龚元明. 基于微信小程序的液压抓斗物联云平台开发[J]. 智能计算机与应用, 2022, 12(8): 106-109.
- [2] 高洪涛. 从零开始学微信小程序开发[M]. 北京: 电子工业出版社, 2017.
- [3] 微信开发文档. 小程序开发[EB/OL]. <https://developers.weixin.qq.com/miniprogram/dev/framework/>.
- [4] Flexbox Playground 参照标准[S/OL]. <https://coding.imweb.io/demo/flex/index.html>.
- [5] 唐永祥, 许嘉浩, 乔波. 基于MongoDB和微信小程序的点餐系统的设计与实现[J]. 电脑与信息技术, 2021, 29(1): 25-27.
- [6] 初学者秒懂的随机数的生成[EB/OL]. https://blog.csdn.net/qq_43553184/article/details/84988011.
- [7] 关于微信小程序的多选和全选实现[EB/OL]. https://blog.csdn.net/qq_42365534/article/details/119138035.
- [8] 龙华. 基于Android和微信小程序的实验器材设计[J]. 现代计算机, 2022, 28(2): 97-100.

The design and realization of the online answering system based on Wechat small program

Cen Junkang, Liu Wei*

(Guilin Rehabilitation and Rehabilitation Center, Guilin 541000, China)

Abstract: The system meets the needs of current employers to answer questions on-line as traction, the establishment of “Paperless” Wechat small program repeated practice, strengthen the user’s learning and memory. Based on developer tools, cloud functions, cloud storage and cloud database resources are used in cloud development environment. Focusing on the single-choice + multiple-choice + test mode, through the compiler code, to achieve page jump, array value-passing, circular nesting modules, build a set of topics, answer questions, unified division, access and other multi-functional platform in one. With the .wxss style beautifies, causes the small procedure to be more succinct and complete.

Keywords: Wechat small program; online answer questions; jump; array; loop

文章编号: 1007-1423(2023)10-0087-06

DOI: 10.3969/j.issn.1007-1423.2023.10.015

基于 STM32 的物联网智能家居控制系统

肖榆瀚*, 柯玮翔

(江西理工大学理学院, 赣州 341000)

摘要: 基于人们对物联网智能家居所需功能的普遍性和差异性, 设计了一套基于 STM32 的互联网智能家居控制系统。以 STM32 为主控制器, 结合各模块实现与智能家居系统的交互, 对智能家居的家居安防、家庭监控、家电控制; 结合蓝牙模块实现对智能家居的本地离线控制; 自行设计 Android APP 采用 MQTT 物联网通信协议与 WIFI 模块、云服务器(one NET)进行通信, 以实现智能家居的远程在线控制; 并设计光耦隔离电路以避免低压控制电路和外部高压电路相互造成干扰。

关键词: 物联网; STM32; 智能家居

0 引言

随着嵌入式技术、传感器技术、无线通信技术、智能交互技术的飞速发展以及物联网世界的到来, 智能家居技术也迎来强劲的技术革新。

智能家居概念起源于上世纪 80 年代初的美国, 彼时的智能家居仅仅指代家用电器带来的住宅电子化^[1]; 80 年代中, 智能家居的含义进一步发展为将家用电器、通信设备与安全防范设备综合为一体的住宅自动化; 80 年代末, 通信技术的发展促进智能家居的发展, 出现了通过总线技术对住宅中各设备进行监控与管理的商用系统。智能家居技术在各国不断发展, 不断拓展开发新技术以满足人们对智能家居的安全性需求、舒适性需求、个性化需求, 逐渐从以产品为中心、以场景为中心, 走向以用户为中心。

在我国, 智能家居的发展历经萌芽期、开创期、徘徊期、融合演变期, 并在近几年迎来

爆发期。对比曾经, 智能家居产品从无到有、智能家居技术从发展到革新、智能家居产业从低迷到兴盛。但对比国外, 仍有一定差距, 如缺乏集中化的运行平台、不同智能家居产品之间的交互性、兼容性有待提高等。

近年来, 物联网技术发展迅猛, 逐渐成为智能家居的核心技术, 旨在通过通信网络将所有设备连接起来, 实现家庭信息的交互和共享, 从而构建统一的管理操作系统^[2]。传统的蓝牙、ZigBee、Wi-Fi 等无线通信技术存在通信距离短、功耗大、连接少等缺点, 在一定程度上限制了智能家居的“智能”, 而物联网通信协议为智能家居带来了更多可能。

在本文中我们主要叙述自行设计的基于 STM32 的物联网智能家居控制系统, 通过项目实践研究 MQTT 物联网通讯协议为智能家居带来的功能升级。该系统以 STM32 为主控制器, 结合指纹门禁、智能窗帘、传感器监测等模块实现家居安防; 结合串口屏、语音控制模块实现

收稿日期: 2023-02-19 修稿日期: 2023-02-25

基金项目: 国家级大学生创新训练计划项目(202110407020)

作者简介: *通信作者: 肖榆瀚(2002—), 女, 湖南邵阳人, 在读本科, 研究方向为信息与通信工程, E-mail: 282974173@qq.com; 柯玮翔(2002—), 男, 广东潮汕人, 在读本科, 研究方向为嵌入式软件开发

与控制系统的交互；结合蓝牙模块实现对智能家居的本地控制；结合 Wi-Fi 模块、云服务器、MQTT 物联网通信协议实现 Andriod APP 对智能家居的远程控制；设计光耦隔离电路以避免低压控制电路和外部高压电路相互造成干扰。

1 系统概述

基于 STM32 的物联网智能家居控制系统主要由感知控制层、网络传输层、应用服务层构成，其系统结构框架如图 1 所示。

感知控制层^[3]是实现物联网的基础，主要功能是采集物理世界信息、执行控制指令。系统的感知控制层以 STM32 控制器为核心，感知部分拓展连接有各传感器^[4]，用于感知室内环境的对应参数。控制模块拓展连接有指纹门禁、智能窗帘、马达、继电器等。当传感器监测到的环境参数超过设定阈值时，主控制器执行预

设功能，如自动开启加湿器、控制窗帘等；并通过 ESP8266 无线模块实现物理实体和网络的连接，以进一步将环境参数和警报反馈至应用服务层的 APP 上。

网络传输层是连接感知层和应用层的桥梁，主要功能是接受感知控制层的信息并将信息传送给应用服务层。系统的传输层通过 MQTT 物联网通信协议实现系统的控制电路和 one NET 云服务器的连接，并在二者间进行环境数据和控制命令的传输。

应用服务层即通过远程终端 (Android APP) 实现用户对家居环境参数的实时查询和家居设备的远程控制。

除此之外，在智能家居内部设计有串口屏、语音模块可实现对家居监控和家电控制的多样化近端交互。在无网络即离线情况下，该系统也设计有 Android APP 对家居的蓝牙控制。

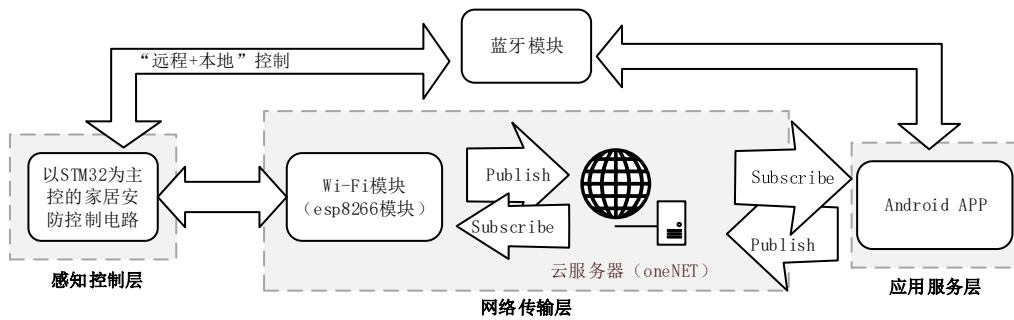


图 1 系统结构框架

2 MQTT 通信协议

消息队列遥测传输协议^[5] (message queuing telemetry transport, MQTT)，是一种基于发布/订阅 (publish/subscribe) 模式的“轻量级”通讯协议，该协议构建于 TCP/IP 协议上，属于应用层。MQTT Broker 也称为 MQTT 消息服务器。MQTT Broker 负责接收来自客户端的网络连接，并处理客户端的订阅/取消订阅 (subscribe/unsubscribe)、消息发布 (publish) 请求。MQTT 客户端 (client) 既可以向服务端发布消息，也可以订阅

消息，因而 MQTT 客户端的身份可以是消息发布者，也可以是消息订阅者。订阅者订阅一个主题如“House”，当消息发布者向服务器发送主题同为“House”的信息，服务器就会将这个信息转发给订阅者，进而实现通信。

3 系统设计

系统设计主要包括智能家居控制电路以及远近端的控制设计，其中智能家居控制电路又包含监测装置、执行装置、人机交互装置，结构框架如图 2 所示。

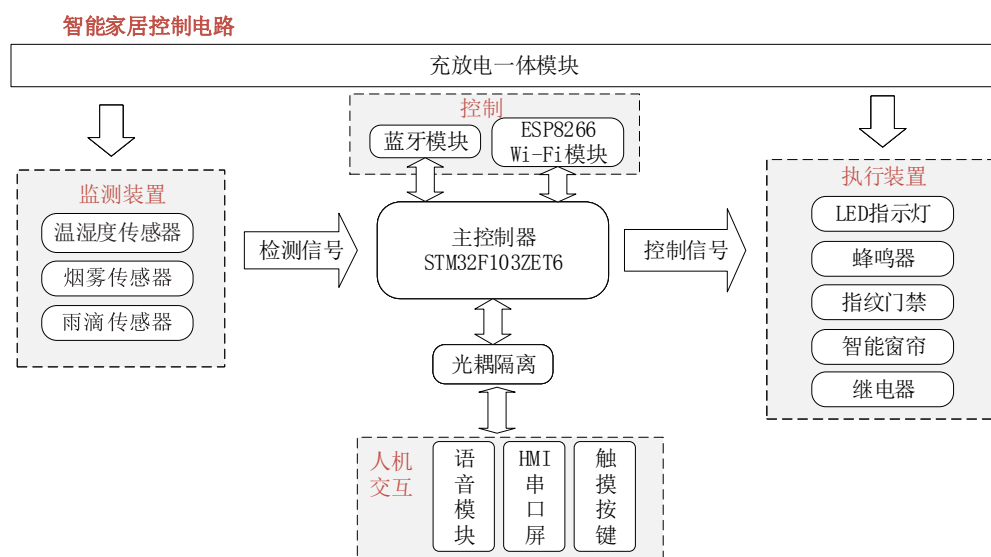


图2 智能家居控制电路结构框架

3.1 监测装置

监测装置包括DHT11(温湿度传感器)、MQ-2(烟雾传感器)、雨滴传感器,实现对室内温度、湿度、空气中可燃气体等数据的获取。执行装置包括指纹门禁、智能窗帘。在此基础上,用户可以根据自己需求接入不同的设备,如电灯、风扇、空调等,通过继电器即可对其实现控制。仅仅需要简单的电路上的接入以及软件上的设置就可以实现对已接入设备的控制。

3.2 执行装置

3.2.1 指纹门禁

指纹门禁功能主要通过电容式半导体指纹识别ATK-301电容模块实现,即通过指纹模块实现判断检索,STM32对指纹模块返回的信息进行读取和处理,然后控制电磁的吸合实现门锁的开与关。

3.2.2 智能窗帘

考虑到市面上的智能窗帘与本项目预期功能有所差异,故自行设计以STC15作为主控的智能窗帘,该智能窗帘采用步进电机作为主要执行器,相比传统利用开关来控制的窗帘,智能窗帘的控制更加高效、精准。

智能窗帘在机械设计方面使用了口径宽于

窗帘杆的挂钩,通过调节挂钩的齿轮进而改变齿轮的高低,通过背板使挂钩稳定位置可使其兼容于市面上各口径的窗帘杆。同时使用封闭半圆设计,使得挂钩不会轻易脱落(如图3所示)。

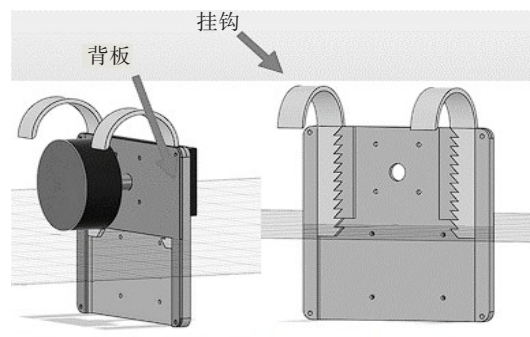


图3 智能窗帘机械结构3D图

手机除了实现对智能窗帘定时开关的控制,还能显示智能窗帘供电电源的电量以便使用者能够及时充电。

由于步进电机需要的输入电压为12V,而ADC采集的电压最大不能超过3.3V,因此以运放LM324为基础,将原有的信号降低4倍,使得恰好符合ADC输入端口所限制的0~3.3V电压范围,同时又极大地发挥出ADC的电压范围。但以上措施保护不具备完善性,因此在放大器的输出端口增加了钳位电路,防止因输入电压

过高导致输出电压超过3.3 V。

通过查找18650电池的电池电压与电量之间的关系，得到关系表格，后使用多项式拟合得到以下关系，如图4所示。

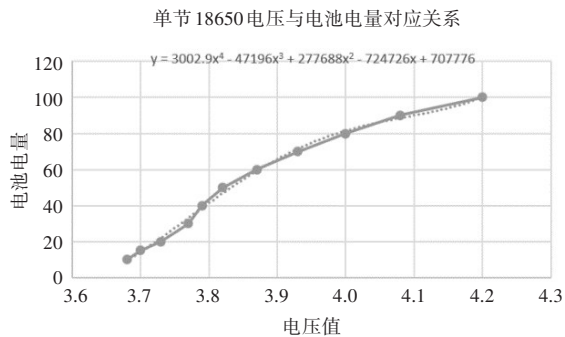


图4 18650电池电压与电量关系

此外，为了避免智能窗帘每次掉电都需重新启动初始化，采用EEPROM记忆元件存储、读取参数。

3.3 人机交互

3.3.1 触摸按键

触摸按键功能通过电容传感TTP224B模块实现。该装置内置了接触式传感器的调节器，静态传感方法可以适应多种情况。人机界面通过非导电介质控制面板连接。本项目利用上述特点实现输出继电器控制与电容触摸按键的交互控制。

3.3.2 USART HMI 串口屏

HMI(human machine interface)也称为人机界面。人机界面是系统和用户之间进行交互和信息交换的媒介，它实现信息的内部形式与人类可以接受形式之间的转换。串口HMI就是设备封装好HMI的底层功能以后，通过串口(USART RS232)与用户MCU进行交互。USART HMI串口屏的界面可直接在软件上设计和编写，通过SD卡或USB转TTL进行下载。与主控制器的交互通过串口即可实现。

