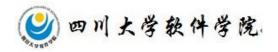


四川大學

生物医学分布式数据计算与管理系统 开发报告

题	目_	《生物医学分布式数据计算与管理系统》			
学	院 _	建筑与环境学院			
专	业_	工程力学-软件工程交叉实验班			
学生好	姓名。	白马腾			
学	号_	209141470416	年级 _2019 级		
指导	教师	宋万	忠		

二0二一年 六 月 十八 日



目录

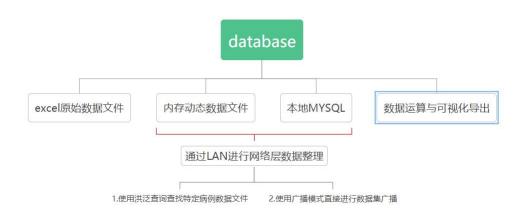
一 、	功能模块与说明	1
二、	实现功能的流程及说明	2
三、	数据库设计与说明	6
四、	项目开发使用的知识点列表	11
五、	项目开发中遇到的问题和体会————————————————————————————————————	13
六、	附件一: 代码	附1
七、	附件二: 视频 (PPT)	附2



项目功能模块简介

该系统是为了高效进行数据信息分析,可视化运算,逻辑数据组织管理而开发的分布式体系。 系统有三个主要的模块构成:

- 1.基于逻辑数据存储的信息运算模块
- 2.基于局域网单项加密下的签名数据共享与搜索模块
- 3.基于 MYSQL 的数据库管理模块



信息运算模块:

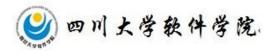
其中逻辑数据存储模块使用逻辑类方法对原始 database 数据进行自动归类整理,构建该模块的使用类变量声明如下:

使用病例类 patient class 逻辑组织病例数据

类名: 方法含义: 变量名: 变量含义: 类内方法: 病人性别 增添新的数据时期 patient add period() sex 病人姓名 show() 显示病例状态 name ID 病人 ID 存储 show F() 显示计算各节点F值 period list[] 病人时期数据集 show K() 显示计算各节点 K 值 d see point list[] 病人血管观测点间距储存数组 calculation 病例可运算标识 period length 时期数组遍历长度

使用时期类 period class 逻辑组织病例数据

类名: 变量名: 变量含义: 类内方法: 方法含义:



period	data top12 stent_list[] Fx[] Fy[] Fy[] Fz[] kx[] ky[]	定义支架时期 定义骨骼位置点 支架列表 反解各节点 x 方向上力大小 反解各节点 y 方向上力大小 反解各节点 z 方向上力大小 反解各节点 z 方向上力大小 反解各节点 F 值大小 结合空间位移,反解 kx 值大小 结合空间位移,反解 ky 值大小	add_stent () show_3D ()	增添支架函数 3D 显示模型函数
	kz[]	结合空间位移,反解 kz 值大小		

使用支架类 stent class 逻辑组织支架结构数据

类名: 变量名: 类内方法: 方法含义:

stent stent_type 标识支架类型(近端,远端) add_huan() 增加支架环个数

stent_shape 标识支架结构形状 huan list[] 存放支架环数据

使用环类 huan class 逻辑组织支架环数据

类名:变量含义:类内方法:方法含义:huanpoint_list[]存放峰值点数据initialization(支架中点值初始

point_mid 存放支架中心点数据

number points 支架环峰值个数(1,5,8)

使用点类 point class 逻辑组织底层点数据

类名: 变量名: 变量含义: 类内方法: 方法含义: point x 点 x 坐标 无

y 点 y 坐标 z 点 z 坐标

形态学特征量分析:

完成逻辑组织数据封装后,可建立不同病人病人不同时期统一运算方法,对观察的特征量进行运算。原有的统计数据方法受限于数据量少无法描述支架与人体中的空间分布状态,因此只能选取少量特征点进行 X, Y, Z 方向上的位移量拆解,无法对支架 z 方向对血管的硬化约束,支架本身的弹性回直效应,支架环向偏转,支架径向变化量,支架环向压缩情况分析。

为探究 1. 支架 xy, yz, xz 三平面内回直效应参数

- 2. 支架各环轴向偏移量大小,限制型支架交叠部分滑移参数
- 3. 支架各环直径,形状变化规律
- 4. 支架环向偏转角方向大小

本文研究中将支架空间描述拆分为四个基本描述量

1. 支架环各时期间轴向偏转角 【描述统计误差】

2. 支架环延轴位移量 【描述支架滑动,相对滑动情况】

3. 支架环向偏转角 【描述支架旋转情况】

4. 支架径向变化量 【描述支架直径变化,径向回弹情况】



研究的标志性参考量:

