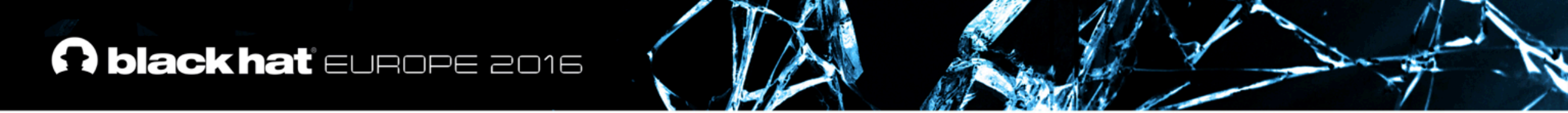

black hat®
EUROPE 2016



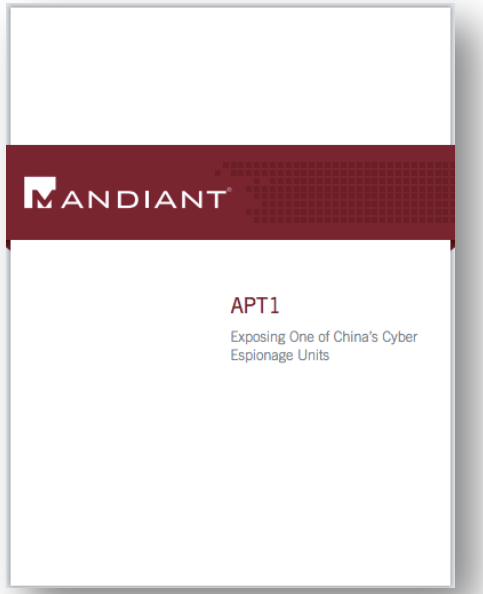
Chasing Foxes by the Numbers: Patterns of Life and Activity in Hacker Forums