3.3.3 语音交互

语音交互功能通过LD3322模块实现。可用AIoT快速产品化平台(智能公元)对模块进行个性化设计唤醒词、命令词、应答语、每个命令

词对应的控制详情等。

智能物联网(AIoT)在智能公元网站进行开发，共经历创建产品→编辑和下载SDK→编译和下载固件→测试服务四部分。于简单应用而言，前面两步即可。首先是创建产品，选择和LD3322语音模块相对应的纯离线方案，模组为SU-03T；其次是编辑和下载SDK，此部分包括麦克风配置情况、Pin脚配置、唤醒词、命令词、发音人配置、其它配置，等等，并可上传录制的音频文件设置个性化音频、允许模块自主学习等。设置完毕后生成SDK并进行下载，按照步骤结合相应软件和串口烧录相关文件即可。其中通信的关键即是对USRT1_TX引脚的控制和参数进行设计，被指定语音词触发后是否向主控发送数据，若发送则向主控发送什么数据。

3.4 控制

3.4.1 本地控制

本地控制基于蓝牙BLE4.2协议+3.0标准，使用多连模式SPP可以让多个从机设备与一个IOS手机之间实现高速通信，加快控制的响应速度。为防止单片机设备与手机设备使用蓝牙进行数据传输时数据错乱，通过在有效数据中加入帧头0XA5帧尾0X5A、一个字节的校验位、固定数据包长度的方式来达到同步的目的。

3.4.2 远程控制

系统结合Wi-Fi模块、云服务器、MQTT物联网通信协议实现Andriod APP对智能家居的远程控制。ESP8266EX由乐鑫公司开发，拥有完整且自成体系的Wi-Fi网络功能，既能够独立应用，也可以作为从机搭载于其它主机MCU运行。内置TCP/IP、MQTT多种协议栈，支持Station/SoftAP/SoftAP+Station三种Wi-Fi模式。可通过ESP8266系列Wi-Fi模组内置的One NET的MQTT协议，轻松接入One NET，并进行数据收发，topic发布和订阅以及接收。

考虑便利性和简易性，我们使用APP Inventor开发Andriod APP。Google APP Inventor是一个完全在线开发的Android编程环境，抛弃复杂的程式代码而使用积木式的堆叠法来完成Android的开发。云服务器使用one NET中国移动物联网开放平台。

ESP8266 Wi-Fi 模块和云服务器的连接通过 AT 指令实现。主要指令有：

```
AT+CWJAP=ssid,password 接入 AP
AT+IOTCFG=devid,proid,auth_info 登录 one NET
```

其中：devid 为设备 ID,proid 为产品 ID,auth_info 为鉴权信息。

```
AT+IOTSUB=xxxx 订阅 Topic
AT+IOTPUB=topic,msg 发布 Topic
AT+IOTSEND=a,b,c 发送数据
```

其中：a 为数据类型 (=0 数值类型；=1 字符串类型)，b 为数据流名，c 为数据值。

Andriod APP 和云服务器的连接则是通过 APP Inventor 开发平台的拓展 MQTT 组件实现。先连接云服务器的 IP 地址 183.230.40.39，进而通过堆叠 MQTT 相关代码块以实现和云服务器的连接、主题的订阅、消息的发送等功能。部分代码块如图 5 所示。

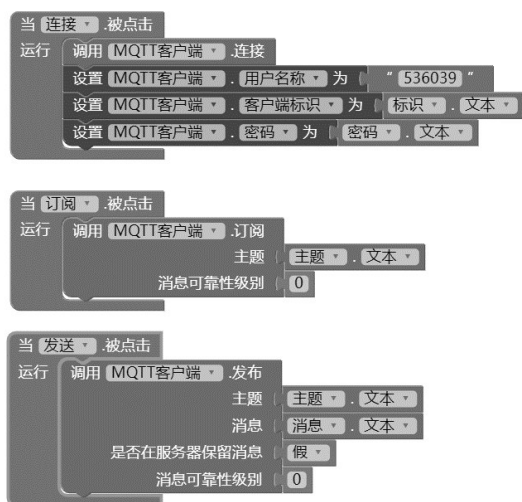


图 5 APP Inventor 代码块

用户名称对应 one NET 平台的产品 ID；客户端标识对应 one NET 平台的设备 ID；密码则是对应 one NET 平台对应设备的鉴权信息。

Wi-Fi 模块和 Andriod APP 作为 one NET 云服务器的客户端，三者通过主题的发布、订阅实现通信。可通过 one NET 官方提供的 MQTT Stimulate Device 和自己的设计的 Andriod APP 对该功能进行测试。二者分别连接同一产品的两

个设备：House1 和 House2→House2 订阅名为 main 的主题→House1 向 main 主题发布消息→House 2 实时接收该消息。

3.5 光耦隔离

在电路安全隐患方面，为防止其中一个设备损坏而引起整个系统工作失效，设计有光耦隔离电路。在本设计中，高速光耦 6N137 两端供电完全不同，一方面兼容了不同通信电压如：3.3 V、5 V、12 V 设备之间的通信，另一方面最大程度上保证了通信过程中两个设备之间的安全。每隔一段时间检查设备是否离线或失效，并将此信息发送给用户的手机端，提醒用户某个设备损坏。在此隔离电路硬件设计方面，使用了插件，更加方便了用户的更换，节约大量成本。

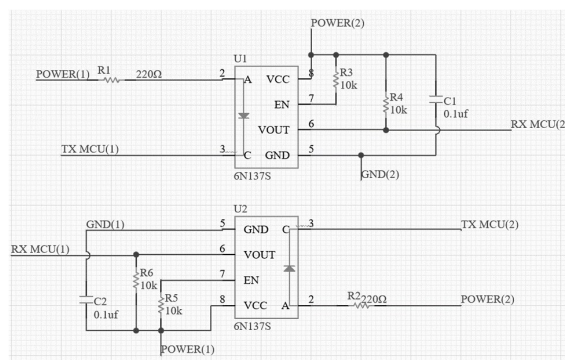


图 6 光耦隔离原理

4 系统软件设计

所有程序采用中断方式进行数据处理，主程序中仅外设初始化一次，后进入空循环。为了节省 CPU 算力，减少 CPU 程序跑飞的可能性，所有操作均位于中断中执行。其中包含：定时器溢出中断、电平中断、数据接收中断。

电平中断：处理由触摸按键引起的电平中断，改变需要控制设备的开启或关闭标志，可燃气体检测，若检测存在可燃气，则发出警报。

数据接收中断：对手机、HMI 串口屏、语言模块传输过来的数据信息的处理。

定时器溢出中断：每隔一定时间，依次执行单总线通信对温湿度获取、使用 ADC 对电压

获取从而得到电量、将数据传至 One Net 上、通过标志位开与闭对执行装置进行控制。

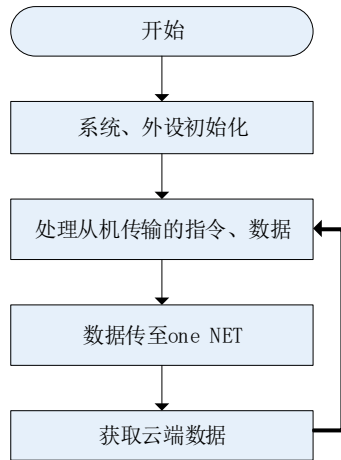


图 7 程序框架

5 结语

本文设计了基于 STM32 的物联网智能家居控制系统，该系统以 STM32 为主控制器，集家居安防、多样化控制、网络智能于一体的智能家居控制系统。经测试，该系统拥有丰富交互方式、多样化功能、个性化设计供使用者选择，

支持本地控制，支持基于 MQTT 协议的远程控制，可广泛适用于普通住宅家庭。但是整个系统在功能上还有待完善：

(1) 目前采用的是监测模块直接将监测信息输入到主控，主控进而将收集到的数据通过 ESP8266 Wi-Fi 模块接入物联网。可进一步加入 Wi-Fi 模块将每个监测模块作为 ESP8266 的从机，直接与主机进行通信进而接入物联网。

(2) APP Inventor 设计的 Andriod APP 有自身局限性，而通过编程原生开发 APP 具有功能强大、用户体验性好等特点。

参考文献：

- [1] 朱敏玲, 李宁. 智能家居发展现状及未来浅析[J]. 电视技术, 2015, 39(4): 82-85, 96.
- [2] 张侠丹. 中国智能家居行业研究[J]. 未来与发展, 2021, 45(12): 14-19.
- [3] 廉小亲, 安飒, 王俐伟, 等. 智能家居发展及关键技术综述[J]. 测控技术, 2018, 37(11): 1-4, 15.
- [4] 孙亭, 杨永田, 李立宏. 无线传感器网络技术发展现状[J]. 电子技术应用, 2006(6): 1-5, 11.
- [5] 任亨, 马跃, 杨海波, 等. 基于 MQTT 协议的消息推送服务器[J]. 计算机系统应用, 2014, 23(3): 77-82.

Intelligent home control system for IoT based on STM32

Xiao Yujing*, Ke Weixiang

(School of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: Based on the universality and difference of the functions people need for the smart home of the Internet of Things, this project designs a set of Internet smart home control system based on STM32. The system uses STM32 as the main controller and combines each module to realize the interaction with the smart home system, home security, home monitoring and home appliance control of the smart home. Combining Bluetooth module to realize the local offline control of smart home; the self-designed Android APP uses MQTT Internet of Things communication protocol to communicate with WIFI module and cloud server(One NET), so as to realize remote online control of smart home. The optical coupling isolation circuit is designed to avoid the interference between the low voltage control circuit and the external high voltage circuit.

Keywords: Internet of Things; STM32; intelligent home

文章编号: 1007-1423(2023)10-0093-06

DOI: 10.3969/j.issn.1007-1423.2023.10.016

汽车零部件订单智能管理系统开发设计

黄梅香*

(柳州铁道职业技术学院建筑技术学院, 柳州 545000)

摘要: 在新一代信息技术迅猛发展背景下, 设计并开发出一个适合中小汽车零部件企业的订单智能管理系统, 主要实现汽车零部件交易过程中对采购订单、销售订单进行智能化管理, 包括订单的实时智能推送, 订单及订单相关资源的自动匹配, 汽车零部件信息的实时更新, 供货商、买家和企业员工的信息管理, 销售业绩的统计分析以及不同员工对系统访问权限的设置等功能, 帮助汽车零部件企业提升订单管理效率, 节约成本。

关键词: 订单智能管理系统; 系统设计; 汽车零部件

0 引言

我国经济的快速发展, 带动了汽车行业的蓬勃发展, 汽车行业主要特点是汽车零部件的数量、种类繁多, 然而, 大部分零部件不是由整车厂自己制造, 而是国内的汽车零部件中小企业生产。通常零配件企业都是按照整车厂的订单进行生产, 订单管理的好坏直接影响着管理成本的高低。现有的订单管理系统过于老旧: 整车厂随时根据市场需求调整订单, 同一订单前后调整没有建立任何联系; 订单没有配套订单相关产品的图纸、文件、材料情况, 需人工查找, 既费时又不准, 造成汽配企业物料、库存等资源的浪费。随着市场竞争日益激烈, 现有的订单管理系统越来越无法适应现代企业的要求。

系统利用人工智能和数据挖掘技术, 帮助企业实现订单与订单相应材料的自动匹配、对同一订单相互关联并突出标明修改之处, 大大节约汽配企业的时间、提高管理效率、降低成本, 为其他中小型企业的订单智能化系统建设提供借鉴。

1 系统总体设计

1.1 系统需求分析

订单管理是公司各种业务处理的数据源, 一个好的订单管理系统可以实现以订单定计划、以订单定产量、以订单定采购等多种业务, 因此订单管理在整个供应链管理中处于关键地位^[1]。订单管理主要是对订单进行全方位的管理, 既包括对订单的接收、审核、处理(订单所需原材料的自动匹配)以及自动识别订单前后的变动情况并进行记录, 实现对订单完成情况的实时跟踪和调整, 也包括对客户信息和产品信息的管理。并且, 该系统还需要具备完善的查询功能, 便于工作人员随时随地利用本系统对订单进行实时的在线查询与编辑、发货标记处理、商品出入库登记等操作, 使订单处理过程变得智能、简单、准确、高效, 帮助企业对订单实现订单所需材料的自动匹配、对同一订单相互关联并突出标明修改之处, 可以节约管理时间、提高管理效率, 从而降低成本。

收稿日期: 2023-01-31 修稿日期: 2023-02-23

基金项目: 2022 年度柳州铁道职业技术学院校级科技课题立项项目(2022-KJC07): 汽车零部件订单管理智能系统研发——以柳州悠进电装有限公司为例

作者简介: *通信作者: 黄梅香(1981—), 女, 广西全州人, 高级工程师(信息系统项目管理师), 硕士研究生, 研究方向为计算机技术, E-mail: 66659563@qq.com

1.2 系统主要功能模块设计

订单管理智能系统，主要包括管理系统和零件存取设备两大部分，零件存取设备与管理系统通过数据连接。

1.2.1 管理系统设计

管理系统的功能模块主要包括系统管理模块、物资管理模块、生产管理模块、质量管理模块、仓库管理模块、售后服务管理模块共六大模块。系统管理模块主要用于用户管理和权限管理，也可用于控制其他各模块；物资管理

模块主要用于订单管理和物料管理；生产管理模块主要用于生成零件配件二维条码和储存生产数据；质量管理模块主要用于储存质量数据和查询数据；仓库管理模块主要用于零件配件入库和出库管理；售后服务管理模块主要用于零件配件的售后信息查询和信息回复；各模块分别与中央处理器连接。管理系统还包括用于保存录像数据的录像数据库，录像数据库也与中央处理器电连接。订单管理系统的各功能模块如图 1 所示。

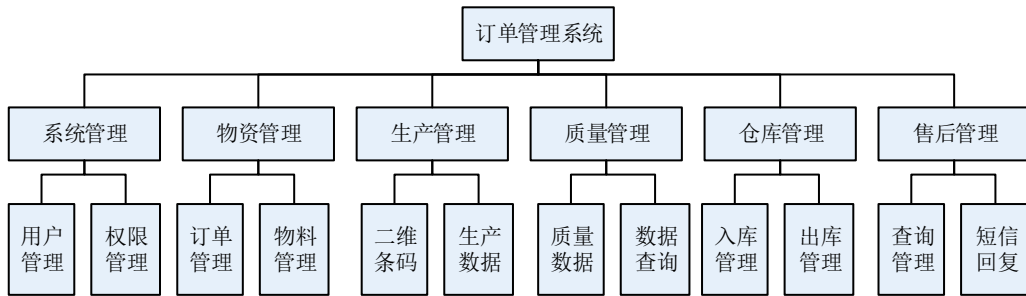


图 1 订单管理系统功能模块设计

1.2.2 零件存取设备设计

零件存取设备包括用于存放零件的库存柜，用于显示数据信息的显示屏，用于输入数据的输入装置，用于与管理系统数据连接的接口，用于控制显示屏、输入装置以及接口的数据处理器，用于扫描零件二维码的 CCD 扫描器以及监控器、验证装置以及当输入密码次数超过指定次数的报警系统等。具体设计结构为：显示屏、输入装置、数据处理器安装于库存柜上；库存柜包括柜体、柜门和电子锁；柜门设置在柜体正面，电子锁用于将柜门锁于柜体上；显示屏、输入装置、接口、电子锁、CCD 扫描器、监控器以及验证装置分别与数据处理器电连接。