1. 支架中心点绝对空间位移

2. 支架环节中心点间距变化情况

3. 支架环节间角度变化情况

4. 支架环回直分布情况

【构建预测模型的重要参考量】

【描述支架长度变化原因】

【描述弹性回直效应】

【描述回直效应的产生机理与特征】

用以与传统方法、mimics 计算统计方法做对比的参考量:

1. 支架末端 z 值位移

2. 支架环截面形状提取

3. 支架环直径分析

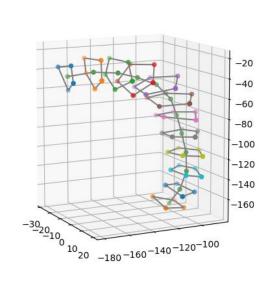
4. 支架弯曲度分析

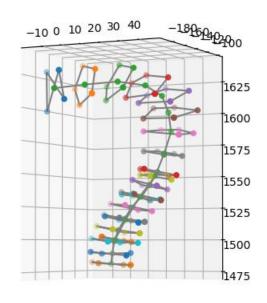
【与古典法对比】

【与mimics 自带截面提取工具对比】

【与 mimics 三种拟合圆方法对比】

【与mimics 支架弯曲度计算对比】





图一

图二

图一: 调用 show_3D()方法展示限制型支架观察点结构模型

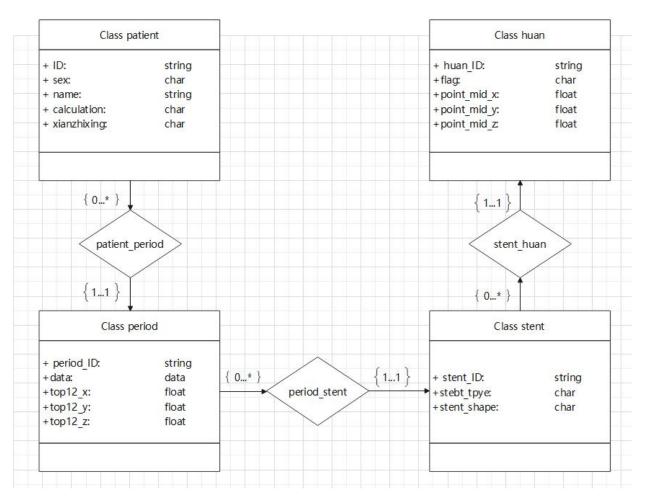
图二:调用 show_3D()方法展示非限制型支架观察点结构模型

数据完备性思考:

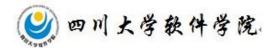
形态学观测的目的是描述,使用各种不同的方法得到能够客观反应事物变化规律的模型。古典方法着重于"特征",使用能够直观观测到的特征量或由某种算法将直观特征量组合而成的新特征量来描述模型或阐明规律。然而人体是一种极端复杂的非线性系统,很难逐一计算统计各个能观测到的特征量的统计学指标并对其生理意义进行逐步分析。采用原有的:提取-检验-分析-再提取需要付出庞大的时间与人力成本。

造成这种结果的原因我认为是对数据"基本性"与"原子性"的忽视。假设将基础的,不可由其他数据表示的,独立的数据集合命名为原子数据集 $\Phi = \{\alpha 1, \alpha 2, \alpha 3...\}$,将通过广义函数 $f(\Phi)$ 构成的新数据集称为 $\Phi = \{f1(\Phi), f2(\Phi), f3(\Phi)...\}$,将由广义函数本身定义的集合称为 $F = \{f1, f2, f3...\}$ 如果能够确定 Φ 集合的范围,并确定集合规模,那么使用计算机对 Φ 集合与 Φ 集合进行存储统计必然优于对 Φ 集合进行的逐一统计。

因此,建立原子数据构成的数据集o并开发广义函数算法集合F,是设计这次项目的基本目的。



SQL 数据口架构 E-R 图



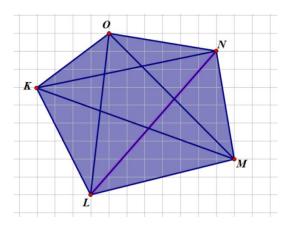
代码模块小览:

模块代码截图:

图三:该模块对数据结构进行定义,将 excel 原始统计数据录入至内存中

图四: 进行主函数选择调用过程中的选择语句

支架径向变化量



图五: 支架截面直径运算示意图

支架的径向变化量使用支架环空间五点(K, L, M, N, 0)中距离最远的两点距离作为参考。对比支架环在随时间变化过程中该参考量的变化趋势与范围。

支架轴向位移

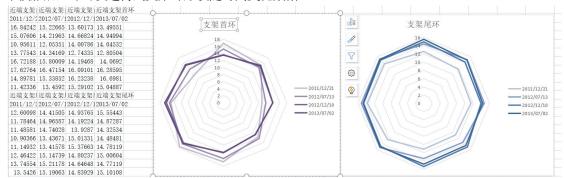
古典方法中对支架滑移使用 z 值直接相减的方法,该方法计算位于降主动脉部分时由于方向适量与 z 向夹角 sata 极小,因此具有较高的精度。但若病患血管方向矢量与 Z 呈较大夹角时(sata>30 度),理论上使统计值和实际滑移相差比例>15.2%,将会严重影响统计结果。

由于在 1.4a 部分中计算除了各环的方向矢量,因此利用方向矢量 I 点乘归一化的支架环中心位移,即可近似算出支架环轴向位移量。

支架环形状投影分析

单一的使用最大线性直径或使用圆拟合方法不足以准确的描述支架环形状变化规律。且单变量描述会损失支架环的形状参数,例如无法描述其由圆变为等面积椭圆的受压过程,因此,找到支架环形的快速投影输出方法并进行二维的平面分析是必要的。

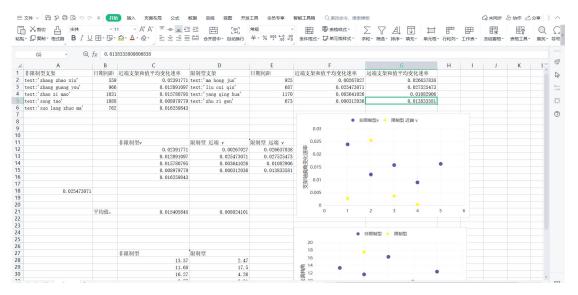
由与支架环首尾受压现象并不明显,因此可近似认为支架环中心与各环峰的夹角等距,计算矢量半径并进行二维投影后,即可快速得到支架环形状随时间变化数据。



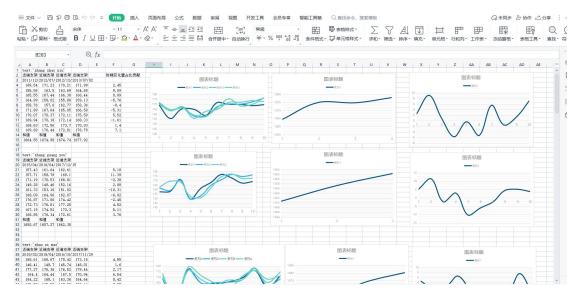
图六: 支架环状投影数据分析

支架回直夹角与分布分析

支架回直角采用图十三中矢量夹角进行计算,导出数据利用 python 直接输出为数据报表进行展示。使用 period 层级的 flag 标记位对报表进行信息标识。



图七: 支架回直数据导出



图八: 支架回直分布状态数据导出



小结:

逻辑数据层的构建能够极大程度上帮助进行信息的快速统计,能够有效的解决样本数据集过大等问题。使用逻辑层数据进行数据分级,建立的本地的 SQL 数据库信息,使用统一管理语句实现了对数据库的高效查询与检索。包括 ini db(), unite db()等批量处理结构方法。