Christopher Ahlberg, Co-Founder & CEO

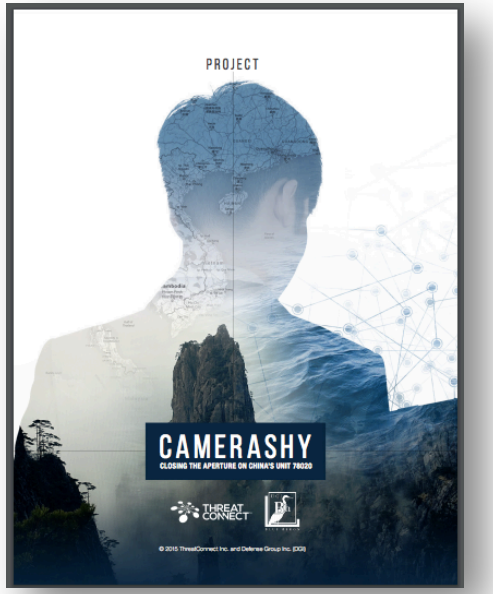
Recorded Future

c@recordedfuture.com
@cahlberg

Attribution [Often] Based on Sloppy Handle Usage



UglyGorilla
SuperHard

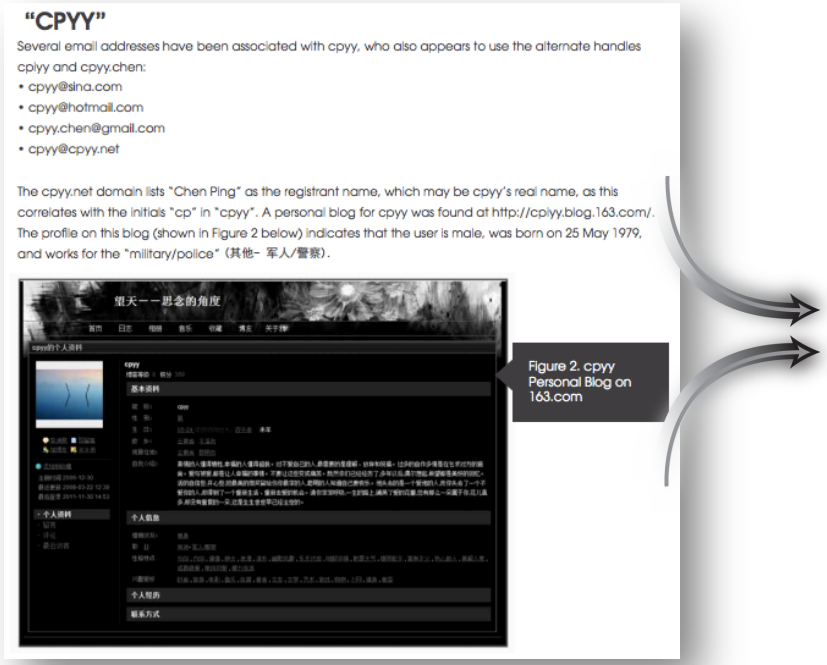


GreenSky27

QQ identity
5054-3533

cpyy

Handles and Forums key to attribution



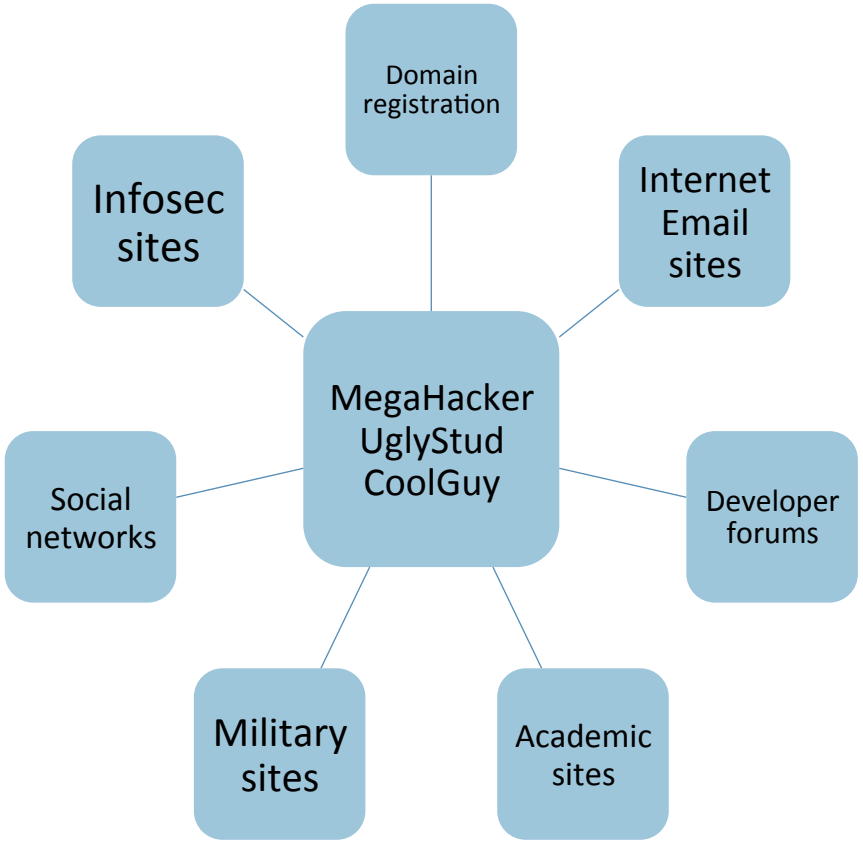
"CPYY"

Several email addresses have been associated with cpyy, who also appears to use the alternate handles cpiyy and cpyy.chen:

- cpyy@sina.com
- cpyy@hotmail.com
- cpyy.chen@gmail.com
- cpyy@cpyy.net

The cpyy.net domain lists "Chen Ping" as the registrant name, which may be cpyy's real name, as this correlates with the initials "cp" in "cpyy". A personal blog for cpyy was found at http://cpiyy.blog.163.com/. The profile on this blog (shown in Figure 2 below) indicates that the user is male, was born on 25 May 1979, and works for the "military/police" (其他- 军人/警察).