2 系统主要功能实现

2.1 管理系统的主要功能实现

2.1.1 系统管理模块

系统管理模块主要是对用户进行管理，提供用户权限的设置，用户的增加、修改和删除等，并提供用户密码的设置、修改以及协同管

理各管理模块的功能。

2.1.2 物资管理模块

物资管理模块主要对业务的订单和物资进行管理，按照订单的需求对订单量进行增加、修改或取消，对订单中零件适用的车型、图纸、物料等进行自动匹配，在客户与订单之间建立信息关联，如图 2 所示。并且提供对订单信息进行查询的服务，订单信息查询功能既可以对所有订单进行查询，也可以对订单中单个订单信息进行查询，将查询结果推送到包含关键字(如含物料，工期空闲等)相匹配的企业。

2.1.3 生产管理模块

生产管理模块主要是通过扫描生产线上的零配件二维码，生成生产管理数据表并在界面显示出相关条码信息。其中，二维码模块包括产品二维码的生成页面和包装二维码的生成页面。主要是根据订单编号、产品编号、产品名称、生产日期、生产场地和适用车型等信息，通过后台二维码生成算法，生成 Data Matrix 图片 [2]。除了自动扫描以外，为了避免扫描码因

褶皱或不清造成扫描不成功的问题，也可以通过手动增添、修改或删除生产数据，同时提供查询生产数据的功能。

2.1.4 质量管理模块

质量管理模块主要提供零件检测数据，用户选择了日期之后，系统会根据日期自动查询数据库，如果查询结果含有当天的生产数据，系统则自动将生产数据包括产品的名称、图号、批次、流向等导入页面，这些信息也可在半成

品追溯中进行查看(如图3所示)；如果没有查询日期，生产数据同样会弹出提示对话框提醒用户结束检测。质量管理模块还提供相关数据的查验、修改、删除和异常提醒功能(如图4所示)，涵盖生产整个过程，包括原材料追溯(如图5所示)、加工过程追溯、产成品追溯等。

2.1.5 仓库管理模块

仓库管理模块主要用于对产品进行入库与出货跟踪管理，零件完成检测并包装以后，通



图2 客户物料关系

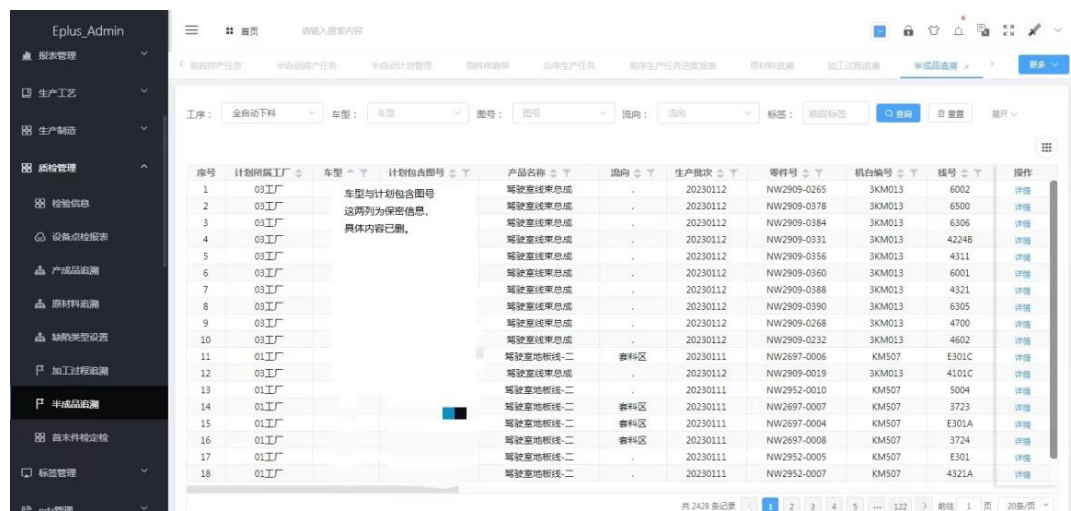


图3 半成品数据查询

物料号: 20402010073 批次: 20221228 流水号: 05050043

工序类型	设备号	批次	看板号	任务量	完成量	操作人	操作时间
全自动	KM707	20221227	DSXSF-X-0046	250	250		2022-12-29 01:14:54
全自动	KM403	20221227	DSXSF-X-0045	250	250	此列为保密信息, 具体内容已删。	2022-12-29 01:19:23
全自动	KM403	20221227	DSXSF-X-0044	250	250		2022-12-29 01:26:32
全自动	KM207	20230111	DSXSF-X-0044	500	500		2023-01-12 04:32:42

资源信息
 检验信息
 物料信息
 异常消耗
 设备信息
 缺陷及返修信息
 工艺信息
 标签信息

模具编号	模具名称	现压接次数	当前压接次数
4105	8240-0552	608103	603053

图 4 异常信息提醒

Eplur_Admin

生产材料追溯

标签: 原材料标签 物料号: 物料号 批次: 批次 流水号: 流水号

序号	物料号	批次	规格	流水号	数量	入库时间	入库人	出库时间	出库人	操作
1	20402010059	2023011201	AVSS 0.3 黄	001	7	2023-01-12 08:...				
2	20402010073	20221228	AVSS 0.3 灰	05050043	2207	2023-01-12 04:...	此列为保密信息, 具体内容已删。	2022-12-31 22:...		此列为保密信息, 具体内容已删。
3	20402010073	20221228	AVSS 0.3 灰	05050043	2207	2023-01-12 04:...		2022-12-31 22:...		
4	20402010087	2023011201	AVSS 0.3 兰	001	500	2023-01-12 04:...				
5	20402010126	20220816	AVSS 0.3 棕	03110004	1650	2023-01-12 09:...		2022-12-31 21:...		
6	20402010126	20220816	AVSS 0.3 棕	03110004	1650	2023-01-12 09:...		2022-12-31 22:...		
7	20402030040	20221011	AVSS 0.5 黑	01150042	140	2023-01-12 09:...				
8	20402030104	20221027	AVSS 0.5 绿	07040002	0	2023-01-12 02:...		2022-12-31 21:...		
9	20402030104	20221027	AVSS 0.5 绿	07040002	0	2023-01-12 02:...		2022-12-31 22:...		
10	20402030104	20221027	AVSS 0.5 绿	07040002	0	2023-01-12 02:...		2022-12-31 22:...		
11	20402030118	20220507	AVSS 0.5 浅绿	03110012	1014	2023-01-12 08:...				
12	20402030118	2022082701	AVSS 0.5 浅绿	001	61	2023-01-12 03:...		2022-12-31 22:...		
13	20402030118	2022082701	AVSS 0.5 浅绿	001	61	2023-01-12 03:...		2022-12-31 22:...		
14	20402030143	20220824	AVSS 0.5 棕	0001	29	2023-01-12 00:...				
15	20402060144	20220622	AVSS 1.25 棕	04080002	103	2023-01-12 00:...				
16	20402060144	20221209	AVSS 1.25 棕	04080003	514	2023-01-12 01:...				
17	20402060144	20221209	AVSS 1.25 棕	04080001	302	2023-01-12 01:...				
18	20402060144	2022122101	AVSS 1.25 棕	001	0	2023-01-12 01:...				
19	20402250045	20230106	AVSSH 0.5F 黄	0001	0	2023-01-12 00:...				

共 649 条记录

图 5 原材料追溯

过扫描器对包装二维码进行扫描, 进入仓库管理系统。产品包装箱入库、出库或库存盘点时通过扫描二维码, 可以准确获取产品的相关信息, 包括生产日期、库存数量、零件型号等。

2.1.6 售后服务管理模块

售后服务管理模块, 对于需要了解二维码信息的客户, 提供对零件二维码信息的查询服务。客户可以通过移动网络发送零件二维码到管理系统获取相关零件二维码信息, 也可以直接通过 Internet 登录公司指定的网站, 登录管理系统将零件二维码的相关信息发送给客户。

2.2 零件存取设备的主要功能实现

零件存取设备的主要功能实现如图 6 所示。零件存取设备的显示屏 4 受数据处理器控制, 当显示屏 4 被点亮后, 显示零件型号列表; 员工通

过触摸显示屏 4 或通过输入装置 3 控制, 挑选所需零件型号并确认所需数量, 输入密码, 数据处理器确认密码正确, 打开电子锁 8; 员工打开柜门 2, 领取所需的零件; 领用后关闭柜门 2, 电子锁 8 自动上锁, 即完成零件的领用。

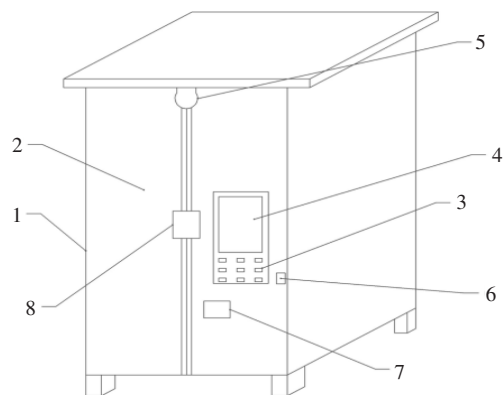


图 6 零件存取设备的主要功能实现

3 系统功能实现采用的主要技术

系统主要围绕如何解决汽车零配件行业订单管理中存在的问题而设置方案进行系统开发，在系统开发过程中应用的主要技术如下：

(1) 确定智能推送与订单匹配系统用以进行深度学习的基础框架：输入输出维数与损失函数。进入实地详细了解用以评价企业订单的各项参数，用状态空间法进行表示，以得到维数。在此基础上进一步了解各项参数对评判的影响基准，以此构建相关损失函数，便于后续建立深度学习框架时进行反向求导的正确性。

(2) 系统采用了最新的基于 Pytorch 库函数的架构，在 Linux 系统上通过 Python 语言建立性能更好的程序。该库函数集成了大量有关深度学习的调用函数，结构简单清晰，用途明确，便于系统的维护和升级，同时便于第三方人员调试与维护，为跨平台操作提供便利，还便于扩展，为企业未来的发展提供有力保障。

(3) 智能推送系统，建立了以 CRNN(CNN+RNN+CTC)为基础的文字识别系统，通过对订单特定区域的文字识别与事先输入/通过深度学习归纳总结出的关键字进行字符串匹配，将匹配结果为适配的订单推送到包含关键字(如含原材料，工期空闲等)的企业。

(4) 订单适配系统，建立了实时可视化输入系统以输入关键字。在此基础上，建立了一个对所有关键字进行过学习的深度学习分类系统。针对企业要求精准度高、泛用性强的特点，采用了在输入关键字与输出分类之间建立全连接层与池化层的 ANN 神经网络，并通过(1)中确立的状态空间维度与损失函数确定其输入输出层，每一隐含层维数，以及最后用以求各维权重的反向传播函数。针对企业要求速率快、结构清晰，则采用二分树结构对输入的状态空间进行了分类。

4 系统评价及展望

4.1 系统评价

整个系统分为系统管理模块、物资管理模

块、生产管理模块、质量管理模块、仓库管理模块、售后服务管理模块共六大模块，每个模块又细分为两个子模块，该系统最突出的优点在于建立了整车厂与零配件企业间订单的实时智能推送并实现了订单所需相关原材料的自动匹配。系统建成后具备以下三个特点：

(1) 反应速度：本系统每次切换当前页面的时间基本都在 1 s 以内，反应速度非常快。

(2) 可扩展性：此程序总体架构采用 B/S 架构模式，便于系统今后根据业务的发展进行扩展。

(3) 可操作性：该系统界面友好，便于操作。

4.2 系统展望

该系统成功建立了整车厂与零配件企业间订单的实时智能推送，实现了动态协调生产，并将多次调整的同一订单相互关联，突出显示订单前后的变化之处及实时更新，同时建立了订单与相关资源的匹配选择模型，挖掘出订单与制造资源执行过程的最佳匹配资源，最大限度地节省了人工选择匹配的时间及人为因素造成的错误，同时还实现了不同人员的权限设置，保证系统的安全可用。但本系统也存在不足之处，例如实际生产过程中当出现首要匹配原材料库存不足时，该采用哪些备用材料来替代，在这方面的功能目前还未能实现，有待今后的进一步完善。

参考文献：

- [1] 刘哲,姜玲杰.基于ERP的汽车生产企业订单管理系统的设计与实现[J].物流技术,2015,34(9):259-261.
- [2] 付先林.基于Data Matrix二维条码的汽车零件管理系统的研究与实现[D].湖北:湖北工业大学,2012.

(下转第 103 页)

文章编号: 1007-1423(2023)10-0098-06

DOI: 10.3969/j.issn.1007-1423.2023.10.017

基于 Funcode 平台的《华夏宝藏》游戏设计与开发

李晓蕾, 王守佳, 刘钰周, 赵晨贺, 黄瀚霖, 王 猛*

(济宁医学院医学信息工程学院, 日照 276826)

摘要: 基于 Funcode 平台开发一款 2D 历史文化类游戏《华夏宝藏》, 采用面向过程的程序设计方法, 运用 C 语言进行程序编写, 结合 Funcode 平台提供的函数接口、开发环境以及独立编写的函数实现游戏功能。将游戏与历史文物宣传有机结合, 尝试探索新的传播传统文化的途径。

关键词: Funcode; 历史文化; 游戏开发; 软件设计

0 引言

随着我国社会经济的快速发展, 弘扬我国优秀的传统文化, 提升我国文化软实力逐渐提上日程。《中共中央关于制定国民经济和社会发展的第十四个五年规划和二〇三五年远景目标的建议》^[1]中给我们指出了 2035 年的文化发展目标——建成文化强国。游戏作为青少年的主要娱乐形式, 在宣传中华优秀传统文化方面拥有巨大优势。本作品以博物馆陈列的文物为切入点, 以中华文明发展史为脉络, 依托 Funcode 平台制作游戏, 寓教于乐, 在轻松愉悦的氛围中向青少年普及文物知识和传统文化。Funcode 是由上海锐格软件有限公司开发的一款支持 C、C++、Java、Python 等语言的基于游戏和动画界面的程序设计开发平台^[2]。该开发平台为本作品提供了各式素材、API 以及相关页面布局, 简化了游戏的开发过程^[3], 再结合使用自行编写的专用函数, 《华夏宝藏》游戏由此诞生。

1 游戏背景

《华夏宝藏》的游戏设计灵感来自于中央电视台热播的文博探索类综艺节目——《国家宝藏》, 作品选取了九件象征中华优秀传统文化发展演变的历史文物, 将其融入到一个完整的剧情故事中, 使整个游戏更加完整、饱满。在对目前流行的游戏调研后发现, Q 版中国风的游戏形象在青少年群体中被接受程度较高, 因此作品在游戏界面上选用 Q 版中国风更能迎合青少年受众。同时, 作为一款益智过关类游戏, 作品保留了此类游戏一贯的递进式游戏难度, 保证了游戏的可玩性与趣味性。在满足以上条件的基础上, 作品还加入一些幽默诙谐的人物对话, 调节玩家游戏过程中的紧张感。

2 游戏设计

在游戏设计上, 秉持依托中华文明发展脉络溯源而上、以游戏的方式展示文物传承历史的设计理念, 作品以明清为始发地, 以殷商为

收稿日期: 2023-01-14 修稿日期: 2023-02-10

基金项目: 2021 年山东省大学生创新创业训练计划项目(S202110443006)

