Time skew analysis using web cookies

Björgvin Ragnarsson

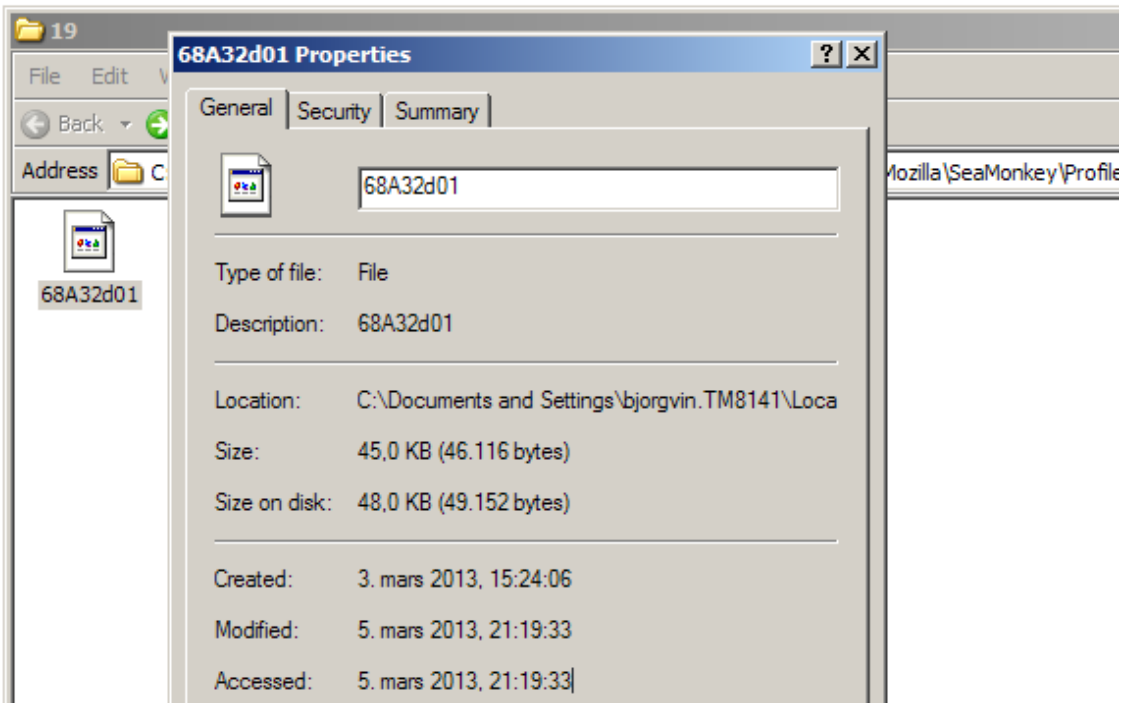
07-03-2013

The problem

- Timestamps are important for forensics...
- ...but the timekeeper is unreliable
- How far off was the system clock when the timestamp was created?

11 years ago: the solution

```
58  <div class="clocknow">
59    <span class="stream">klukkan er</span>
60    <span id="LiveClock" data-date="Tue Mar 05 2013 20:16:09 GMT">20:16</span>
61  </div>
62  <hr class="stream" />
63
64  </div>
65  </div>
66
67
68
69    <div class="middle">
70      <div class="wrap">
71
72  <div class="pgmain">
73  <div class="wrap">
74  <div class="tophead">
75
76  </div>
```



11 years ago: problems

- Manual work
- Dynamic or static timestamps?
- Is the server time reliable?

Deriving skew from cookies (1/3)

```
HTTP/1.0 200 OK
Date: Fri, 21 Sep 2012 05:51:31 GMT
Status: 200 OK
Set-Cookie:
   productId=17;
   expires=Fri, 28-Sep-12 05:51:31 GMT;
   domain=example.com
```

Deriving skew from cookies (2/3)

```
id:             9768
baseDomain:     example.com
name:           productId
value:          17
host:           example.com
path:           /
expiry:         1348811491
creationTime:   1348206691
```

Deriving skew from cookies (3/3)

```
Set-Cookie:
    productId=17;
    Max-Age=604800;
    domain=example.com;
```

Algorithm 1: ranking possible skews

For each cookie in a browser cookie DB:

1. Find probability that it usable
2. Calculate possible skews
3. Add probability to the rank of each possible skew

Processing the corpus

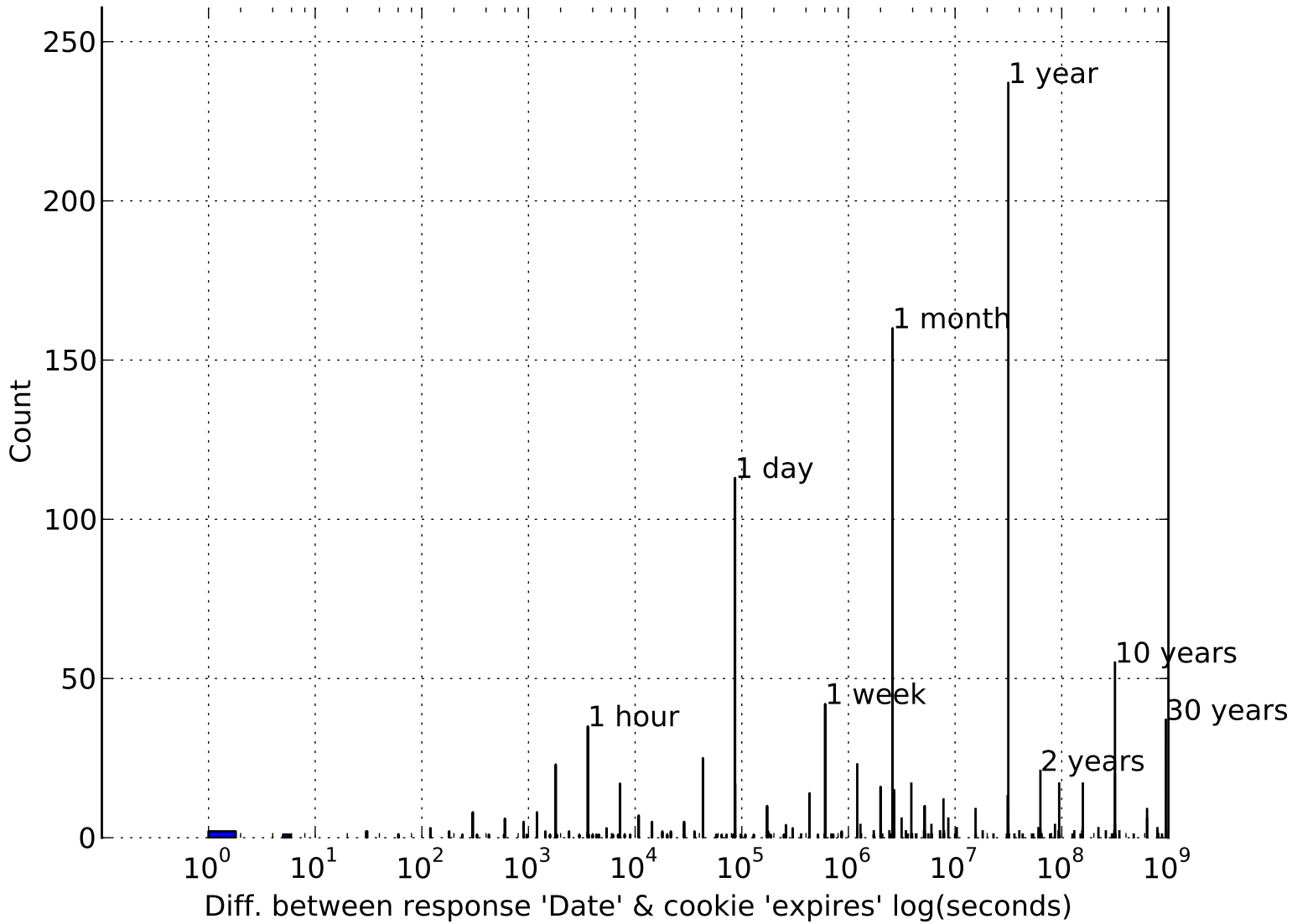
Web sites requested	10.000
Number of User agents used	14
Cookies in responses	59.453
Cookies with both `Max-Age` and `expires`	481
Cookies with only `Max-Age`	355
Cookies with only `expires`	28.764