Figure 2. cpyy Personal Blog on 163.com



If Better Actor Opsec, Then What?

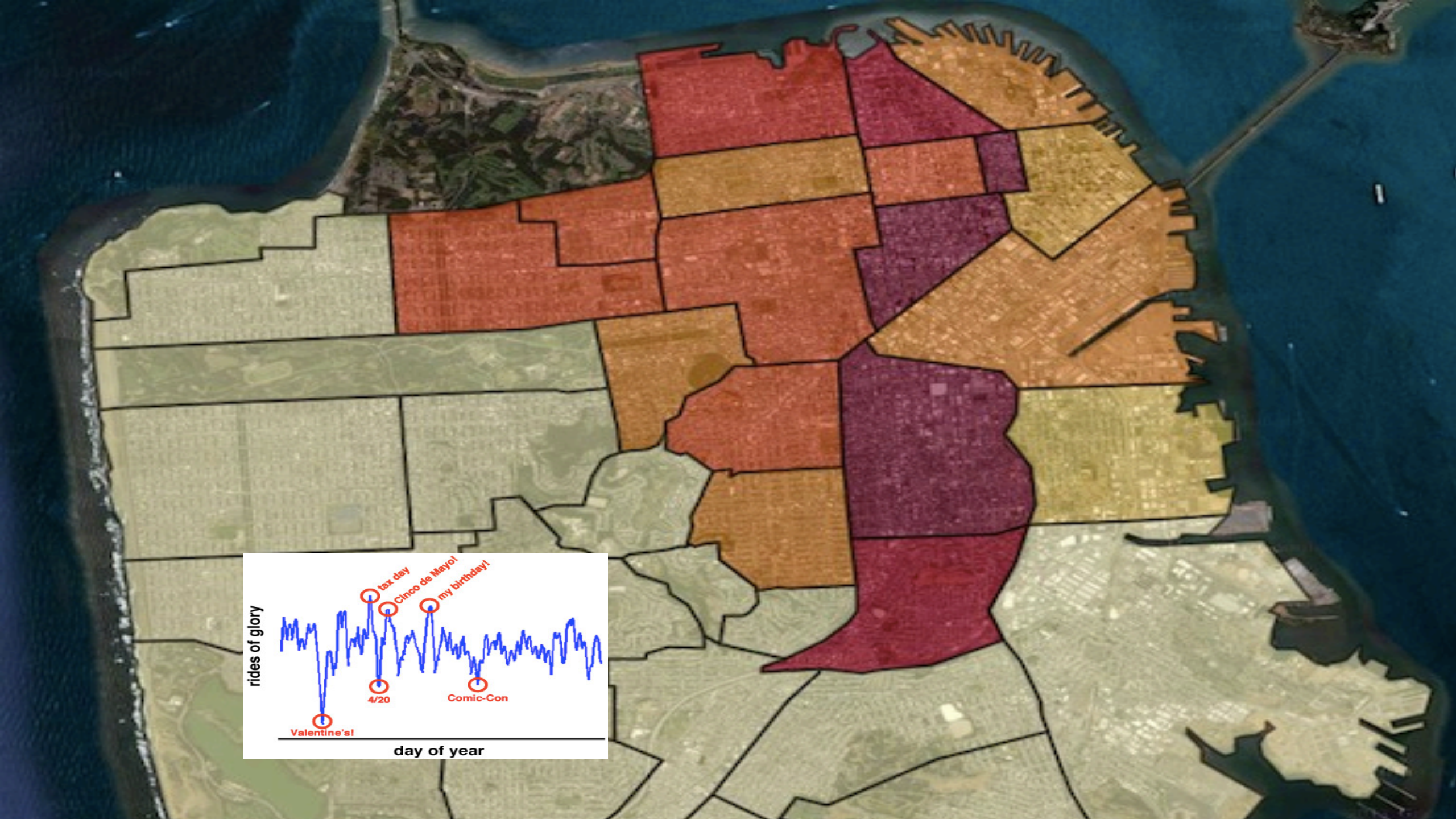
Forum Targeting

