

系统 Log 日志分析

数据提供了1000个随机抽取的样本用户，在4周内的电脑系统日志。behavior文件夹中是按日期归档的样本行为日志，例如，样本行为日志“0AB6BBBEDFF24EC8BAAC905F45AE314C_2012-05-07_21-22-38.txt”中，记录了ID为0AB6BBBEDFF24EC8BAAC905F45AE314C的样本在2012-05-07这天，21点22分38秒这次开机以来的行为日志。

在行为日志中，以“字段<=>值”的格式记录数据，若一行中有多个字段、值对，则以“[=]”分隔。例如，一个典型的样本行为日志如下：

```
Last<=>26037
L_Start<=>2012-08-06 08-49-16
...
T<=>8933[=]P<=>360chrome.exe[=]I<=>1224[=]U<=>item.taobao.com/item.htm[=]A<=>205e0[=]B<=>203b2[=]V<=>5.3.0.922
T<=>8935[=]P<=>explorer.exe[=]I<=>4052[=]W<=>10302[=]V<=>6.1.7600.16385
T<=>8952[=]P<=>Foxmail.exe[=]I<=>4608[=]W<=>404a0[=]V<=>7.0.1.91
T<=>8960[=]P<=>explorer.exe[=]I<=>4052[=]W<=>60c24[=]V<=>6.1.7600.16385
T<=>8986[=]P<=>iSeeQView.exe[=]I<=>8436[=]W<=>20d36[=]V<=>1.2[=]N<=> 看图精灵
[=]C<=>北京远图科技
T<=>8988[=]P<=>explorer.exe[=]I<=>4052[=]W<=>60c24[=]V<=>6.1.7600.16385
...
```

样本用户每次开机时，都会形成一个对应的日志文件，文件的开头2行会记录以下2个值：

Last表示日志的最后一次更新时间点距离开机时间的秒数；

L_Start表示样本用户的开机时间（年-月-日 时-分-秒）。

接下来，数据采集程序会以2秒一次的频率扫描样本用户计算机的当前焦点窗口，若焦点窗口发生变化，则会在日志中追加一条记录。其中：

T表示该记录的时间点，以距离开机时间的秒数形式记录，例如8933表示距离开机时间的8933秒后；

P表示焦点窗口的窗口进程名；

I表示焦点窗口的窗口进程号；

U表示浏览器窗口的地址栏内容，即当前浏览网页的URL（为了保护样本隐私，已将URL中第1个问号之后的参数部分截断）；

A表示浏览器窗口的地址栏句柄；

B表示浏览器窗口的当前标签页句柄；

V表示焦点窗口对应的程序版本号；

W表示非浏览器焦点窗口的窗口句柄；

N表示焦点窗口对应的程序名称（仅在第一次出现时记录）；

C表示焦点窗口对应的程序所属公司名（仅在第一次出现时记录）。

根据对数据文件的分析， 请回答以下问题：

- 1、 请问在所有这些数据中，哪 10 个程序（P）使用次数最多，它们的使用次数分别是多少？
- 2、 请问在所有这些数据中，排序输出每个浏览器程序分别访问了多少次网页。

问题 2 提示：根据浏览器窗口的地址栏内容（U），按照浏览器程序（P）进行分组排序。