Table: Statistics on the HTTP Header Survey, 2012/09/22

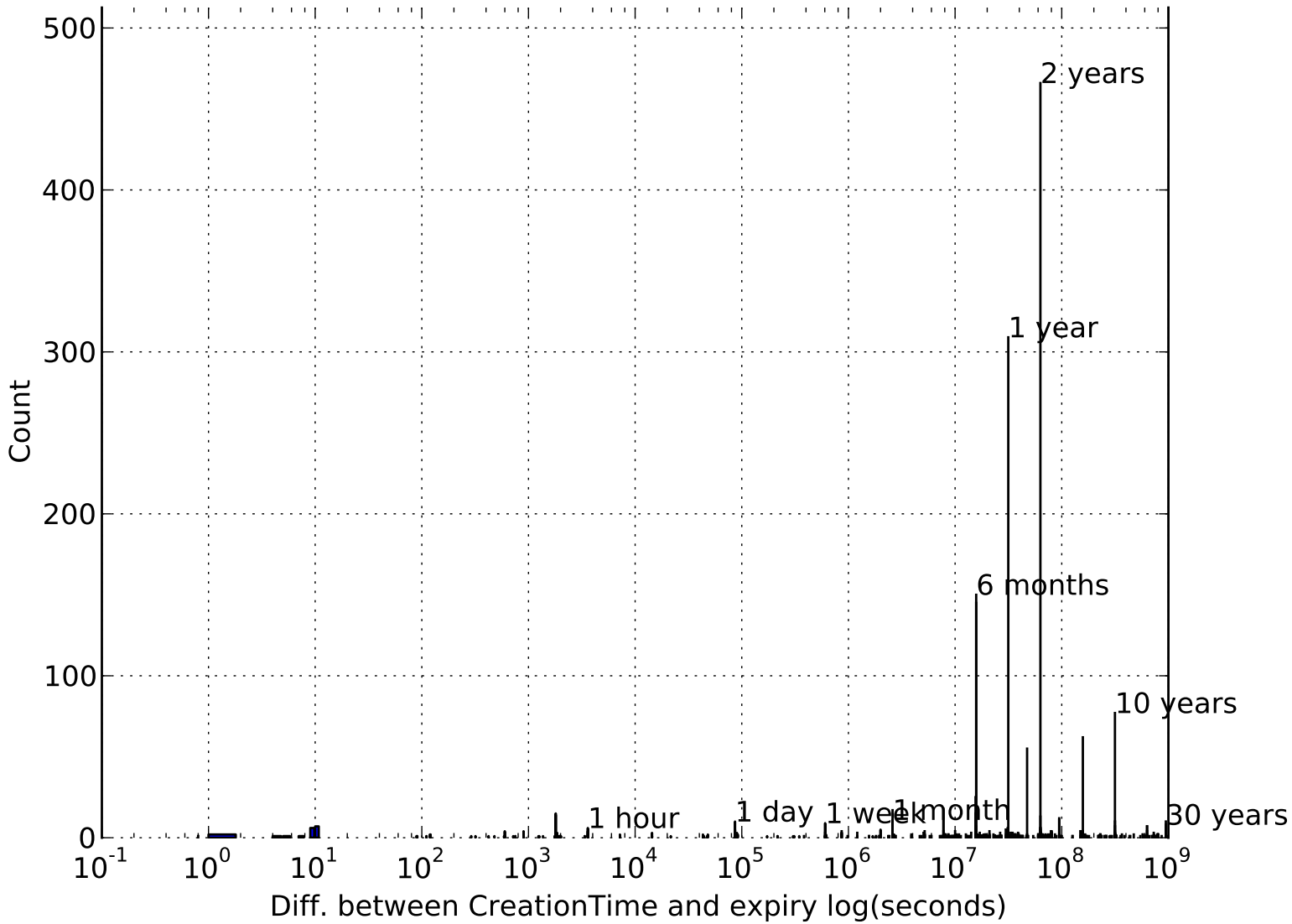
Processing the corpus: Frequency of bad expiry dates

95	2019-12-23 23:50:00
67	1970-01-01 00:00:01
16	2020-02-19 14:28:00
13	1970-01-01 00:00:10
10	2019-12-31 23:00:00
10	1970-01-01 00:00:00
9	2096-10-02 07:06:40
9	2037-12-31 23:55:55
8	2038-01-19 03:14:07
7	1970-01-01 12:00:01

Processing the corpus: Acquiring server deltas



Processing the corpus: Comparison to a Firefox DB



Ranking possible skews: results

```
$ skewy.py -c 83sback.sqlite -z top10k.db \
-j 0.2 -m 0.028 -bdl BDL.csv -p
```

	skew	rank	cookiecount	cookieratio
1	-83	0.31	1104	0.39
2	63071917	0.26	936	0.33
3	86317	0.22	780	0.27
4	31535917	0.20	719	0.25
5	-31449683	0.19	677	0.24

...