作者简介: 李晓蕾(2000—), 女, 山东德州人, 本科生, 研究方向为网络安全与软件开发; 王守佳(1999—), 男, 河南濮阳人, 本科生, 研究方向为网络安全与软件开发; 刘钰周(2001—), 男, 山东济南人, 本科生, 研究方向为网络安全与软件开发; 赵晨贺(2000—), 男, 山东邹城人, 本科生, 研究方向为网络安全与软件开发; 黄瀚霖(2000—), 男, 山东青州人, 本科生, 研究方向为网络安全与软件开发; *通信作者: 王猛(1981—), 男, 山东济宁人, 硕士, 副教授, 研究方向为计算机网络及信息安全, E-mail: apollo636@foxmail.com

终点站，以中华文明的重大历史事件为节点，以历史事件中存留的文物为依托，构筑起了《华夏宝藏》这款游戏的主体框架。作品选定明永乐青花三足炉、昭陵六骏浮雕、真珠舍利宝幢、藏医四部医典、敦煌莫高窟、秦跪射武士俑、三圣像、青铜纵目面具以及殷墟YH127甲骨窖穴九件文物来展现中华文明波澜壮阔的历史，为玩家擘画中华文化的宏伟蓝图。

《华夏宝藏》设计了九个不同的游戏关卡，并根据文物出土或馆藏位置的历史名称设计关卡名称和美术风格，如图1所示。为提升游戏的趣味性与可玩性，游戏在设计时就引入技能机制以及文物图鉴模块。同时通过运用对话系统讲述一个文物失窃、玩家了解并把文物找回的小故事，将游戏与文物衔接在一起，将一条清晰的游戏主线展现在玩家面前。



图1 关卡选择界面

《华夏宝藏》通过上述游戏设计，再加上存档、读档系统，共同构成了游戏的基本流程，如图2所示。

3 游戏实现

3.1 UI设计

UI是玩家与游戏交互的基础。《华夏宝藏》力图让游戏的操作变得舒适简单、自由、美观。色彩是玩家对游戏UI设计的第一印象。通过调研发现，不同年龄、不同性别、不同区域的人对色彩的喜好存在明显差别。综合来看，橙黄色区间的暖色系更能被青少年群体接受，另一方面这也与“历史”这一主题相呼应，因此选择橙黄色作为游戏色彩的主基调能更受玩家喜爱，也可以带给玩家更沉浸的游戏体验。字体

是玩家对游戏UI的第二印象，相较于端庄正式的宋体字形而言，选取简洁可爱的圆体字形能带给玩家轻松、活跃的感觉。另外在游戏地图选择界面的左上角，设置有玩家头像、血量显示以及等级显示三个直观展示主角状态的显示框，玩家可以及时、准确地查看当前游戏角色的状态。游戏的主要界面如图3和图4所示。

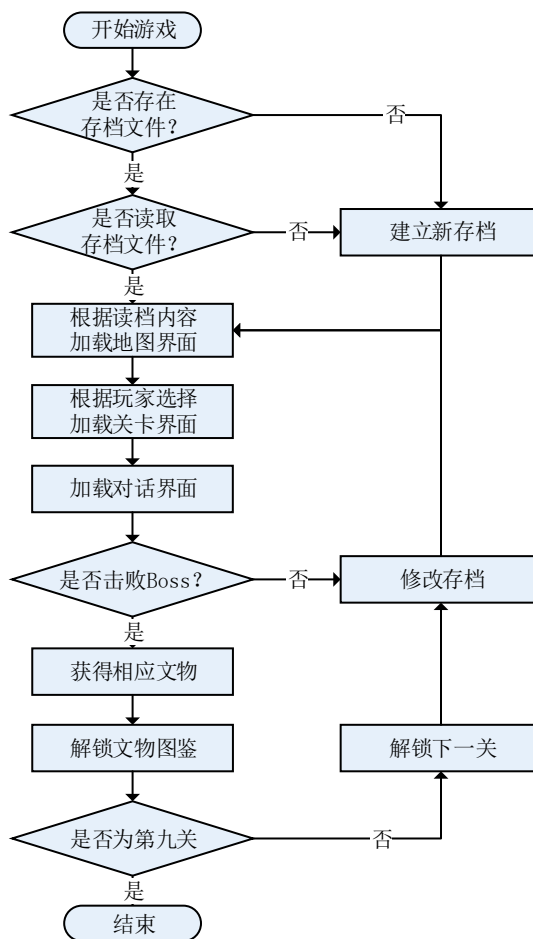


图2 游戏流程



图3 游戏主界面



图4 游戏关卡界面



图6 攻击Boss效果图

3.2 关卡设计

关卡是游戏内容的载体，关卡的数目和内容设计直接影响一款游戏的可玩性和趣味性。《华夏宝藏》共设计了九个不同的关卡，每一个关卡都有不同的小兵以及专属的关卡Boss，玩家需要击败关卡Boss才可成功通关。游戏开始时仅解锁第一关，击败关卡Boss通关成功后才可解锁第二关，后面的关卡以此类推。

《华夏宝藏》设计了关卡选择界面，如图1所示。在关卡选择界面中，玩家可以通过鼠标点击已解锁的关卡进行游戏。关卡的通关难度与游戏的可玩性息息相关，《华夏宝藏》为不同的关卡设置了不同的难度，具体表现为：关卡中小兵和Boss属性的阶梯型增强，保证游戏关卡的难度随关卡的提升而逐步提升。作品也为游戏角色等级成长做了平衡处理，使玩家击杀关卡小兵和Boss为游戏角色提供的经验值随关卡的提升而提升。图5和图6分别为攻击小兵和Boss的效果展示。

每一个关卡中Boss以及小兵的属性数值均以第一关卡中设置的基础属性数值为模板，再经过公式计算之后获得。通过合理设置关卡间属性数值的成长，使小兵和Boss的属性成长较为平滑，以确保游戏的平衡性。如表1所示。



图5 攻击小兵效果图

表1 游戏中关卡Boss和小兵属性数值变化规则

数值类型	小兵	Boss
生命值	生命值= $100 \times \text{当前关卡数}$	生命值= $120 \times \text{当前关卡数} + 130$
攻击力	攻击力= $5 \times \text{当前关卡数} + 5$	攻击力= $10 \times \text{当前关卡数} + 40$
防御力	防御力= $2 \times \text{当前关卡数} + 9$	防御力= $10 \times \text{当前关卡数} + 60$

3.3 存档系统

《华夏宝藏》的可游戏时间较长，拥有角色属性数值成长和收集系统，为了解决玩家一旦中断游戏，游戏数据丢失，玩家再次游戏必须从头开始的问题，就需要为《华夏宝藏》设计存档系统，对游戏数据进行保存。当玩家进入或退出游戏关卡、击败关卡Boss或血量清零挑战失败时，游戏将会自动进行存档。游戏通过stdio.h头文件中的freopen()函数、fopen()函数、fclose()函数、printf()函数、scanf()函数配合实现文件的读写操作，进行关卡进度数据以及游戏角色数据的保存。

存档方式：建立一个文本文件，将要保存的信息按一定规则写入，下次再保存时将会把上次写入的信息覆盖。

读档方式：玩家点击游戏主界面“旧的回忆”按钮，游戏将读取保存在文本文件的存档数据，并根据所读入的数据加载“游戏地图选择界面”的信息，如果读取失败，将加载初始化的“游戏地图选择界面”。

关卡进度数据：此数据类型主要包括当前玩家所解锁的关卡进度。

游戏角色数据：此数据类型主要包括当前游戏角色的等级、经验值、血量等角色属性数据。

3.4 技能系统

技能是一种游戏术语，指游戏中普通攻击之外的产生攻击、移动、辅助增益等效果的主动操作。《华夏宝藏》为游戏角色设计了位移型、治疗型、小范围伤害型和大范围伤害型四种技能，游戏角色在初始状态下(即角色1级时)不具备任何技能，技能将在角色等级达到2级、3级、4级和6级时陆续解锁。

为了确保游戏的可玩性，作品对四种技能设定了不同的冷却时间。位移型技能伤害较低，功能性较强，冷却时间较短，仅为4秒。治疗型技能作为纯功能性技能，冷却时间为5秒。小范围伤害型技能伤害较高，是游戏后期的主要输出手段，拥有6秒冷却时间。大范围伤害型技能伤害极高，拥有长达12秒的冷却时间。通过不同的冷却时间来限制单局游戏技能释放次数，平衡游戏难度、提高游戏的操作性和趣味性。

3.5 对话模块

对话模块是《华夏宝藏》游戏剧情的载体。主要作用是帮助玩家了解游戏的故事背景，引导玩家完成游戏。《华夏宝藏》共设计了18个对话场景，分别在玩家首次遭遇Boss或首次击杀Boss时触发。触发对话模块后，游戏流程立即暂停，游戏将通过访问记录当前关卡数的变量g_WhichGq以及记录Boss血量的变量B_Blood来判断具体加载哪一个对话模块。

对话模块是游戏趣味性的重要组成部分，也是串联历史文物与游戏主线剧情的重要纽带。对话模块的加入放缓了游戏节奏，在玩家紧张的游戏过程中，提供一个相对舒缓的时间窗口，使玩家在游戏过程中产生的紧张感得以释放，为玩家提供更加舒适的游戏体验。

3.6 图鉴系统

《华夏宝藏》的图鉴模块相当于文物收集与介绍系统，是衔接游戏剧情与历史文物的重要机制。

图鉴系统拥有单独的面板，当文物未解锁时，图鉴面板显示为未知状态，如图7所示。只有文物被解锁时，图鉴面板才会显示文物信息，如图8所示。



图7 未解锁的图鉴面板



图8 已解锁的图鉴面板

调查研究博士兼游戏设计师卢卡斯·布莱尔(Lucas Blair)^[4]认为成就系统会影响玩家的行动表现、积极性和态度。《华夏宝藏》采用图鉴收集的方式形成别样的成就系统，以此来刺激玩家行动，提升玩家的游戏乐趣，激发玩家的主动性和收集欲。作品通过在图鉴中运用简洁、生动的文字，活灵活现地介绍文物深厚的历史底蕴和其所代表的博大的文化内涵，让玩家在收集文物图鉴过程中增进对文物的了解，从而达到更好的宣传中华传统文化的效果。

3.7 伤害和经验系统

过高或过低的游戏难度，都会给玩家带来不佳的游戏体验，当这种不佳的游戏体验达到一个阈值，就很可能导致玩家放弃游戏。为了使游戏难度适中，保证玩家在通关的同时具备一定的挑战性，游戏引入了护甲减免系统以及相对完善的经验系统。

游戏为所有小兵和Boss设置了成长性的护甲值，同时引入护甲计算函数，以及攻击力浮动，当玩家操纵的角色受到或造成伤害首先会计算护甲减免值，之后才会计算实际造成的伤害。攻击力浮动则是为了保证攻击伤害数值上下的轻微波动，来增加游戏的不确定性，而不

是每次都造成一个固定的伤害数值。护甲减免函数如以下所示：

实际伤害=

$$\text{攻击伤害} * \frac{1}{1 + \frac{\text{护甲值}}{100}} + (\text{rand}() \% 3 - 2)$$

玩家击败小兵和Boss后将会获得经验奖励，经验积满后人物将会升级，玩家升级后各项数值将会按照成长函数进行提升。

3.8 交互系统

游戏交互系统是玩家良好游戏体验的根本保障。《华夏宝藏》通过游戏中动画、音效等形式将游戏基本信息传递给玩家，使游戏更易上手，帮助玩家快速适应游戏操作。

游戏按钮是否拥有良好的交互性往往决定了玩家对游戏操作性的第一印象，而Funcode平台并没有直接提供实现按钮交互功能的函数。要实现游戏良好的按键交互性，就需要让玩家清楚地知道按钮是否被选中。作品通过将鼠标移动到按钮上时，按钮图片放大显示，以此表示此按钮目前正处于选中状态。在代码层面，作品采用Funcode平台所提供的OnMouseMove()函数获得鼠标位置，使用dIsPointInSprite()函数判断鼠标所在位置是否位于按钮图片范围内，利用dSetSpriteScale()函数改变图片放大倍数。每次鼠标移动时，通过调用dIsPointInSprite()函数判断鼠标是否在按钮图片内。如果鼠标指针在图片内，则将按钮标记置为1，反之按钮标记置为0。当检测到有按钮标记置为1时，立即调用dSetSpriteScale()函数来实现图片的放大显示，同时播放1次鼠标选中音效。反之当检测到有按钮标记置为0时，调用dSetSpriteScale()函数将图片还原至原本大小。当OnMouseUp()函数捕捉到鼠标点击操作时，可以直接通过判断哪个按钮标记为1确定是点击了哪一个按钮，进而确定应该调用什么函数、执行什么操作。

动效和音效有着提升游戏美感，影响玩家视听感受的作用。《华夏宝藏》引入了缓入、缓出、上浮消失等动画效果，这些动画效果广泛应用在游戏界面动态切换，游戏角色、小兵和Boss的攻击效果显示、游戏角色升级展示等方

面。在代码层面，Funcode平台提供了dSpriteMoveTo()函数和dDeleteSprite()函数。dDeleteSprite()函数可以使图片按照给定速度移动到给定坐标点，dSpriteMoveTo()函数可以在游戏中删除图片。利用dSetSpritePositionX()和dSetSpritePositionY()获得图片的位置，并将参数传递给dSpriteMoveTo()函数就可以轻松实现缓入缓出效果。若在短暂延迟后调用dDeleteSprite()函数删除图片，就可以实现上浮消失的动画效果。

4 结语

在游戏风格上，《华夏宝藏》坚持贴合青少年群体的Q版中国风风格，从游戏设计到代码实现均独立完成。素材选择上坚持风格一致的基本原则，所有素材都经过严格的筛选，采用了大量独立制作的图片素材，部分图片及动画素材取自于网络。游戏代码经过反复测试修改，直至游戏运行流畅，无明显bug，基本做到高可读性和高扩展性。游戏数值系统经过多次设计优化，游戏难度设计得愈加合理，游戏的可玩性也大大提高。游戏中加入的独立设计的对话系统丰富了游戏剧情，游戏的趣味性大大加强。《华夏宝藏》采用中国风的游戏设计，通过游戏的方式融入中华悠久历史文化的宣传教育，拓宽了中华优秀传统文化传播的途径，实现了寓教于乐的目的。

参考文献：

- [1] 中共中央关于制定国民经济和社会发展第十个五年规划和二〇三五年远景目标的建议[A/OL]. http://www.gov.cn/zhengce/2020-11/03/content_5556991.htm.
- [2] 刘硕,王海森,李其伟,等.基于Funcode平台的儿童益智类游戏设计与开发[J].软件,2020,41(8):52-55.
- [3] 陈强,刘洋.基于FunCode平台的“太空星战”游戏开发与设计[J].电脑编程技巧与维护,2020(1):150-152.
- [4] BLAIR L. Lucas Blair 对游戏成就系统设计的建议[EB/OL]. <https://www.docin.com/p-544252242.html>.

Design and development of “Huaxia Treasure-Trove” game based on Funcode platform

Li Xiaolei, Wang Shoujia, Liu Yuzhou, Zhao Chenhe, Huang Hanlin, Wang Meng*

(School of Medical Information Engineering, Jining Medical University, Rizhao 276826, China)

Abstract: Developed a 2D historical and cultural game “Huaxia Treasure-Trove” based on the Funcode platform. The game used a process-oriented programming method, used C language for programming, and combined the function interface, development environment and independently written functions provided by the Funcode platform to realize the game function. Combined games with the propaganda of historical relics and tried to explore new ways to spread traditional culture.