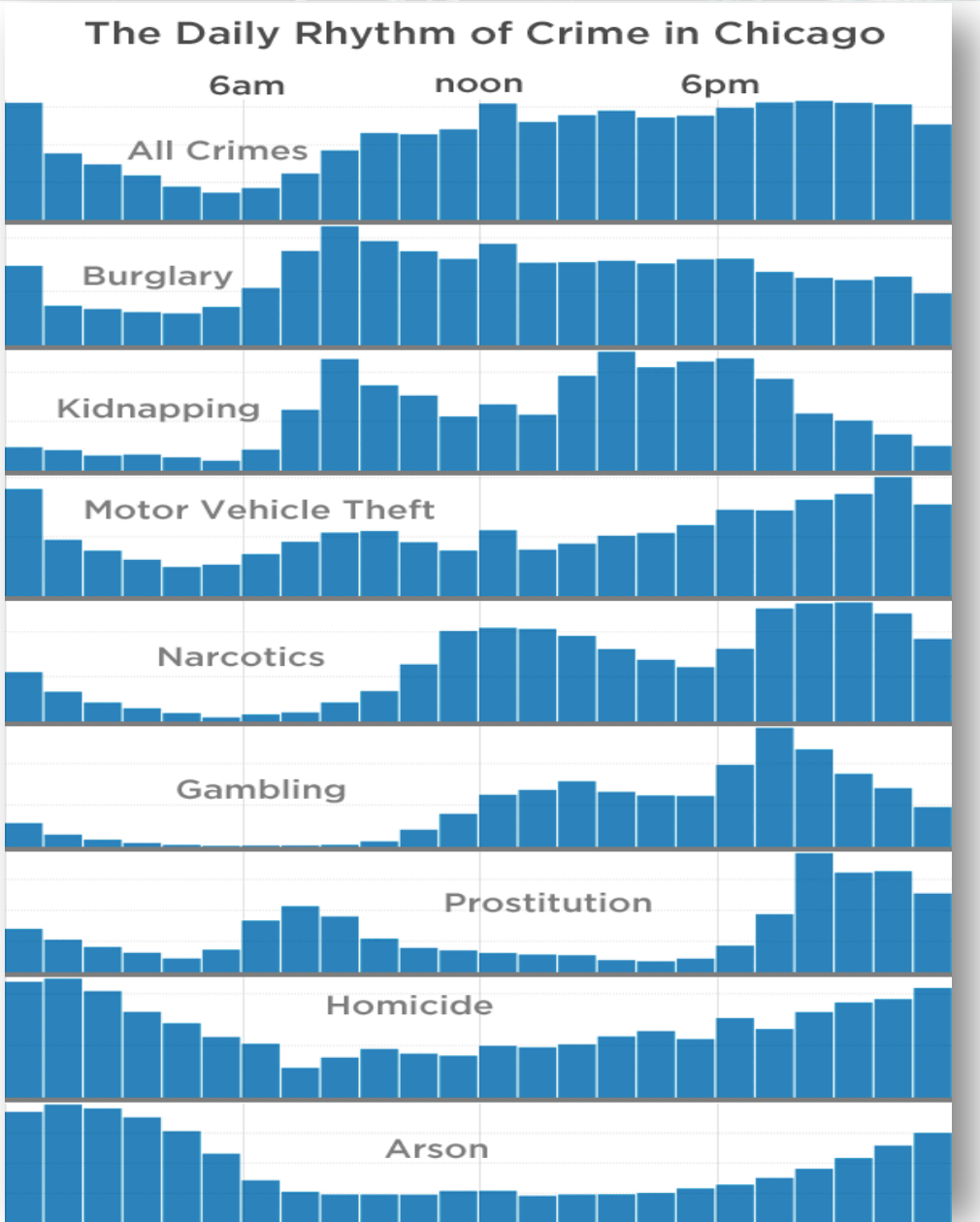
Key handles in forums

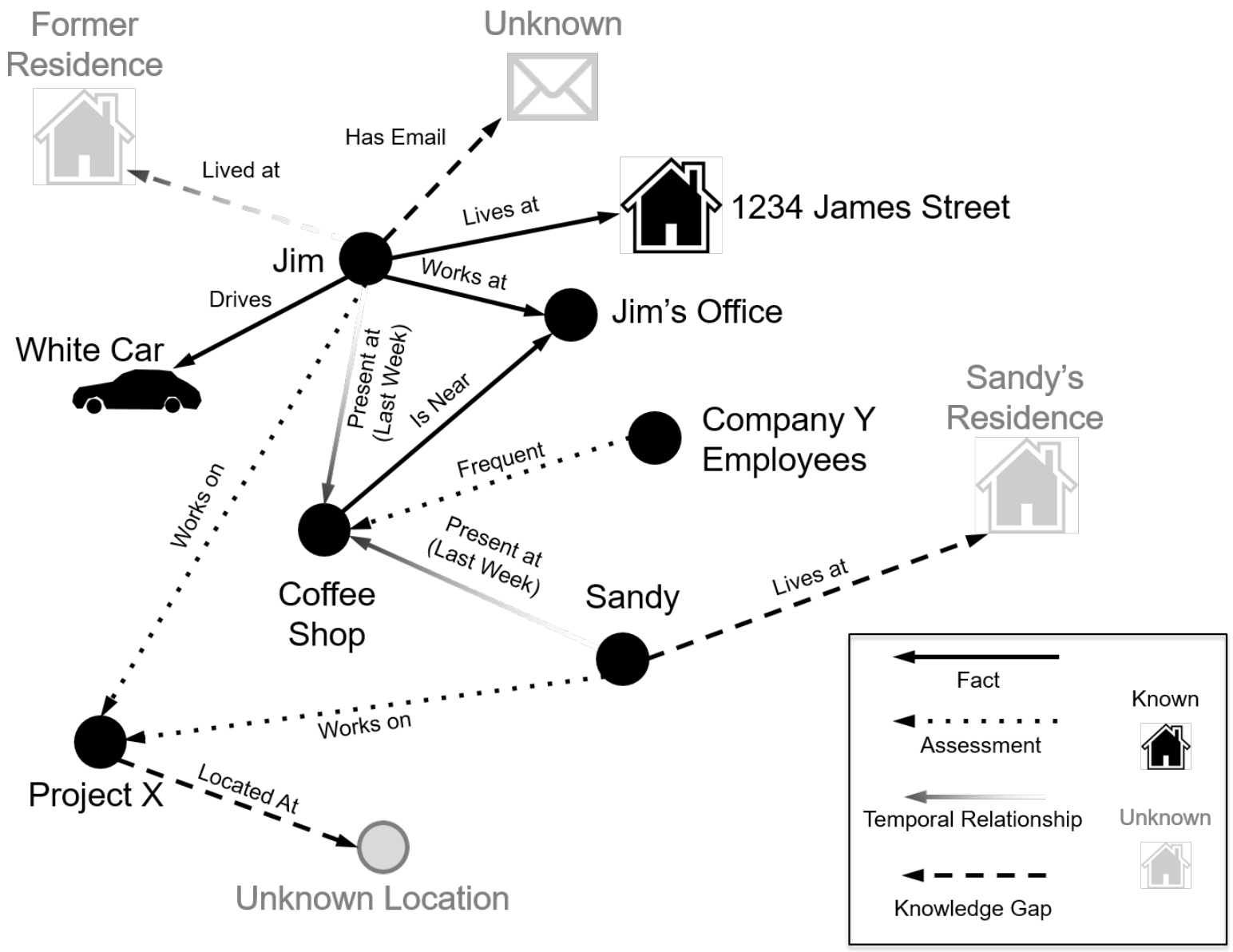
Handle hopping within and Across forums

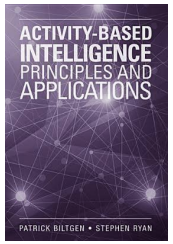
Crews of handles

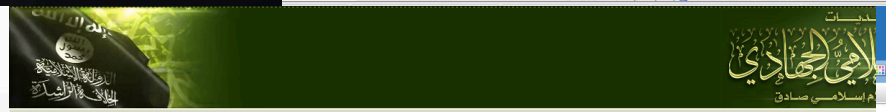
Can we apply some $$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{1/2}}$$?

