

Result of getting user info on PTT

Speaker: Yin-Hong, Hsu

Aug 23, 2016



Outline

Previously

What I've done

The Flow Of Getting User Info

Data format

Somebody's data



Previously

[徵求] 5/26 11點半鹿安橋行車記錄器
<https://www.ptt.cc/bbs/Gossiping/M.1464435196.A.46B.html>
smiledolphin
barbarian72：分類啦幹
jetalpha：標題請加縣市地區，內容請加聯絡人跟聯絡方式，幫高調
XXXXGAY：幫你置底幾天
torunwind：先幫你高調
ArcueidY：4X好壞XD
smiledolphin：對不起啦，我不會用這個...我不是故意的！
lion9177408：幫高調
strike519：+ U
monkey1101：高調
iceyeman：先縮電視台數做出優質節目外銷其他國比較快啦幹！
nebbiabards：！！！花鳥風月！
coon182：老婆推錯地方了（拖回置底
aghgna：動把我困在家裡關起來後就不再理會我

我們的朋友施柏毅
5/26當天中午11:35分左右，在鹿港東石里鹽埕巷靠近彰濱新聯絡道路（鹿安橋）附近發生嚴重車禍，雖經緊急搶救，仍不幸辭世。
對方駕駛肇事逃逸，目前還找不到人
不知道在那時間點是否有經過這邊的駕駛朋友、有錄到相關可以影像的、是否可以麻煩提供行車記錄。
幫忙我們的朋友一起找出肇事逃逸者
謝謝大家.....

Figure: get post data by http

Previously

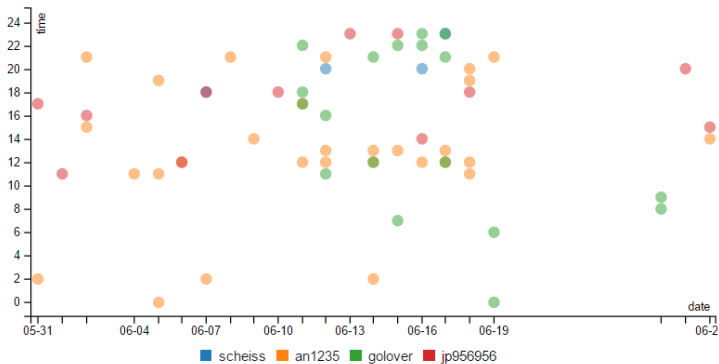


Figure: analyze poster's activity

What I've done

- ▶ Parse user info from userinfo page by telnet
- ▶ Record all data into database



Process

- ▶ Login
- ▶ Menu
- ▶ Talk
- ▶ Users
- ▶ Info
- ▶ Back & Down
- ▶ Info



P T T 小天使 招募新人中！

招生日期：8/26~9/07
報名看板：NewAngels

報名資格：

註冊滿一年文章滿300篇 或
註冊滿兩年文章滿150篇

退文數兩篇(含)以下

登入 300次以上
(日期均以2016/09/07為準)

歡迎來到批踢踢實業坊，現在有【52201】人按兩下找小天使問問題！

請輸入代號，或以 guest 參觀，或以 new 註冊：

2016 棒球圓夢計畫・球員募集中！》》》 <http://baseballgift.weebly.com>

Figure: Login

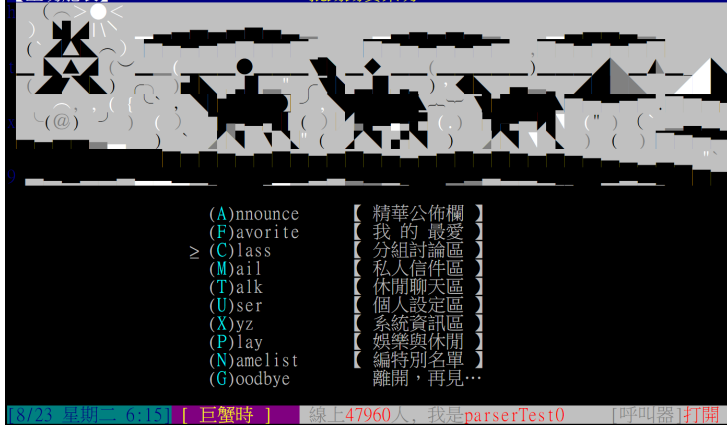


Figure: Menu



≥ (U)sers
 (Q)uery
 (T)alk
 (C)hat
 (P)ager
 (W)atch Pet
 (A)Change Angel
 (D)isplay