Keywords: Funcode; history and culture; game development; software development

(上接第97页)

Development and design of auto parts order intelligent management system

Huang Meixiang*

(College of Building Technology, Liuzhou Railway Vocational Technical College, Liuzhou 545000, China)

Abstract: In the context of the rapid development of the new generation of information technology, an order intelligent management system suitable for small and medium-sized auto parts enterprises has been designed and developed. This system mainly realizes the intelligent management of purchase orders and sales orders in the process of auto parts trading, including real-time intelligent push of orders, automatic matching of orders and order-related resources, real-time update of auto parts information, information management of suppliers, buyers and enterprise employees, statistical analysis of sales performance, and setting of system access rights of different employees help auto parts enterprises improve order management efficiency and save costs.

Keywords: order intelligent management system; system design; auto parts

文章编号: 1007-1423(2023)10-0104-05

DOI: 10.3969/j.issn.1007-1423.2023.10.018

一种电波产品业务平台智能监控系统

刘晓雷*, 刘书志, 李静静

(中国电子科技集团公司第二十二研究所, 青岛 266100)

摘要: 为及时监测电波产品业务平台运行状态, 解决复杂电波业务系统下的故障无法及时判断的问题, 构建了一种基于多 Agent 技术和故障诊断专家系统的智能监控系统。通过分析业务平台架构和业务逻辑, 对核心业务环节构建多层次混合式多 Agent 模型, 设计了基于结果共享的黑板模式多 Agent 协同和通信方法, 从而准确获取业务故障节点运行状态信息。采用基于数据融合的专家系统进行故障诊断, 在构建故障决策树的基础上, 对运行状态信息进行特征向量提取, 通过匹配专家诊断知识表, 完成故障诊断。系统应用后, 能够有效地识别电波业务平台中的故障, 具有良好的业务监控效果。

关键词: 多 Agent; 混合式; 黑板模式; 专家系统; 故障诊断

0 引言

近年来, 多 Agent 技术和故障诊断技术在企业复杂业务系统中的应用逐步增多。多 Agent 架构在处理企事业内部业务系统众多、数据共享、交互能力差等问题时具有良好的效果^[1]。在管理问题复杂性不断提高情况下, 基于规范的 Agent 混合结构模型被提出并验证了可行性^[2]。在多数数据中心的信息系统环境下, 利用多 Agent 技术设计的面向大型分布式系统的智能监控系统能够充分满足运维需求^[3]。在逻辑复杂的监控系统中, 基于故障树技术的专家系统能够有效进行故障诊断^[4]。基于多 Agent 技术和专家系统构造的监控系统, 可以有效诊断网络故障类型^[5]。

本文在上述基础上构建了基于多 Agent 的电波产品业务平台智能监控系统, 系统针对业务平台复杂性、分布式及多网络等特点, 构建了多层次混合式多 Agent 模型, 通过设计给予黑板模式的 Agent 协同和通信协议, 将多 Agent 模型与业务系统运行情况有机结合, 从而实现对业务平台的全方位监控, 能够提前或实时发现故障

问题, 保障了电波产品业务平台持续制备电波环境产品的能力。

1 电波业务平台智能监控系统概述

电波产品业务平台包括了全国多个观测站点的汇集、数据处理、质量控制、产品制备以及产品分发等业务。业务平台具有分布式、多平台、多网络、多业务系统的复杂系统特点, 业务平台下设备众多, 分布在数据单向流通的多张网络当中; 网络之间通过多个数据交换平台进行交互, 存在一定数据依赖性; 数据和产品种类繁多, 业务之间关联性强, 业务流程较为复杂; 不同数据和产品包含各自的多个业务系统, 系统之间耦合性较强。由于该业务平台的复杂化, 传统模式的监控系统难以有效监测业务平台运行状态, 无法适应针对平台特点做到全方面监测, 难以及时发现故障问题。

业务平台智能监控系统监控对象包括分布在三张网络中的多个软件系统、硬件平台、数据和产品、传感器等。监控系统以业务平台在网络 A、网络 B、网络 C 三张网络中各自的数据

收稿日期: 2023-01-12 修稿日期: 2023-02-19

作者简介: *通信作者: 刘晓雷(1990—), 男, 山东烟台人, 工程师, 主要研究方向为应用系统研发, E-mail: liuxiaolei666666@163.com; 刘书志(1988—), 男, 山东诸城人, 高级工程师, 主要研究方向为应用系统研发; 李静静(1984—), 女, 山东潍坊人, 高级工程师, 主要研究方向为应用系统研发

和产品为最终监测目标，确保数据和产品完善，并最终发送到指定用户端。三张网络拥有相对独立的产品制备和发送平台，网络间的数据单向流通并且依次部分依赖。业务平台智能监控系统在此基础上分别对三张网络进行独立监控，并将监控信息单向共享汇集，形成最终诊断结果。

为确保监控平台资源占用最小，系统以用户产品送达为主要监控导向，监控系统分别对业务流程节点、各类数据产品和用户产品进行定时检测，如果发现数据或产品存在问题，则对产品依赖的数据和软硬件平台等预设检测节点进行检测，并将检测后的处理信息通过专家系统进行汇集诊断，最终得出故障信息。

2 基于多 Agent 的监控系统结构模型

传统的多 Agent 系统结构难以实现对复杂网络和逻辑业务平台下的全面监控，本系统结合业务平台特性设计了多层次混合式多 Agent 系统结构模型，能够充分解决复杂电波产品业务平台的故障监控和诊断难题。根据混合式多 Agent 功能结构^[6]将 Agent 分为三类，信息采集级 Agent、管理级 Agent 和全局级 Agent，同类 Agent 之间功能独立互相协作，上级 Agent 对下级 Agent 进行管理。根据多网络结构的业务特点将 Agent 模型分为两个层次，第一层为整体层，由三个子网络多 Agent 系统和全局体 Agent 组成；第二层为三个子网络多 Agent 系统，分别由信息采集 Agent 和管理 Agent 组成，每个子网络多 Agent 系统既能够执行单独的任务，也向全局 Agent 系统进行信息交互，智能监控系统多 Agent 结构模型如图 1 所示。该模型克服了网络数据单向流通和复杂业务平台的障碍，能够较为准确地实现全面的业务平台监控，为故障诊断提供了必需的信息。

2.1 信息采集级 Agent

针对监控对象的类别、网络位置以及所处系统分布位置的区别，设置不同的信息采集级 Agent。监控系统的监控对象包括服务器、网络设备、专用设备、数据库、应用软件、公共平台软件、数据文件、数据产品包、机房环境信息等。在系统中每一类监控对象均有多个，对每一个监控对象的监控要素进行指标属性统计，

将监控属性和模块相同的监控对象设置一个 Agent 进行监控管理，主要设置数据监控 Agent、FTP 监控 Agent、光闸监控 Agent、传输平台监控 Agent、软件监控 Agent、环境监控 Agent、设备监控 Agent，等等。

信息采集级 Agent 的功能是接收上级 Agent 指令，采集预设节点的监控信息，并将收集到的信息与管理 Agent 进行反馈交互，并能根据自我判断或上级指令进行软件和服务重启等操作。信息采集级 Agent 分布在三张网络中，根据每张网络中不同的业务逻辑，设置不同类别的 Agent 对预设节点进行相关数据采集。

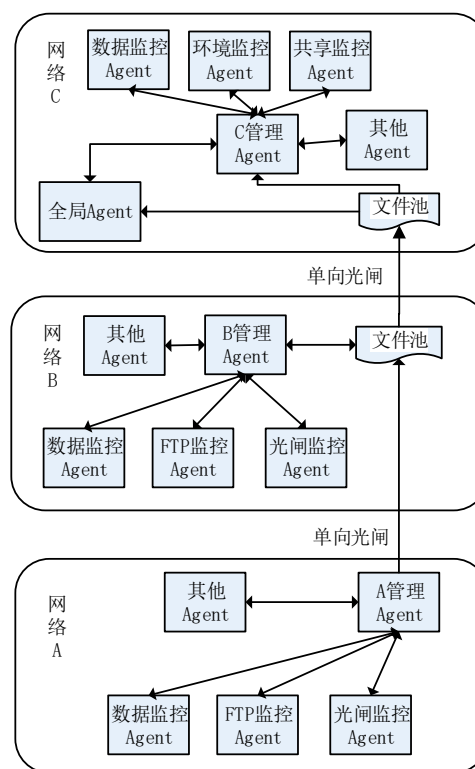


图 1 智能监控系统多 Agent 结构模型

2.2 管理级 Agent

主要负责监控数据和产品的完备性，根据监控结果制定任务并分发给信息采集级 Agent，在收集到的数据进行分析后，与管理级 Agent 和全局级 Agent 进行数据沟通。管理级 Agent 在三张网络中各设置一个，分别负责控制各自网络中的信息采集 Agent，并能根据收集到的信息进行局部故障判断。其中 A 管理 Agent 将处理后的

信息分别发送给B管理Agent、C管理Agent和全局Agent。B管理Agent将处理后的信息发送给C管理Agent和全局Agent。C管理Agent与全局Agent间进行信息交互。

2.3 全局级Agent

负责收集三个管理级Agent的信息、判断系统是否存在故障、诊断出具体故障原因。通过对系统的运行状态进行判断,分析各个预设节点信息,结合历史故障数据,诊断出具体的系统故障问题,并将问题反馈到显示终端。

3 多Agent间的协作

单一的多Agent协同交互模式难以满足复杂业务平台多网络和数据单向流通环境的要求,智能监控系统采用基于结果共享的黑板模式进行多Agent间系统通信。

3.1 Agent工作逻辑

基于多监控系统结构模型的核心工作机制是管理级Agent,主要负责对信息采集Agent的信息进行处理,将任务分解下达,实现各Agent之间的相互协同处理,并将处理结果进行共享,最后通过全局实现对整体目标的控制。

管理级Agent主要包括消息模块、数据模块、控制模块、诊断模块、共享模块。消息模块主要实现Agent之间的通信,实现消息的发送和接收。数据模块用于对黑板数据区域进行数据读写。控制模块用于进行任务决策与分发。诊断模块用于对数据进行分析,判断监控结果。共享模块用于管理级Agent直接的数据分析结果共享。管理级Agent内部工作逻辑如图2所示。

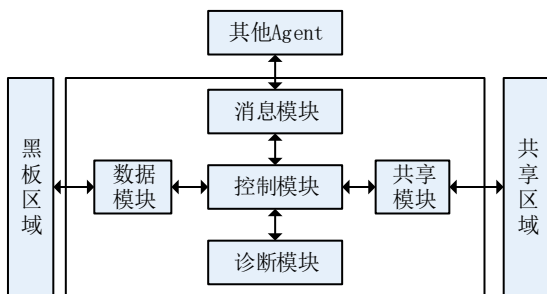


图2 管理Agent工作逻辑

3.2 多Agent通信与协同

同一网络下的管理级Agent和信息采集级Agent之间通信采用黑板模式^[7]进行信息共享,由管理级Agent作为黑板控制Agent,数据监控Agent获取到产品故障信息后反馈到黑板数据区域,管理级Agent从黑板数据区域获取数据进行任务分解,信息采集Agent从管理级Agent获取到任务后采集信息反馈到黑板。为克服大规模信息采集级Agent带来的巨大容量的共享信息,在黑板通信区域按照监控业务逻辑类别,设计了多个消息共享队列,供不同Agent进行读写,黑板通信模型如图3所示。

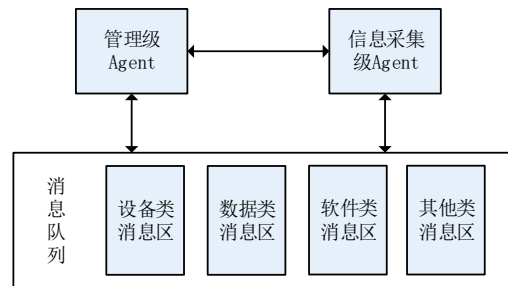


图3 黑板模式通信模型

不同网络下的管理级Agent之间、管理级Agent和全局Agent间采用单向结果共享模式进行协作,A管理Agent通过定时发送心跳包向B管理Agent、C管理Agent和全局Agent进行采集处理信息共享,B管理Agent通过定时发送心跳包向C管理Agent和全局Agent进行信息共享,C管理Agent以黑板模式方式向全局Agent进行信息共享。心跳包以XML文件形式,通过单向光闸设备进行单向数据同步^[8],单向结果共享模型如图4所示。

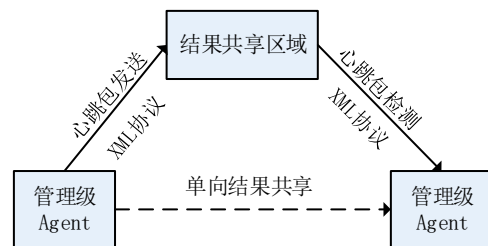


图4 单向结果共享模型

4 故障诊断

电波数据业务平台设备和软件较多，很多故障由于监测节点众多和业务逻辑复杂无法直接诊断故障信息、判定故障来源，同一监测节点信息出现异常甚至会来源于不同的故障。系统采用基于数据融合的专家故障诊断方法，将每个 Agent 提取的设备、软件或业务流程节点信息设为故障点，形成历史系统故障状态集合特征矩阵 Z ，将 Z 按照故障规则进行数据融合，建立历史故障专家知识库，形成专家诊断知识表 Y 。根据故障规则将系统状态集合 X 进行特征提取后与专家诊断知识表 Y 进行匹配，完成智能故障信息诊断^[9]。

4.1 建立专家知识库

第一步收集历史故障记录信息，将历史故障中每个 Agent 提取的信息节点构建对应关系矩阵 Z ，如式(1)所示。

$$Z = \begin{pmatrix} Z_{11} & \cdots & Z_{1m} \\ \vdots & \ddots & \vdots \\ Z_{n1} & \cdots & Z_{nm} \end{pmatrix} \quad (1)$$

其中： Z_{ij} 表示第 i 个 Agent 的第 j 个节点故障信息。取值为0和1，有故障为1，其他为0，节点数不足的用0补位。

第二步根据故障规则建立故障节点决策树，对于具有从属关系的节点建立对应关系，例如 Z_{11} 节点故障会导致 Z_{12} 、 Z_{13} 、 Z_{14} 节点获取到的信息均为故障信息， Z_{11} 就是 Z_{12} 、 Z_{13} 、 Z_{14} 的上级节点。同一节可能会有多个上级节点，在构造决策树时需按照信息最大化原则进行^[10]。

第三步对矩阵 Z 进行数据融合，按照节点决策树对矩阵 Z 进行特征提取，将上级决策树为1的下级节点值设置为0，形成专家诊断知识表 Y 。

4.2 知识推理

将每个故障点的信息节点进行提取，构建系统状态集合形成故障数据特征向量集 X_i ，所有故障数据特征向量集进行合成后形成系统故障状态集合特征矩阵 X ，如式(2)所示。

$$X = \begin{pmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nm} \end{pmatrix} \quad (2)$$