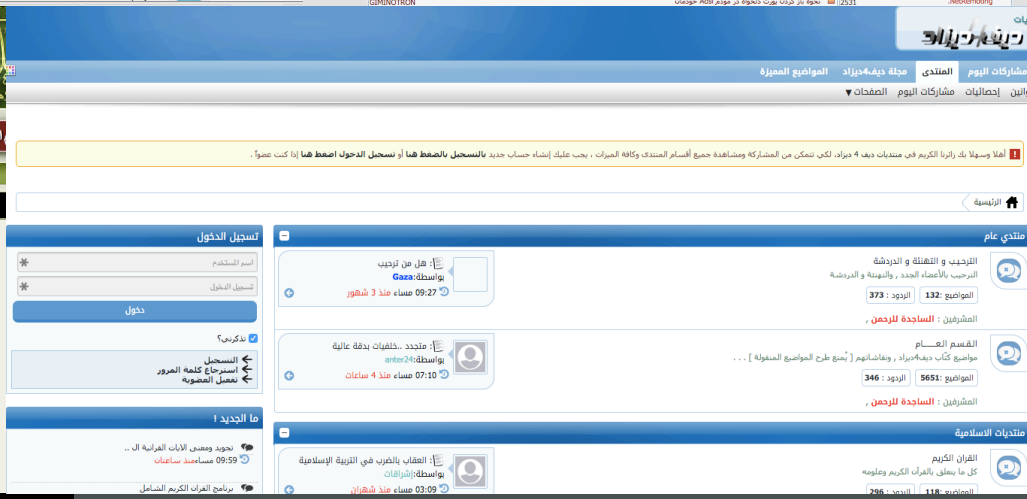

tax day
Cinco de Mayo!
my birthday!
4/20
Comic-Con
Valentine's!
rides of glory
day of year


The Daily Rhythm of Crime in Chicago
6am
noon
6pm
All Crimes
Burglary
Kidnapping
Motor Vehicle Theft
Narcotics
Gambling
Prostitution
Homicide
Arson