内存数据结构上,通过 python 对类级分层数据的控制,实现的深度优先检索与本地数据计算,统计信息绘制等方法。

基于 MYSQL 的数据库管理上,通过 MYSQL 的表结构数据分层,实现了数据库统一查询实现与快速更新数据方法。

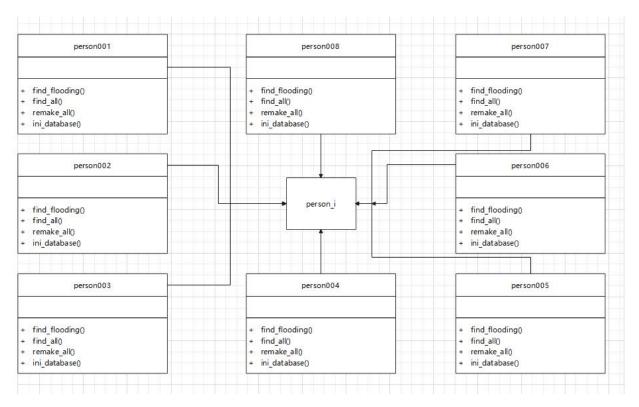
基于 excel 原数据库结构上,使用 excel 进行快速更迭,实现了高效的原始信息录入方法。 通过这三种不同的数据管理结构,实现了本地数据的完备性更新,但并不能完整的进行分布 式数据运算、不同主机间的数据快速更新、数据节点过于庞大时的洪泛信息查询等必要信息统计 整理过程。



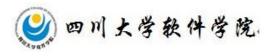
局域网共享与搜索模块:

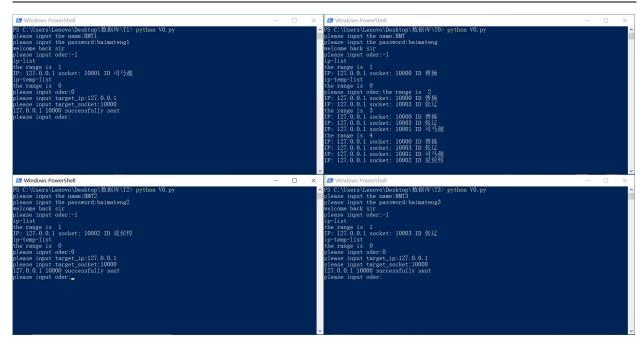
该模块的开发为了解决课题组内分布式数据无法快速共享,统计,深度搜索等问题。构建了两种数据统计方法: 1. 基于 C/S 结构的中心式广播模式。2. 基于迭代法的洪泛查找模式。

1. 中心式广播模式:

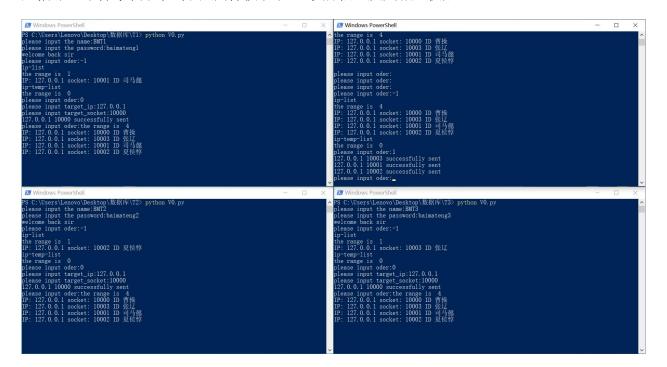


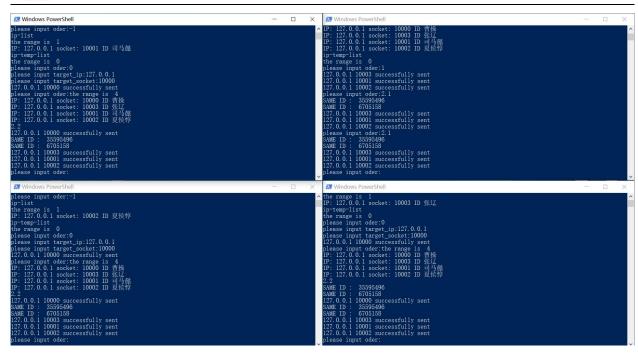
中心广播所使用的方式是基于 C/S 模型的中心查找+广播信息模式,使用的结构来源于 C/S 模型的思想,但为了保证局部网络的健壮性,在设计本项目时,使用"去中心化"理念:每台运行主机都会开启 Sever 与 Client 两个线程模型 (在 1.2 版本时将 Client 线程并入主线程保证 matlib 函数库的并发调用),当用户选择执行中心化链接方法时,可以选取任意一台电脑作为本次链接的中心机,或选取多台电脑同时进行作为核心。





作为中心机的电脑 Sever 线程启动并进行信息搜集,制作本地 IP 列表。输入"-1"查看本地 IP,ini 信息。在获取全部信息后中心节点选择是否将本地的 IP 表格进行负反馈广播,选择反馈广播即意味着每个分布式节点同样获取了 IP 执行信息及其相应权限。