将 X 按照故障节点决策树进行特征提取后，与专家诊断知识表 Y 进行匹配，采用如式(3)所示的几何距离 L 进行故障判定匹配，若完全匹配则能直接判定故障原因；若不能完全匹配，则选取 L 最小值为相近故障。并对故障进行分析记录，记入专家诊断知识表^[11]。知识推理过程如图5所示。

$$L = \sqrt{\sum_{\substack{0 < i < m \\ 0 < j < n}}^n (X_{ij} - Y_{ij})^2} \quad (3)$$

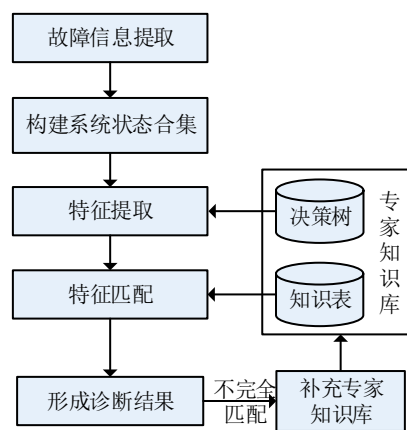


图5 知识推理过程

5 结语

随着我国电波观测业务的发展，电波数据业务平台的构成和功能也在不断复杂化，为保障平台的持续运行，保障电波数据和产品的完整性，智能监控系统尤为重要。本文结合多 Agent 系统和专家系统故障诊断的方法，设计了一种针对复杂数据业务平台的监控系统，可在复杂网络和分布式环境下的数据产品制备系统中充分应用，能够较好地保障业务系统运行的持续性，具备较好的应用基础。

参考文献：

- [1] 杨长春, 蒋冬冬, 吕何新. 基于多 Agent 协同的数据增量交换模型[J]. 计算机应用与软件, 2017, 34(1):85-89, 105.
- [2] 杨海龙, 赵军. 基于规范的 Agent 混合结构模型研究[J]. 计算机应用与软件, 2016, 33(3):6-9, 37.
- [3] 段立明. 面向大型分布式系统的智能监控系统设计

- 与实现[D]. 沈阳:中国科学院大学(中国科学院沈阳计算技术研究所),2018.
- [4] 徐宣,贺国,潘兴隆,等. 某型监控系统故障诊断专家系统知识获取和故障树推理分析[J]. 舰船电子工程,2016,36(12):133-137.
- [5] 史春光,刘波,彭进霖,等. 基于多Agent的分布式IP网故障诊断系统设计与实现[J]. 装备学院学报,2016,27(6):95-99.
- [6] 齐一名. 基于多Agent的磨矿过程智能控制系统研究[D]. 吉林:吉林大学,2014.
- [7] 李帅. 基于WebSocket的多Agent通信机制设计与应用[D]. 大连:大连海事大学,2019.
- [8] 赵荣康,孔祥瑞,梁蓉蓉. 不同安全等级网络之间的数据交换方案研究与实现[J]. 信息安全研究,2020,6(4):338-344.
- [9] 杨明. 自适应迭代滤波分解方法及其在滚动轴承故障诊断中的应用[D]. 长沙:湖南大学,2019.
- [10] 朱红超. 基于粗糙集与决策树理论的救援等级决策算法[J]. 科技信息,2010,(35):530-531.
- [11] 张志勇,秦文超,钟伯录. 基于互联网的集中润滑系统在港口设备中的应用[J]. 港口科技,2017(5):15-21.

An intelligent monitoring system for radio product business platform

Liu Xiaolei*, Liu Shuzhi, Li Jingjing

(The 22nd Research Institute of CETC, Qingdao 266100, China)

Abstract: In order to timely monitor the operation status of radio wave product business platform and solve the problem that the fault under complex radio wave business system cannot be judged in time, an intelligent monitoring system based on multi-agent technology and fault diagnosis expert system is constructed. By analyzing the business platform architecture and business logic, constructs a multi-level hybrid and multi-agent model for the core business links, and designs a multi-agent cooperation and communication method based on the blackboard mode of result sharing, so as to accurately obtain the operation status information of business fault nodes. Adopt expert system based on data fusion for fault diagnosis, based on the fault decision tree, the feature vector of operation state information is extracted, and the fault diagnosis is completed by matching the expert diagnostic knowledge rules. After the application of the system, it can effectively identify the faults in the radio wave service platform and has a good service monitoring effect.

Keywords: multi-agent; hybrid; blackboard model; expert system; fault diagnosis

文章编号: 1007-1423(2023)10-0109-05

DOI: 10.3969/j.issn.1007-1423.2023.10.019

基于 Web 的“湘赣边”线上全景漫游系统开发

邱望^{1,2*}, 熊钰¹, 梁莉菁¹

(1. 萍乡学院信息与计算机工程学院, 萍乡 337000; 2. 江西汇梦网络信息科技有限公司, 萍乡 337000)

摘要:“云旅游”和全景漫游能够为用户带来良好的线上虚拟旅游体验。为促进人们对“湘赣边”区域的了解, 设计实现了线上全景漫游系统。借助 DataV 数据可视化工具和 ECharts 实现网页地图功能; 使用无人机航拍和鱼眼相机采集原始图像; 使用 PTGui 软件合成全景图; 使用 WebGL 技术和 Three.js 组件实现了网页全景漫游和交互功能。

关键词: 线上旅游; 全景漫游; 软件开发; 虚拟现实

0 引言

2021 年, 国家发改委印发的《湘赣边区域合作示范区建设总体方案》指出, 湘赣边区域合作示范区是湘鄂赣革命老区的中心区域, 也是长江流域的重要生态屏障区域。湘赣边区域位于湖南、江西两省交界地带, 包括 24 个县(市、区)。党的十八大以来, 两省在建设湘赣边区域中积极探索, 推动了多项合作。《方案》还提出, 湘赣边区域内有着丰富的旅游资源, 应支持发展智慧旅游, 利用互联网、大数据和虚拟现实(VR)等新技术, “云旅游”成为一种流行的虚拟旅游方式^[1]。

VR 技术作为当前最热门的一项数字信息技术之一, 在旅游业中具有广阔的应用前景。相比于传统的网页图片浏览模式, VR 技术中的 360 全景漫游为浏览者提供了动态交互和完整展示的功能, 具有更好的展示效果, 因此该技术也逐渐被应用到各种虚拟场景展示的功能中。目前, 针对湘赣边两省多地的区域内旅游资源融合发展的应用较少, 人们对区域内的县市有哪些特色旅游资源不了解, 不利于推动区域内旅游资源的整合与协同发展。本文将基于网页

地图可视化和 360 全景技术, 设计一款“湘赣边”线上云旅游系统, 整合区域内旅游资源, 促进人们对“湘赣边”区域的了解^[2]。

1 “湘赣边”数字化旅游系统设计

数字化旅游在很多城市已经有应用, 各项技术也较为成熟, 因此在选取景点时应突出特色, 提升软件的吸引力。另外, 用户体验是软件被用户接受的最重要影响因素之一, 只有充分考虑用户使用软件的舒适度, 软件才能被更多人使用^[3]。

1.1 系统功能需求分析

本系统作为一项展示型的网站, 暂时不需要动态的数据管理功能, 因此主要以前端开发为主。若后期要进行功能拓展, 本系统也可简单地接入后台和数据库。考虑用户的浏览体验效果, 具备以下的主要功能:

(1) 景点地图。在首页中需要展示两省的地图, 包括各地级市的区域。各个市会显示相应的景点数量, 点击跳转到该市的详情页。在地级市的详情页中, 应包含地图和景点列表, 列表中包括简介和全景漫游页面的跳转链接。

收稿日期: 2023-02-05 修稿日期: 2023-03-07

基金项目: 江西省教育厅科技项目(2021B5419)

作者简介: *通信作者: 邱望(1994—), 男, 江西萍乡人, 硕士, 助教, 研究方向为计算机应用、虚拟现实技术, E-mail: 23020040@pxu.edu.cn; 熊钰(2002—), 女, 江西吉安人, 本科生, 研究方向为虚拟现实技术; 梁莉菁(1970—), 女, 江西萍乡人, 本科, 副教授, 研究方向为数字媒体技术

(2) 全景漫游。需要制作全景图片，在网页中以虚拟3D视角的形式展示。用户可以通过鼠标进行交互，控制视角的选择，视野的缩放。一个景区中一般会有多个景点，还应提供场景的切换功能。

1.2 景点的选择

“湘赣边”包含了非常多的景点，过多的景点无法凸显每个地方的特色，在制作全景图时难度也非常大。另外从用户的角度来看，绝大多数用户只对各地最有特色的几个景点感兴趣，过多的景点实际上不会起到效果。因此需要选取各地最具特色、最受人们喜欢的景点。

分析人们的喜好是一个非常复杂的问题，因此可以借助已有的旅游APP数据来选择。美团APP上不同地区的景点排名是经过海量用户数据计算出来的结果，能够较准确地反映出景点的受欢迎程度，可以据此来选择景点。

1.3 区域地理与景点分布的可视化实现

为了制作更精确且功能丰富的地图，使用ECharts和阿里云DataV数据可视化工具进行开发。ECharts是基于JavaScript的数据可视化图表工具，能够在HTML网页中生成定制化的图表，其中也包括地图的显示^[4]。页面中两省的地图分别用2个div标签作为容器，设置合适的样式：

```
<div id="HuNan" style="width: 600px; height:
700px; margin: 0 auto;float: left;"></div>
<div id="JiangXi" style="width: 600px; height:
700px; margin: 0 auto;float: left;"></div>
```

阿里云DataV.GeoAtlas中提供了国内的地理数据，开发者可以通过其提供的数据接口获取到所需地域的地理坐标数据(http://datav.aliyun.com/portal/school/atlas/area_selector)。以江西省为例，选择“江西省”可得到江西省的地理数据API，如图1所示。



图1 DataV.GeoAtlas中的江西省地理数据源

在HTML的script脚本中，通过jQuery插件获取JSON数据：

```
$.getJSON("https://geo.datav.aliyun.com/areas_v3/
bound/360000_full.json", "", function(ret) {
    getArea1(ret);
})
```

getArea1(ret)方法则是使用ECharts设置和显示地图和区域，关键代码如下：

```
function getArea1(ret) {
    var myChart1 = echarts.init(document.getElementById(
        ('JiangXi')))
    let geoJson = ret;
    echarts.registerMap('jiangxi', ret)
    //注册矢量地图数据
    var option = {
        //地图属性设置
        visualMap: {...}, //省略一些视觉设置代码(如颜色)
        tooltip: {
            //悬浮框
            trigger: 'item', //触发条件
            backgroundColor: 'rgba(255, 170, 255, 0.8)',
            formatter: '{b}<br>景点数量: {c}',
            //自定义显示数据
        },
        series: [
            //省略一些CSS样式设置
            {
                data: JX_Info //地市数据
            }
        ]
    }
    myChart1.setOption(option); //echart应用option设置
    myChart1.on("click", function(params) {
        //为区域中地市添加点击跳转事件
        var subSystem = params.name;
        console.log(subSystem);
        if (params.name == "萍乡市") {
            window.open("./PingXiang.html");
        }
    });
}
```

其中，地市数据以JSON格式保存，“value”表示的是该地市的景点数量，

```
const JX_Info = [
    {name: "新余市",value: "5"},{name: "赣州市",
    value: "12"},
    {name: "吉安市",value: "12"},{name: "抚州市",
    value: "0"},
    {name: "鹰潭市",value: "0"},{name: "上饶市",
    value: "0"},
```

```
{name: "景德镇市",value: "0"},{name: "南昌市",
value: "0"},
{name: "宜春市",value: "15"},{name: "萍乡市",
value: "13"},
{name: "九江市",value: "9"},
];
```

以同样的方式可以实现在HTML页面中显示湖南省以及各市的地图，在浏览器中运行即可得到地图效果。

1.4 地市景点详情页

点击地图上的地市则跳转至该市的景点列表页面，页面使用Bootstrap前端框架，以卡片式的页面布局方式呈现各景点简介，具有简洁美观、响应式的优点。效果如图2所示。



图2 萍乡市部分景点列表

点击按钮，进入该景点的全景漫游页面。

2 景点全景图制作

2.1 制作原理和关键技术

普通的图像是将相机固定在某个位置和视角拍摄的相片，一般相机的视角都会小于 120° ，拍摄的图像只能展示场景的一部分。要在网页中看到完整的场景，则需要制作全景图片。全景图片是指从某个点拍摄到的完整场景的图像，先用相机固定在场景中合适的位置，然后旋转角度拍摄相片，这组相片要包含所有视角下的场景。然后利用专业的全景图片处理软件进行拼接、优化和合成。在网页中全景图像需要利

用WebGL提供的相关功能进行交互，还可以使用网页的VR模式在VR设备中进行浏览，达到沉浸式的体验^[5]。

全景图片的制作通常包括图像采集、拼接、优化和合成等步骤。

2.2 图像采集

在采集场景图像时，为了提高采集效率，应尽量减少图片的数量，因此选择高像素、短焦距、大广角的专业相机。由于室内场景和室外场景具有不同的特点，因此设备和拍摄方法也不用，本文分别以“杨岐山”景区入口和“孽龙洞”入口场景为例分别介绍。

“杨岐山”景区入口是户外场景，应侧重于展示全貌，适合采用俯瞰的视角，因此采用无人机航拍的方式进行拍摄。选定景区大门前的位置，通过手机APP控制无人机飞行至距离地面40米左右的高度，根据光线强度调整参数，选择全景模式拍摄即可得到一组全景图片。

“孽龙洞”入口场景应侧重表现地面的洞口和周围事物，适合采用游客视角，因此使用带“鱼镜头”的相机进行拍摄。拍摄时，将三脚架置于合适位置，将相机的高度调整至人眼的高度，之后旋转相机拍摄各视角的场景。由于超大广角的特点，只需要6张水平视角相片+1张天空相片+1张地面相片即可包含该场景的所有内容。

2.3 全景图片的合成

完成图片的采集后，使用PTGui软件对图片进行处理和合成。PTGui导入一组图片时会自动对齐，但在图片的连接处会出现一些畸变和错误，需要添加控制点进行优化调整。最终合成的全景图如图3所示。



图3 “杨岐山”景区入口全景

3 网页全景漫游交互

在实际的制作中,要亲自去拍摄所有的景点是不可能的,因此可以借助其他人在网上公开发布的虚拟场景完善系统。720云平台(<https://www.720yun.com/>)中提供了场景全景漫游功能,只需要上传全景图即可。由于本系统为非商用,因此部分景点链接到了720云平台中公开的場景,但如果要实现商用,则需用其他方法实现该功能。

本系统使用WebGL和Three.js工具实现网页全景漫游功能。WebGL技术能够借助GPU提升浏览器对于图像的渲染速度,更高效地创建复杂的显示效果。Three.js组件能够支持Web网页中全景图片的360°环绕展示,封装了滑动、旋转、点击、切换等事件,用户可以使用鼠标实现漫游功能^[6]。

