



金融风控的资金流水分析

初赛赛题说明

发布日期 2020-01-17

目 录

1 题目说明.....3

1.1 输入信息3

1.2 输出信息4

1.3 限制条件4

1.4 接口说明4

1.5 排分规则5

1.6 其它说明5

2 运行环境.....6

1 题目说明

通过金融风控的资金流水分析，可有效识别循环转账，辅助公安挖掘洗钱组织，帮助银行预防信用卡诈骗。基于给定的资金流水，检测并输出指定约束条件的所有循环转账，结果准确，用时最短者胜。

- 1.1 输入信息
- 1.2 输出信息
- 1.3 限制条件
- 1.4 接口说明
- 1.5 排分规则
- 1.6 其它说明

1.1 输入信息

输入为包含资金流水的文本文件，每一行代表一次资金交易记录，包含本端账号 ID，对端账号 ID，转账金额，用逗号隔开。

- 本端账号 ID 和对端账号 ID 为一个 32 位的正整数
- 转账金额为一个 32 位的正整数
- 转账记录最多为 28 万条
- 每个账号平均转账记录数 < 10
- 账号 A 给账号 B 最多转账一次

举例如下，其中第一行[1,2,100]表示 ID 为 1 的账户给 ID 为 2 的账户转账 100 元：

1,2,100
1,3,100
2,4,90
3,4,50
4,1,95

2,5,95
5,4,90
4,6,30
6,7,29
7,4,28

1.2 输出信息

输出信息为一个文件，包含如下信息：

- 第一行输出：满足限制条件下的循环转账个数
- 第二行开始：输出所有满足限制条件的循环转账路径详情。

输出循环转账路径要按照指定排序策略进行排序：总体按照循环转账路径长度升序排序；同一级别的路径长度下循环转账账号 ID 序列，按照字典序（ID 转为无符号整数后）升序排序。

举例如下：

4
1,2,4
1,3,4
4,6,7
1,2,5,4

1.3 限制条件

循环转账的路径长度最小为 3（包含 3）最大为 7（包含 7），例如账户 A 给账户 B 转账，账户 B 给账户 A 转账，循环转账的路径长度为 2，不满足循环转账条件。

1.4 接口说明

- 选手需要在代码中读取输入文件/data/test_data.txt，分析后将结果输出到/projects/student/result.txt
- C 语言的编译命令为：gcc -O3 main.c -o test -lpthread
- C 语言的运行命令为：./test
- C++语言的编译命令为：g++ -O3 main.cpp -o test -lpthread
- C++语言的运行命令为：./test
- Python 语言的运行命令为：python3 ./Main.py
- JAVA 语言的编译命令为：javac Main.java

- JAVA 语言的运行命令为: `java Main`

1.5 排分规则

- 结果准确，用时最短者胜出。
- 如果时间一样，先提交的选手排名靠前。
- 选手成绩取个人多次提交里面的最好成绩。

1.6 其它说明

- 可选基于鲲鹏 920 的特点（如：多核，NEON，Cache 大小）进行加速。
- 只能使用 Python 标准库和 `numpy`，其中判题程序使用的是 `numpy 1.17.2` 版本。

`numpy 1.17.2` 在 Euler OS 上的安装方法为：

1. 安装 `python3-devel`

```
wget https://developer.huawei.com/ict/site-euleros/euleros/repo/yum/2.8/os/aarch64/updates/python3-devel-3.7.0-9.h15.eulerosv2r8.aarch64.rpm
```

```
rpm -ivh python3-devel-3.7.0-9.h15.eulerosv2r8.aarch64.rpm --nodeps
```

2. 修改 `pip` 的安装源

创建如下的文件 `~/pip/pip.conf`，并在文件中添加如下配置，保存并退出

```
[global]
```

```
index-url = https://repo.huaweicloud.com/repository/pypi/simple
```

```
trusted-host = repo.huaweicloud.com
```

```
timeout = 120
```

3. 安装 `numpy 1.17.2`

```
pip3 install numpy==1.17.2
```

- JAVA 代码请使用 UTF-8 编码。
- ID 和转账金额都为 32 位的正整数，且小于 2 的 31 次方，选手可以不用考虑范围之外的输入。
- 不能使用网络相关的接口调用，如 Python 里面的 `multiprocessing.Manager` 接口，否则会出现运行错误。
- 打开输入文件时，要使用只读方式打开文件，否则打开文件会失败。
- 通过 `open` 创建文件时需要指定文件权限，将 `open(fileName, O_RDWR | O_CREAT)` 改为 `open(fileName, O_RDWR | O_CREAT, 0666)`，否则会导致编译失败。

2 运行环境

- 选手使用的练习资源：2U4G
- 判题系统使用的判题资源：4U16G
- 操作系统：Euler OS
- 服务器：TaiShan 服务器
- 芯片：鲲鹏 920