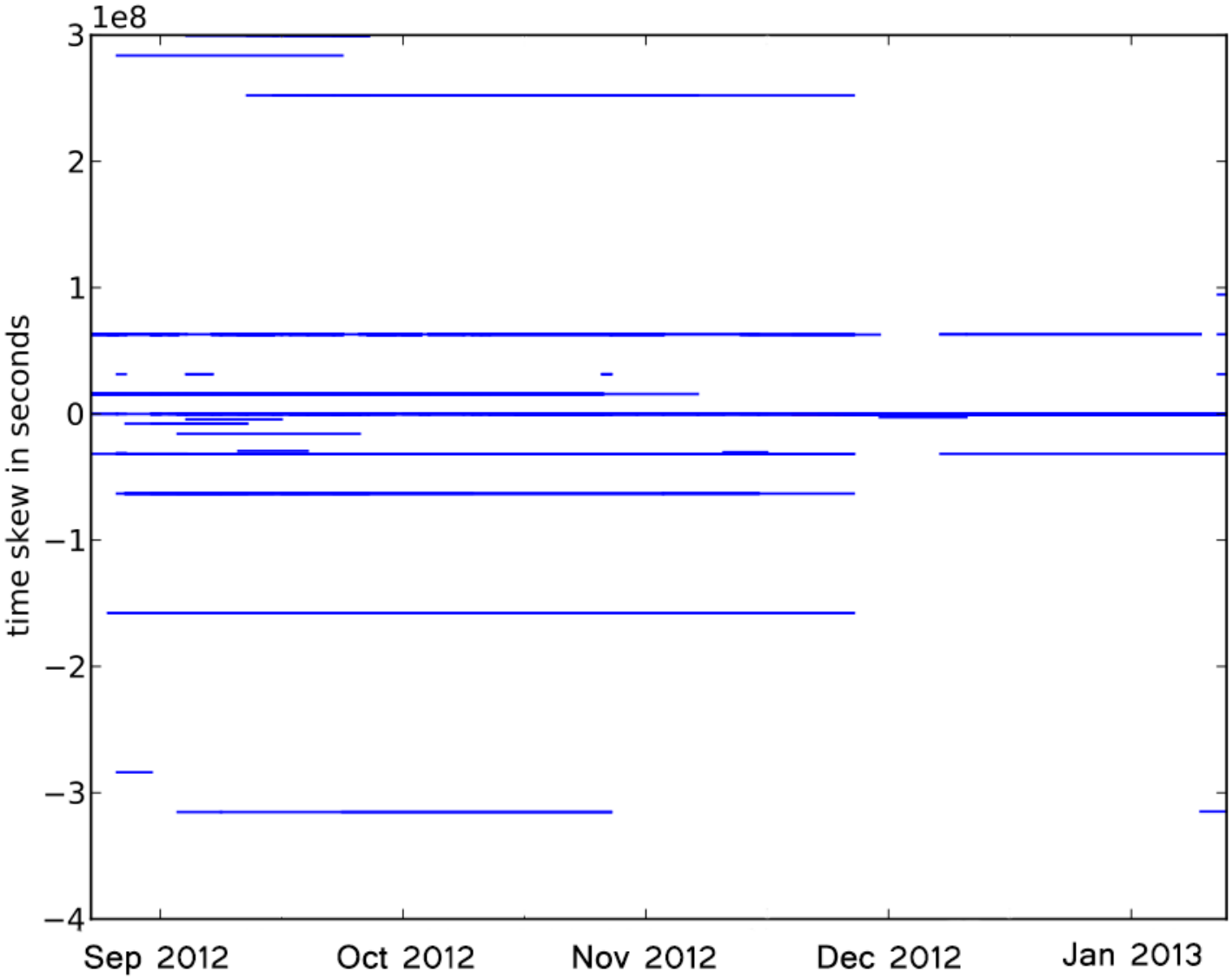
Algorithm 2: finding different skews

Find all groups of 4 cookies which

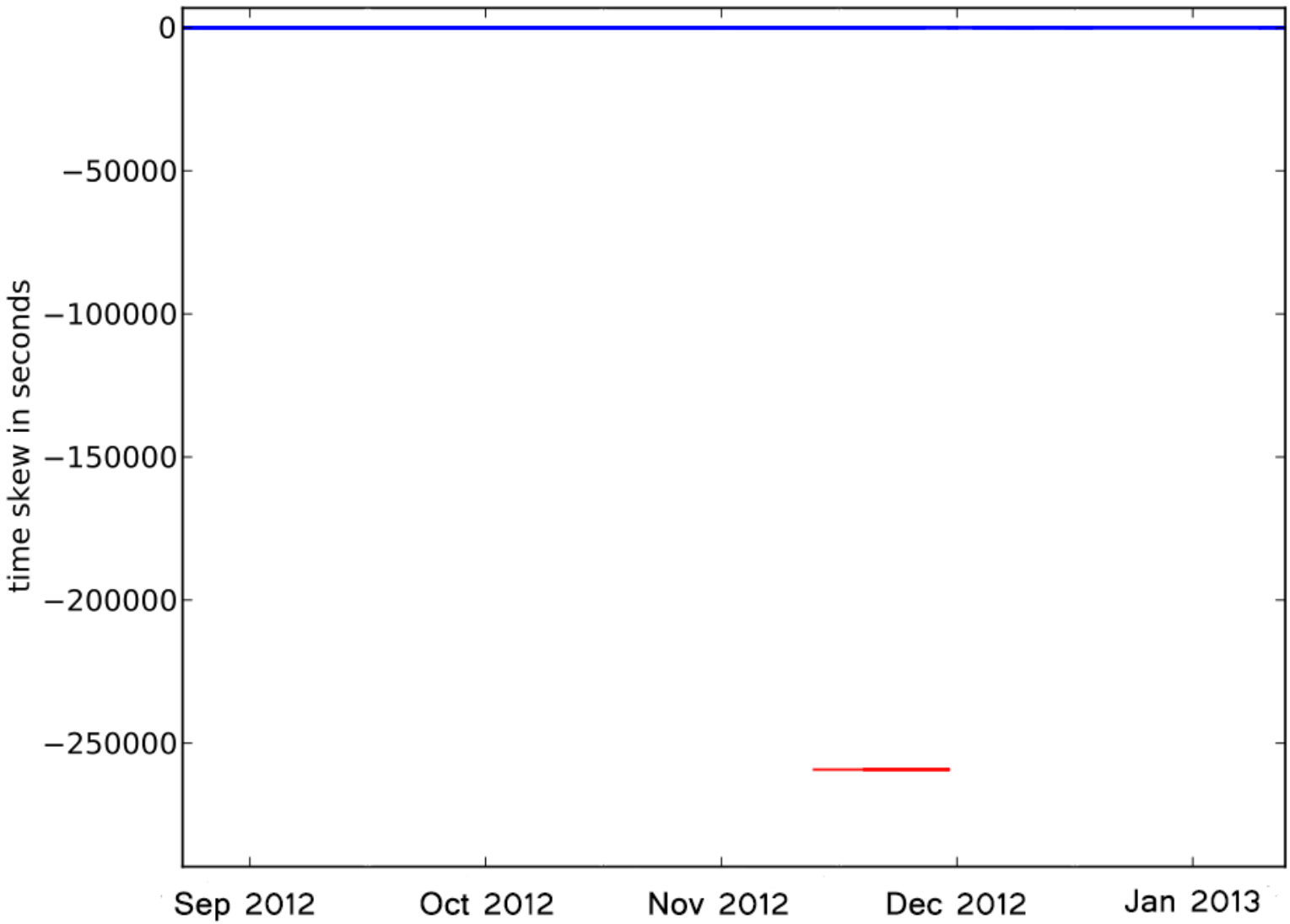
1. have the same possible skews
2. have different deltas
3. are close as possible in creation time

Display the period the group spans

Algorithm 2: Results (1/2)



Algorithm 2: Results (2/2)



Conclusions

- Algorithm 1 ranks the correct skew as #1
- Algorithm 2 needs more work
- More testing is needed