在HTML中加入div用于放置全景图:

```
<div id="container"></div>
```

在<script>中使用three.js加载、设置和渲染全景图。关键代码如下:

```
<script>
container = document.getElementById( 'container' );
//设置全景图的虚拟相机视角
camera = new THREE.PerspectiveCamera(75, width/
height, 1, 1100 );
camera.target = new THREE.Vector3(0, 0, 0 );
scene = new THREE.Scene();
//将全景图(img.jpg)设置为网页3D图片材质
var material = new THREE.MeshBasicMaterial( {
map: new THREE.TextureLoader().load( 'img.jpg' )
});
mesh = new THREE.Mesh( geometry, material );
//渲染场景
scene.add( mesh ); //网页设置
renderer = new THREE.WebGLRenderer();
renderer.setPixelRatio(window.devicePixelRatio );
renderer.setSize(window.innerWidth, window.innerHeight );
container.appendChild( renderer.domElement );
</script>
```

为了实现景区内多个景点之间的切换,可以采用水平排列图片的形式,点击图片跳转到

相应的景点。最终的网页中全景漫游的效果如图4所示。



图4 “杨岐山”景区入口网页全景漫游

4 结语

“云旅游”是一种低成本、高效率、传播方便的线上旅游方式,是数字化和信息化发展的一个分支。本文以湘赣边区域为主题设计了云旅游系统,使用ECharts技术和DataV地理数据实现了网页地理区域的可视化,介绍了全景图的制作方法,利用WebGL和Three.js组件实现了网页全景漫游,这些方法同样适用于其他类似的展示系统。在未来的工作中,可以考虑加入后台和数据库,实现更丰富的功能。

参考文献:

- [1] 赵皇根,李武,李湘鹏.科技文化融合背景下湘赣边红色文化资源的可持续开发研究[J].湖南理工学院学报(自然科学版),2021,34(3):45-50,63.
- [2] 张成文,马少斌,邢玉娟,等.基于VR全景技术的在线虚拟旅游平台的设计与实现[J].兰州文理学院学报(自然科学版),2020,34(4):62-65.
- [3] 刘芳,薛冰洁,林琳,等.基于全景技术的虚拟校园漫游系统研究[J].现代电子技术,2022,45(22):133-137.
- [4] 王好平,王超.基于数据可视化的海图销售实时监控系统设计及实现[J].航海,2021(2):55-57.
- [5] 邱望,梁莉菁,徐远恒.基于WebGL的校园全景漫游系统的设计与实现[J].萍乡学院学报,2020,37(3):57-62.
- [6] 权西瑞,王凯,王小飞,等.基于Three.js的全景漫游产品设计与实现[J].地理空间信息,2022,20(7):71-73,119.

(下转第120页)

文章编号: 1007-1423(2023)10-0113-04

DOI: 10.3969/j.issn.1007-1423.2023.10.020

软件测试管理系统的研究与应用

付春子^{1,2*}, 唐海涛^{1,2,3}, 徐进^{1,2}

(1. 中国电建集团华东勘测设计研究院有限公司数字工程研究院, 杭州 311122;

2. 浙江华东工程数字技术有限公司技术管理部, 杭州 311122;

3. 数字城市 CIM 技术浙江省工程研究中心, 杭州 311122)

摘要: 提出了一种基于软件测试管理体系的管理系统, 阐述了如何通过产品解决方案来实现工程企业数字化产品的测试管理, 及如何在企业内部应用运作, 旨在帮助产研团队, 将质量活动管理起来并同时提升协同效率。通过对测试工作的数字化管理, 实现了高质高效的软件生命周期管理。软件测试管理系统产品的大规模应用, 在工程企业中, 赋予了产品质量活动过程的标准化、数字化, 具有重要的意义。在工程类产品的应用结果表明, 该系统大大提高了工作效率与质量管理能力, 具有广阔的应用前景。

关键词: 软件测试; 管理系统; 质量管理; 协同效率

0 引言

随着近年软件市场的快速发展, 人们对软件的依赖程度越来越高, 这里的软件, 不再局限于传统基于电脑的应用程序, 也包括 Web 系统、APP 客户端、随身设备等产品。而其中软件的质量, 对于产品的发展具有至关重要的作用, 因此每个企业都会成立专门的质量保障团队, 其中角色包括 QA (quality assurance)、QC (quality control)^[1], QA 的角色需要保证产品的交付质量满足需求; QC 角色则是需要建立一套规则体系, 并将这套规则体系实施在整个研发过程中。软件质量的主要保障手段就是软件测试, 一般由专业的测试工程师进行产品测试。测试工程师通过专业的方法论结合实际的项目, 开展一系列的质量活动。而随着产品的复杂度越来越高, 非常需要工具系统对测试活动进行管理。

1 软件测试的基本理论

1.1 软件测试的定义

在 IEEE 国际标准中, 对软件测试进行了详细的定义: “软件测试是在规定的条件下, 使用人工或自动化手段来运行或测试某个系统的过程, 其主要目的是对其是否满足设计要求进行评估”^[2]。

1.2 软件测试的发展

早期的软件开发过程中没有专门的测试阶段, 通常是开发人员自己去完成简单的调试工作。直至 1957 年, 软件测试这一环节的目的演变成发现缺陷。到了 80 年代, 软件测试又增加一部分工作, 对产品进行质量评估。随着技术的发展, 到了 90 年代, 开始出现各种各样的测试工具。到了 21 世纪, 随着互联网的快速发展, 测试技术取得了重大的进步, 从测试工具向测试平台、智能化测试方向演进, 让测试技术又

收稿日期: 2023-02-03 修稿日期: 2023-03-01

作者简介: *通信作者: 付春子(1988—), 女, 黑龙江虎林人, 本科生, 工程师, 研究方向为质量管理、自动化测试, E-mail: fu_cz@hdec.com; 唐海涛(1983—), 男, 江苏连云港人, 硕士研究生, 高级工程师, 研究方向为工程数字化; 徐进(1987—), 男, 河南固始人, 本科生, 助工, 研究方向为软件测试

往前迈了一大步^[3]。

1.3 软件测试模型

软件测试模型中应用比较广泛的有V模型、W模型、H模型。V模型(V-model)是一种用图像表示系统发展生命周期的模式。与V模型不同的是,W模型增加了软件开发阶段中应该同步进行的验证以及确认活动。H模型将软件测试的过程活动定义为独立的流程,并且参与产品各个生命周期^[4]。

1.4 软件测试管理

软件测试管理,实际是对整个软件研发过程中属于测试阶段所有活动以及结果的管理,可以细分为:测试对象管理(即需求管理)、测试计划制定、测试过程监控、缺陷管理、测试度量、测试过程产出物管理以及测试结果的管理等。

其中测试管理的重点是测试用例的管理、缺陷的管理、测试过程的管理、以及团队间信息的流转互通。通过系统的方式将以上内容高效有序地进行整合,可以为软件的研发降低风险并带来极大的便利。

2 测试管理系统的现状分析

当前,市面上的测试管理系统非常多,国内外一些较成熟的产品有:TestDirector、PingCode-Testhub、Zephyr for Jira、禅道、云效Testhub。

TestDirector是Mercury Interactive公司生产的面向企业的测试管理工具。系统中包括了测试管理的各个环节,但是属于商业产品,企业使用需要投入很大的资金成本,只适用于预算充足的企业使用。

Zephyr for Jira是Atlassian公司出品的项目与事务跟踪工具,主要功能包括需求管理、缺陷跟踪、 workflow管理、任务流转等。Jira对本土化的支持不是很友好,行为习惯和国内的用户有一些差距。同时按用户数量收费,且费用较高,仍适用于预算充足的企业使用。PingCode-Testhub是一家国内领先的企业,协作SaaS服务提供商研发的产品,是对标Jira的国内商业化产品,依然有使用成本,且该产品不具备自动化

测试能力,不能满足对自动化体系建设有诉求的团队^[5]。

3 测试管理系统的设计与应用

基于上述测试管理系统的现状分析,并结合企业的实际情况,市面上的产品或多或少都有不满足工程企业的测试管理的需求的地方。工程类产品主要包括大量交付型项目,以及部分基线产品,并且有很多大型项目交付周期很长,过程版本多。所以需要有完整的项目、版本、需求、用例、执行过程、缺陷、测试报告的质量活动闭环管理,需要有接口测试以及接口自动化的回归能力,需要有多种角色权限的管理能力。因此QA中心的成员们在测试基础理论之上,结合实际需求,设计了一套满足以上需求的软件测试管理系统。

3.1 产品架构

该系统主要包括:项目管理、需求管理、测试用例以及缺陷管理、测试过程管理等模块。项目管理是软件管理系统的基础,项目可以代表交付项目、基线产品等。后续的所有内容管理都是建立在项目的基础之上,因此在设计项目管理的基本要素时包括项目基本信息、模块信息、版本信息、项目成员和项目时间轴。版本信息,用来做大型项目或者产品的版本管理,可以对长期的交付过程进行版本拆分管理。需求管理主要用来创建需求,做需求变更的管理以及历史记录查看。

用例管理模块包括用例池管理和用例评审两部分功能。用例池功能主要管理项目中的用例,维护用例信息。用例评审功能主要记录项目团队中各个成员提交的评审记录,便于其他成员查看。测试过程管理各个版本用例的执行过程和进度,并可以自动生成测试报告。整个过程均可可视化,使管理效率提升。系统内预置缺陷的标准模版,进行缺陷的创建、任务流转、缺陷的风险标记等管理,以便于监控缺陷的整个生命周期,达到风险管控的目标。

系统将研发测试流程中的卡点进行了管理,以减少线下的管理沟通环节,包括版本开发提测以及版本测试结果自动评估的能力。版本开发提测主要用于待提测的版本关联需求或者修

复后的缺陷提测。版本测试结果的自动评估则是根据规则配置，系统自动判断版本是否符合质量要求。系统中设计了接口测试、接口自动化、定时巡检等自动化能力，可以大幅提升系统的测试回归时长，以及对系统质量作定时巡检以保证服务可用。测试管理系统的业务工作流程如图1所示。

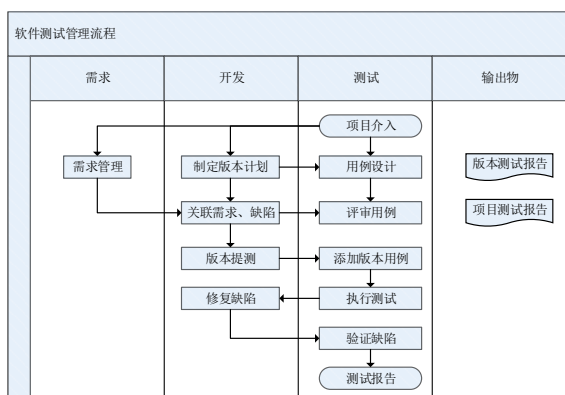


图1 软件测试管理流程

3.2 技术架构

该系统技术架构设计主要划分为5层，分别是IaaS层、数据层、服务层、网关层以及应用层。IaaS层为系统动态提供各类资源。根据软件质量管理业务应用的需求，数据层负责数

据存储，包括软件质量管理元数据、业务数据等，具体包括分布式缓存(redis)^[6]内存数据库以及Minio对象存储(object storage service)^[7]用于存储非结构化数据。软件质量管理体系基于云平台+微服务架构，实现软件质量管理应用服务一体化管理、运维以及弹性伸缩，提供统一平台化支撑服务，包括统一身份认证服务、访问控制、统一配置系统、API网关等。服务层实现软件质量管理体系所需的基础服务支撑。API网关则将客户端的请求分发至对应服务。软件质量管理体系支持多Web终端，比如市场占有率较高的IE、Chrome以及Firefox等浏览器，并提供了标准的RESTful API。

3.3 系统应用

公司设计的这套测试管理系统，是针对我院工程交付项目多、项目周期长、质量风险高、成本高等因素提出的建设方向。该系统定位于产品质量活动过程的标准化、数字化，已经在大型工程数据平台、图形引擎、水利工程系统等多个产品线中应用，得到了不错的应用效果和正向的用户反馈，极大地提升了产研的工作效率，也助推了质量管理水平的发展。如图2所示，是公司在做产品以及项目交付时使用的质量管理体系界面。

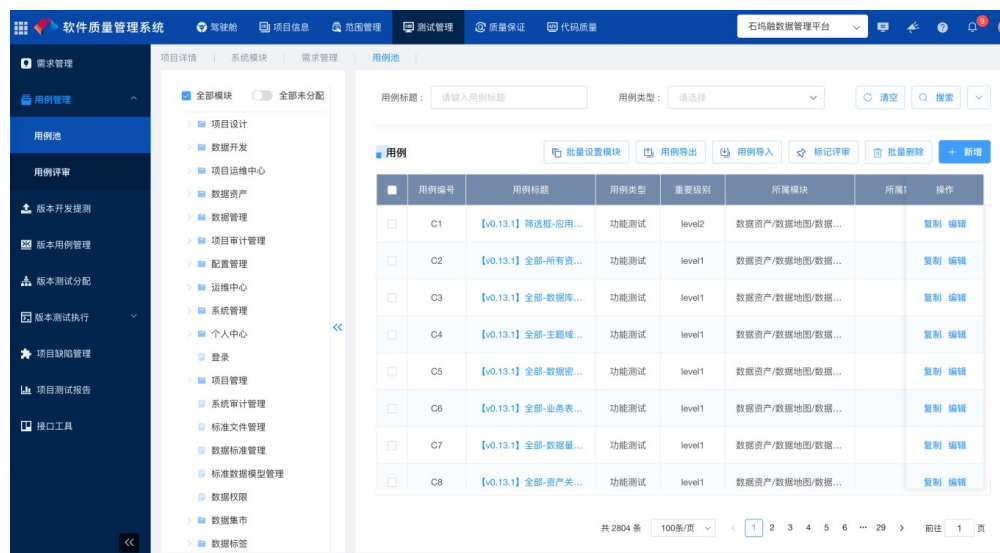


图2 用例管理

4 结语

本项目对于工程企业数字化产品的测试管理带来了重要意义,将线下的产研过程全部数字化、标准化地管理起来,对于不同角色都有清晰的视角可以管理项目。但是也存在一些不足之处,自动化能力暂时提供了服务端的自动化工具,还可以建设UI自动化的能力,作为分层质量管理的手段。相信随着质量手段的多样化,该系统不断地迭代完善,一定会为工程行业的质量管理提供越来越多的便利。

参考文献:

- [1] 贺焯. 软件企业质量保证的基石:QA、QC的良性协作[J]. 程序员,2005(6):59.
- [2] 叶姗姗. 浅析软件测试技术与测试管理[J]. 电子技术与软件工程,2013(16):83.
- [3] 易敏捷. 软件测试国内外发展现状及趋势研究[J]. 电脑知识与技术,2013,9(26):6020-6022.
- [4] 赖祥伟,张为群,邱玉辉,等. 形式化的软件测试模型研究[J]. 计算机科学,2004(11):169-173.
- [5] 陈国志. 软件测试管理系统的研究与实现[D]. 武汉:华中科技大学,2005.
- [6] 曾超宇,李金香. Redis在高速缓存系统中的应用[J]. 微型机与应用,2013,32(12):11-13.
- [7] 李兵,王连忠,司运成. MinIO存储在监控系统中的应用设计[J]. 工业控制计算机,2020,33(9):79-80, 82.