線上使用者列表

查詢網友

找人聊聊

【Ptt多人聊天室】

切換呼叫器

查詢寵物

更換小天使

顯示上幾次熱訊

[8/23 星期二 6:16] [巨蟹時] 線上48069人, 我是parserTest0 [呼叫器]打開

Figure: Talk

【休閒聊天】 批踢踢實業坊						
排序:[嗨! 朋友] 上站人數:48092 我的朋友:1 與我為友:1 板友:0 壞人:0						
編號	P	代號	暱稱	故鄉	動態	發呆
≥ 1		parserTest0	0000	東森媒體科技	使用者名單	
2	*	sanal	蒼天已死	*	閱讀文章	4h36
3		cherry000333	阿庭	中山大學	閱讀文章	0'34
4		cherry000333	阿庭	Giga超網路	閱讀文章	1'00
5		LSC	原來台北是貴州	36.224.52.61	主選單	0'19
6		shaznax	漂漂のIZAM	中華電信	閱讀文章	14h37
7	f	pest	這些分鐘妳有沒有	*	看板列表	0'18
8		rosephills	朱天爺	台灣寬頻通訊顧問	閱讀文章	3h51
9		singetta	婚攝九五	東森媒體科技	看板列表	1'55
10	W	y851503040	松本雨	*	看板列表	2'05
11		al234414	al234414	110.28.70.177	讀信	13'46
12	-	Jerway	哲瑋	*	閱讀文章	1'16
13		luvjp	這就是人參	42.72.41.149	看板列表	0'11
14	*	RealARSLAN	妖魔	*	閱讀文章	2h08
15		ASKA	The 羊男	中華電信	閱讀文章	5h32
16	*	wet	小魚	*	看板列表	32'57
17	-	orion1982	深藍	*	閱讀文章	1'37
18		cindy0617	oreo	223.139.245.124	閱讀文章	0'10
19		juin	成為暴風	66.171.123.202	閱讀文章	1'51
20	*	dadafantasy	鳴啦啦	*	主選單	2'44
(休閒聊天) (TAB/f) 排序/好友 (a/o) 交友 (q/w) 查詢/丟水球 (t/m) 聊天/寫信 (h) 說明						

Figure: Users

【休閒聊天】 批踢踢實業坊
《ID暱稱》 parserTest0 (0000) 《經濟狀況》 家徒四壁 (\$0)
《登入次數》 12 次 (同天內只計一次) 《有效文章》 0 篇 (退:0)
《目前動態》 查詢 parserTest0 《私人信箱》 最近無新信件
《上次上站》 08/23/2016 06:15:14 Tue 《上次故鄉》 118.233.159.81
《五子棋》 0 勝 0 敗 0 和 《象棋戰績》 0 勝 0 敗 0 和

《個人名片》 parserTest0 目前沒有名片

請按任意鍵繼續

Figure: Info

Data format

```
ID:String,  
news:String,  
lastTime:Date,  
lastIP:String,  
time:Date
```

Figure: Data format for user info

JACKY7's record

▼ (1) [_id : 57b83fda9940d6dc462ee3ea...]	{ 7 fields }	Document
_id	57b83fda9940d6dc462ee3ea	ObjectId
ID	JACKY7	String
news	中兩顆水球	String
lastTime	2016-08-20T08:09:26.000Z	Date
lastIP	114.44.216.31	String
time	2016-08-20T11:32:42.927Z	Date
_v	0	Int32
▼ (2) [_id : 57b895abbc2f799d3b4e...]	{ 7 fields }	Document
_id	57b895abbc2f799d3b4e6c68	ObjectId
ID	JACKY7	String
news	中三顆水球	String
lastTime	2016-08-20T08:09:26.000Z	Date
lastIP	114.44.216.31	String
time	2016-08-20T17:38:51.562Z	Date
_v	0	Int32
▼ (3) [_id : 57b9c2abc26997f2209d...]	{ 7 fields }	Document
_id	57b9c2abc26997f2209db42d	ObjectId
ID	JACKY7	String
news	中兩顆水球	String
lastTime	2016-08-21T12:06:27.000Z	Date
lastIP	114.44.216.31	String
time	2016-08-21T15:03:07.797Z	Date
_v	0	Int32
▼ (4) [_id : 57baa5e9361ab5af04ee...]	{ 7 fields }	Document
_id	57baa5e9361ab5af04eeb7e7	ObjectId
ID	JACKY7	String
news	中八顆水球	String
lastTime	2016-08-22T03:18:39.000Z	Date
lastIP	114.44.216.31	String
time	2016-08-22T07:12:41.911Z	Date
_v	0	Int32

Figure: Somebody's data