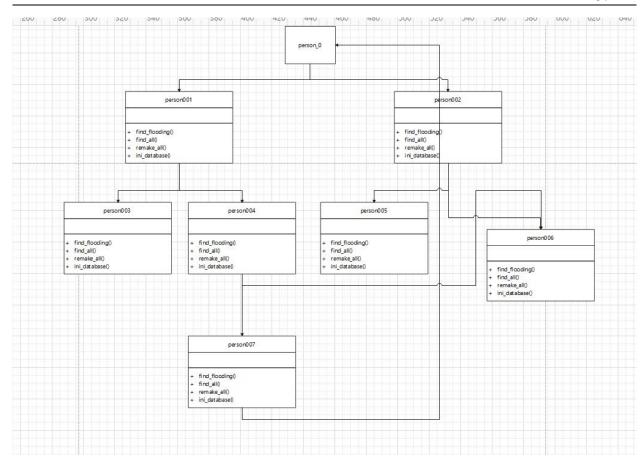


在进行中心化信息聚合之后,具有权限的电脑即可选择对低于自身权限的主机信息进行访问,该过程是一个基于顺序列表的并行 TCP 链接建立过程。建立通用 TCP 链接对课题组内所有database 进行整合后,即可选择进行本地信息运算。

2. 洪泛信息查找:

虽然基于 C/S 模式构建的网络使用多线程的方式部分解决了"去中心"的问题,但不可避免的: C/S 结构本身即说明了这种模式不可能做到绝对意义上的去中心化。同时,C/S 模式在大规模信息检索过程中同样暴漏了其顺序检索所带来的及高成本的时间复杂度消耗。

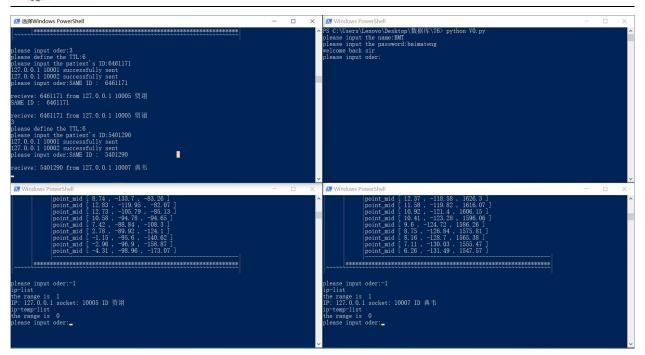
为了解决上述的两点问题,选择 flooding 模式是必要的。



洪泛模型的测试用例网络拓扑结构如上,在洪泛模式下,需要解决两个问题:

- 1. 洪泛模型下的信息传播在遇到环路重复问题
- 2. 洪泛模型下的信息连锁 DDL 问题

针对第二个问题,使用传递报文协议加入 DDL (默认设置为 5) 可以解决,解决问题一使用报文协议: 在报文的标记头尾部增添节点记录,每次节点在处理发送信息的的同时对尾部进行筛选,若发现目标节点已存在于信息节点中,则跳过本次发送。

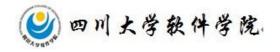


洪泛在 TO 处进行检索,使用 flooding 进行信息传递,对特定病例进行查找后,返回病例数据与病例数据名称。

总结:

如果说本学期的课程真正有趣在哪里,我认为在于其极强的适用性质:数据库的知识让我在建设系统时能使用 SQL 工程对复杂的数据结构进行定义与查询。提供了一组有效的信息结构化组织形式。计算机网络的知识让我实现了 LAN 下分散型数据库的快速整理与查找,并构建了 C/S,FLOODING 两种模式以适用在课题组的不同项目条件下。这是一件有意义的事情,它告诉我使用所学的知识能够极大的化简平时繁重的工作,一个合理建立且有序的体系同样能够方便我们将原本离散的数据高效的组织起来。

这次的大作业最终能够以项目的形式出现同样离不开各位学长学姐,宋老师和其他同学的帮助。代码的总工作量大约在1300-1400 行左右,实现了数据可视化,数据分布式管理,数据库录入与动态查询,支架数据算法实现等。



由于这次的大作业的用例数据集是来自于华西临床医学院的病例数据集,为了保障病患数据 安全性,我将用本机演示向大家演示项目的实操情况与流程,视频与项目源代码,配置环境详见 附件。不便之处请诸位谅解 QAQ。