Research and application of software test management system

Fu Chunzi^{1,2*}, Tang Haitao^{1,2,3}, Xu Jin^{1,2}

- (1. The Digital Engineering Research Institute, Powerchina Huadong Engineering Co., Ltd., Hangzhou 311122, China;
2. Department of Technical Management, Zhejiang Huadong Engineering Digital Technology Co., Ltd., Hangzhou 311122, China;
3. Zhejiang Digital City CIM Technology Research Center, Hangzhou 311122, China)

Abstract: This paper presents a management system based on software test management system. It describes how to implement the test management of digital products in engineering enterprises through product solutions and how to apply and operate them within enterprises. It aims to help production and research teams manage quality activities and improve collaboration efficiency at the same time. Through the digital management of test work, high-quality and efficient software life cycle management is realized. The large-scale application of software test management system products is of great significance to the standardization and digitization of the product quality activity process in engineering enterprises. The application results in engineering products show that the system greatly improves the work efficiency and quality management ability, and has a broad application prospects.

Keywords: software testing; management system; quality management; synergy efficiency

文章编号: 1007-1423(2023)10-0117-04

DOI: 10.3969/j.issn.1007-1423.2023.10.021

无线充电智能小车设计

徐琛浩, 胡艺轩, 丁纵辽, 周飞雪, 张 岚, 赵春锋*

(上海工程技术大学工程训练中心, 上海 201620)

摘要: 根据全国大学生智能汽车竞赛无线充电要求, 限定 50 W 传输功率的无线充电短时间内为智能小车提供能量, 电能的充电与利用效率成为研究热点。以英飞凌单片机 TC264 为控制核心, 配合 LCC 型无线电能传输(WPT), 实现对电能的高效接收与利用。

关键词: 无线充电小车; LCC 拓部结构; 单片机

0 引言

无线充电技术已经越来越成熟, 将其与超级电容组一起运用在小车上, 摆脱了传统小车供电系统, 其结构灵活多变。类似的机器人有运用在餐饮行业的服务机器人和商场中的指引机器人。因此, 无线充电智能小车对无线充电发射端发出的能量的高效利用是我们主要研究的一点。

本文对三轮小车采用 2 F、12 V 超级电容组、32 位双核单片机 TC264 作为主控 MCU, 以此为核心搭建基本外围电路, 在此基础上加入超级电容组电能监控电路、电机驱动电路、信号放大电路。无线充电接收采用 LCC 恒流拓补结构。车身重量控制在 700 g 左右, 通过小车的 3D 结构搭建, 将重心尽可能压低, 处于电动机之上。自动控制方面对小车进行分段 PID 控制, 除去传统视觉寻迹外还加入了视觉路径优化, 达到处理速度较快的同时, 电能利用最大化。图 1 为无线充电智能小车实物图, 图 2 为小车赛道运行图。

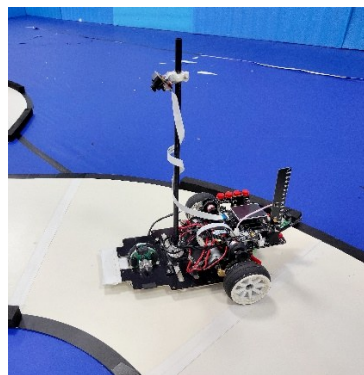


图 1 无线充电智能小车实物

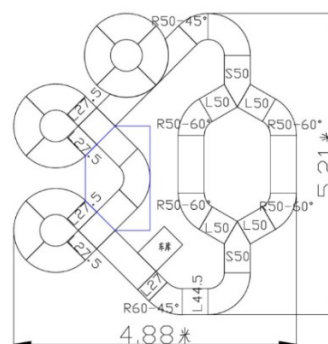


图 2 小车赛道

收稿日期: 2023-01-10 修稿日期: 2023-02-14

基金项目: 上海工程技术大学上海市大学生创新训练项目(CS2224001)

作者简介: 徐琛浩(2002—),男,上海青浦人,在读本科,主要研究方向为自动化;胡艺轩(2002—),女,陕西宝鸡人,在读本科,主要研究方向为自动化;丁纵辽(2001—),女,河南焦作人,在读本科,主要研究方向为机械电子工程;周飞雪(2003—),女,四川成都人,在读本科,主要研究方向为交通管理;张岚(2002—),女,甘肃陇南人,在读本科,主要研究方向为电子封装技术;*通信作者:赵春锋(1980—),男,江苏通州人,副教授,工学硕士,研究方向为智能监测与控制,E-mail: zhaochunfeng2008@126.com

1 小车硬件电路搭建

该无线充电小车以TC264为核心，搭建出基本核心电路，外接ICM20602姿态传感器，512线光电方向编码器、MT9V03X系列光电传感器，2.0寸IPS显示屏、HIP4082电机驱动器。系统框架如图3所示。

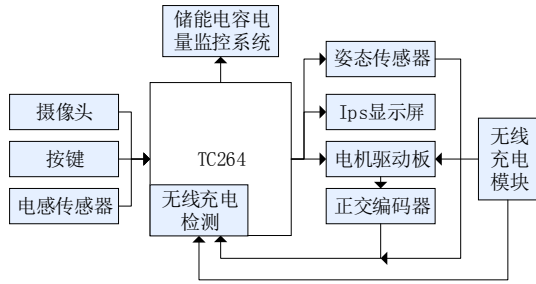


图3 系统框架

电能传输选用双边LCC型WPT(见图4)，由4个电容、2个谐振补偿电感和两个线圈组成。接收电路采用LCC副边拓补结构(见图5)，拓补部分由 C_{ss} 隔直电容、 C_{sp} 副边补偿电容和 $L2$ 接收线圈组成，其在负载波动的情况下仍然能够使得接收电流恒定，从而提高系统的稳定性，通过高频交流电整流滤波后输出直流电给超级电容组充电。结合文献[1]，有如下方程：

$$\begin{cases} Z1 = R1 + j\omega L1 + \frac{(j^{-1}\omega^{-1}C2^{-1} + j\omega L2)j^{-1}\omega^{-1}C1^{-1}}{j\omega L2 + j^{-1}\omega^{-1}C2^{-1} + j^{-1}\omega^{-1}C1^{-1}} \\ Z2 = R2 + j\omega L3 + \frac{j^{-1}\omega^{-1}C3^{-1}(RL + j\omega L4)j^{-1}\omega^{-1}C1^{-1}}{j\omega L4 + RL + j^{-1}\omega^{-1}C4^{-1}} \\ \delta = \frac{\omega M}{\sqrt{RxRy}} \end{cases}$$

其中： ω 为谐振角频率， j 为虚数单位， $Z1$ 为原边等效阻抗， $Z2$ 为副边等效阻抗， δ 为耦合因数。利用双边LCC补偿网络的特性，通过调节谐振角频率和互感系数，匹配原副边阻抗，使系统达到临界耦合状态，可提升谐振网络传输的精确性、稳定性及传输效率。其中互感系数在距离变化较小时为稳定值，根据原边回路等效电阻和副边回路等效电阻匹配双边LCC谐振下的等效阻值，使耦合因数 $\delta=1$ 即可实现最大功率传输。

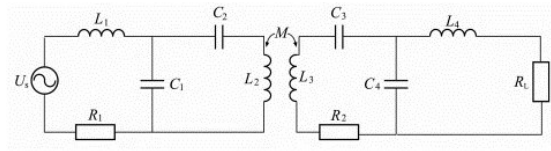


图4 LCC型WPT

根据 $C_{sp} = 1/(\omega X_s)$ 、 $L2 = X_s/\omega$ 、 $C_{ss} = 1/(\omega(\omega Ls - X_s))$ 得出在发射线圈150 kHz频率、 $X_s = 2 \Omega$ 下，副边LCC接收的理论参数。对参数做轻微调整后，最后成功将电流恒定在5 A左右，3.8 s左右将2 F、12 V电容组充满。整流部分选用低功耗型整流二极管，在5 A的情况下，单个二极管损耗功率在1 W左右，压降为0.55 V左右。为消除在未充电情况下前级电路可能对超级电容能量损耗的可能性，加入放倒灌二极管。由文献[2]可知，该系统能够在一定谐振频率下进行与负载无关的恒流输出，然而在实际运用中没有达到电压电流零相位的条件，因此随电能转换时间推移，电流输出能力有下降趋势，同时对于发射端发射50 W的情况，理想最大接收功率为45 W，然而实际接收功率为37.9 W左右，最后电能接收效率大约在76%左右。

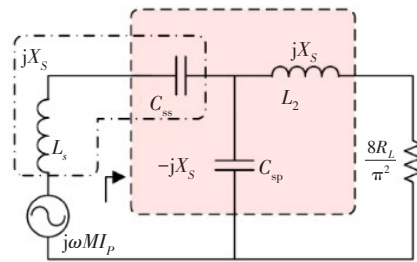


图5 LCC副边拓补结构

电能监控方面，我们选用TLD2331恒流输出芯片，外围电路中放置15颗LED灯，每一颗代表1 V，便于观察小车行驶路径中高功耗的地方。

电机供电方案选用直接连接超级电容组两端。超级电容组的电压处于2 V以下时，小车电机将无法工作，通过 $U=1/2 * C * Uf^2$ 可以得出，能够驱动电机的能量一共是140 J。该超级电容组最多能提供144 J能量，因此没有必要选用boost-buck电源芯片对电机供电。电机的H桥电路中，选用TPH1R403作为4个开关管，其等效

Rds 仅仅 2 mΩ 左右，在电机负载运行中损耗功率较低，为 0.004 W 左右。

2 系统算法

2.1 控制运行

每一款电机都有一个高效转速。因此我们在驱动电机时给出的脉冲调制(pulse width modulation, PWM)不能过高或过低，不能过高或过低，我们在程序上将 PWM 输出最高限制在 70%，当小车平均运行 PWM 在 35%~45% 区间内时，相同路径能够节约更多时间，效率可观。

小车控制算法选用 PID 控制，公式如下所示：

$$u(t) = k_p \left[e(t) + \frac{1}{T_i} \int_0^t e(t) dt + T_d \frac{de(t)}{dt} \right]$$

其结构简单，易上手，对于线性系统效果显著。小车行进中我们选择在直道增量式 PID 电机闭环控制的基础上，加入弯道的误差进行计算位置式 PID。这种串级 PID 能够提高控制质量，相对于并行 PID，能够更有效地抑制不同 PID 之间带来的冲突，以保证在操作条件和负荷发生变化的情况下，拥有更好的控制效果。在电机驱动中，主要体现在 PWM 的脉宽调制。脉宽的变化冲突越少，也就意味着 H 桥的开闭更加缓和，对超级电容组的电能管理也就更有效率。以下依次为增量式与位置式 PID 电机闭环控制实现代码：

```
static float BiasO, PwmO, Last_biasO;
int Incrementalone_PI (float Encoder, float Target)
{
    IncrementPID_Transition();
    BiasO=Target - Encoder;           //计算偏差
    PwmO+=Velocity_KP*(BiasO-Last_biasO)+
        Velocity_KI*BiasO;           //增量式 PI
    PwmO = PwmO >7000? 7000 : PwmO;
    PwmO = PwmO <(-9000)? (-9000) : PwmO;
    Last_biasO=BiasO;                 //保存上一次偏差
    return PwmO;                       //增量输出
}

static float Bias, Pwm, Last_Bias;
float Position_PID(float error)
{
    Bias=error;                         //计算偏差
```

```
Pwm=Position_KP*Bias+Position_KD*(Bias-
    Last_Bias);                         //位置式 PID 控制器
    Last_Bias=Bias;                       //保存上一次偏差
    return Pwm;                             //输出
}
```

但在不同超级电容组电压情况下，电机驱动的扭矩不同，因此针对不同电压情况进分段 PID，这样能够让传统 PID 更加灵活，避免一套 PID 带来的不稳定性，让电机能够在低电压的情况下有相同的扭矩和响应速度。当超级电容组处于 8 V 以下时，电机的扭矩明显不足。部分代码如下：

```
static void PHI_Position_PID_live(void)
{
    else if(Pre_Speed >= 90 && Pre_Speed < 100)
    {
        Position_KP = 0.53;
        Position_KI = 0;
        Position_KD = 0.5;
    }
    else if(Pre_Speed >= 100 && Pre_Speed < 110)
    {
        Position_KP = 0.59;
        Position_KI = 0;
        Position_KD = 0.7;
    }
}
```

在寻迹传感器上我们选用光电式传感器，通过对图像的二值化，得到赛道两边边线，进而得出中线，完成基本寻迹。在弯道的情况下，转弯半径如果过大，会导致路径更长，并且小车采用差速过弯，这样会让电能消耗更快。因此，我们将小车控制在内切的路径上，对于一些圆环、十字回环类元素，也是同理，尽可能挑选出转弯半径最小的路径。经过测试，在相同电能情况下，小车能够经过更多的元素。

2.2 信号检测

无线充电原边线圈与副边线圈在不同位置下的接收电压、互感系数大不相同。我们在遇到发射线圈时需要将小车停在合适的位置来接收最多的电能。这里采取电感电容并联谐振来检测 150 kHz 信号。电感外切发射线圈的瞬间，接收电压会迅速拉低，我们在检测到最低电压

后加入电机编码器计步实现精准停车,在调节过程中,只需调整最低电压检测阈值即可。

3 结语

在实际运用中,我们在短期将电能接收、利用效率达到最大化,使智能小车能够在赛道上行进大约 24 s、40 m 左右,时速约 1.7 m/s。如果需要进一步改善利用效率,需要对接收端 LCC 参数做细致调整,并且需要换成其他高效

率的电机,本文选用的 RS380 直流有刷电机的消耗功率相较于无刷电机更多。

参考文献:

- [1] 魏小钊. LCC-LCC 结构电动汽车无线充电若干问题的研究[D]. 济南:山东大学,2021.
- [2] 程福临,刘春喜,邵翔宇,等. 一种对称式双 LCC 补偿无线电能传输系统参数设计方法[J]. 电工电能新技术,2021,40(12):64-72.

The design of wireless charging intelligent car

Xu Chenhao, Hu Yixuan, Ding Zongliao, Zhou Feixue, Zhang Lan, Zhao Chunfeng*

(Engineering Training Center, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: According to the wireless charging requirements of the National College Students' Smart Car Competition, wireless charging, which is limited to 50 W transmission power, can provide limited energy for smart cars in a short period of time, and the utilization efficiency of electric energy has become a research hotspot. In this paper, Infineon MCU TC264 as the control core, with LCC type Wireless Power Transmission(WPT), efficient reception and utilization of electric energy was realized.

Keywords: wireless charging intelligent car; LCC WPT; MCU

(上接第 112 页)

Development of “Xiang-Gan Bian” online panoramic roaming system based on Web

Qiu Wang^{1, 2*}, Xiong Yu¹, Liang Lijing¹

(1.School of Information and Computer Engineering, Pingxiang University, Pingxiang 337000, China;
2. Jiangxi Huimeng Network Information Technology Co., Ltd., Pingxiang 337000, China)

Abstract: The “Cloud Tourism” and the panoramic tour bring a good online virtual tour experience to the users. In order to promote people's understanding of the “Xianganbian” region, an online panoramic tour system was designed and implemented. The web map function was implemented with the help of DataV data visualization tool and ECharts component. PTGui software was used to synthesize the panoramic images. Web page panoramic roaming and interactive functions were implemented using WebGL technology and Three.js component.

Keywords: online tourism; panoramic tourist; software development; virtual reality