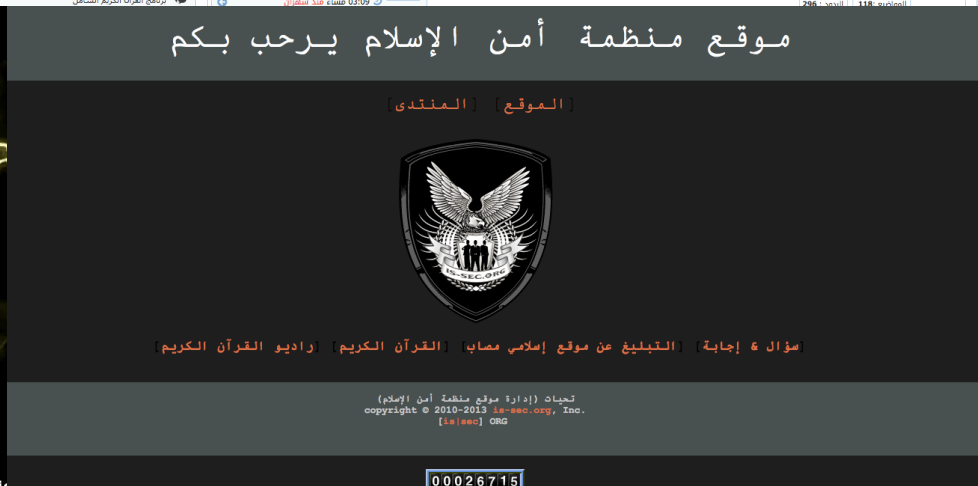








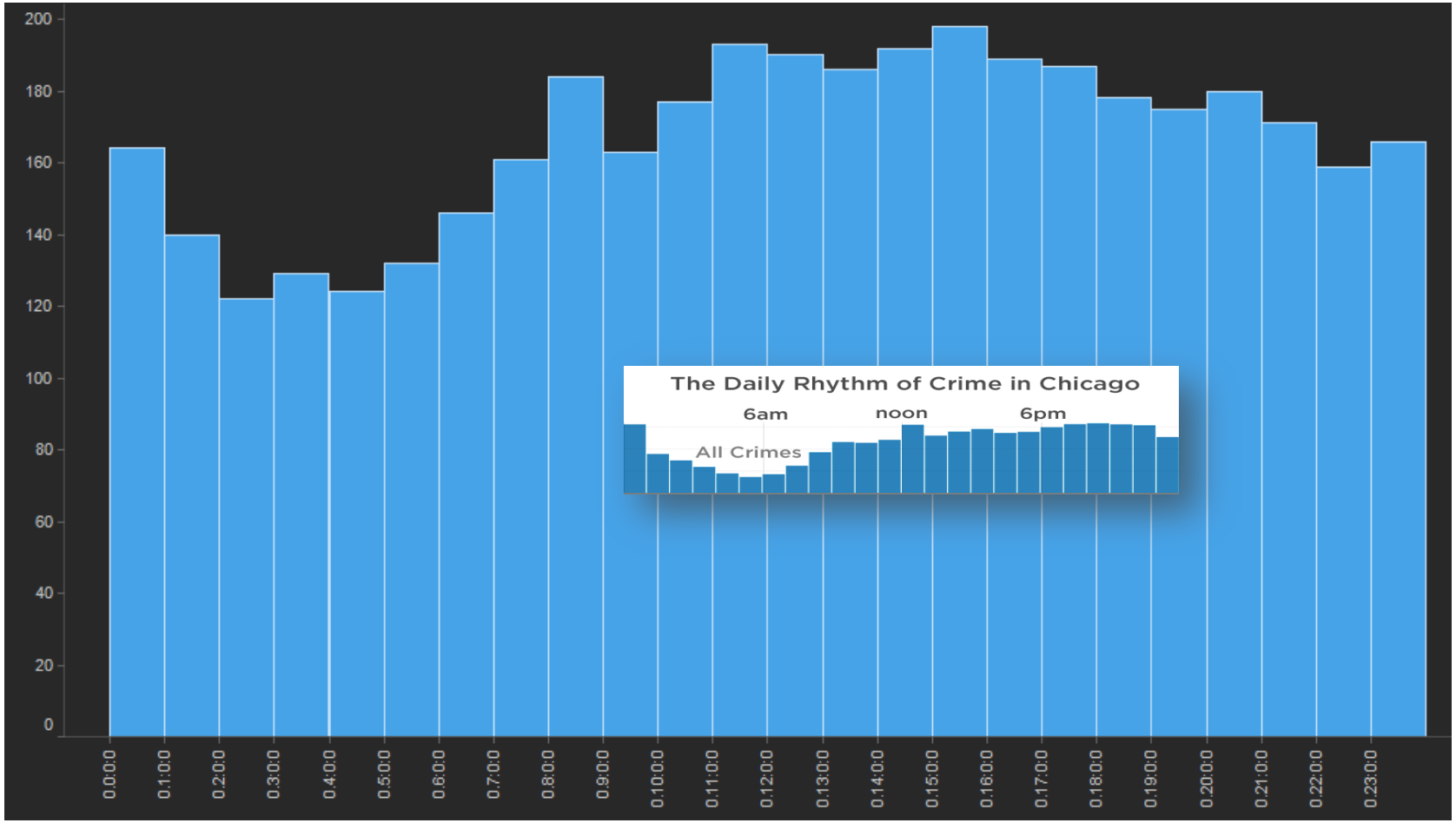
الشكل سهل . أما المضمون . فمن المستحيل تقليده



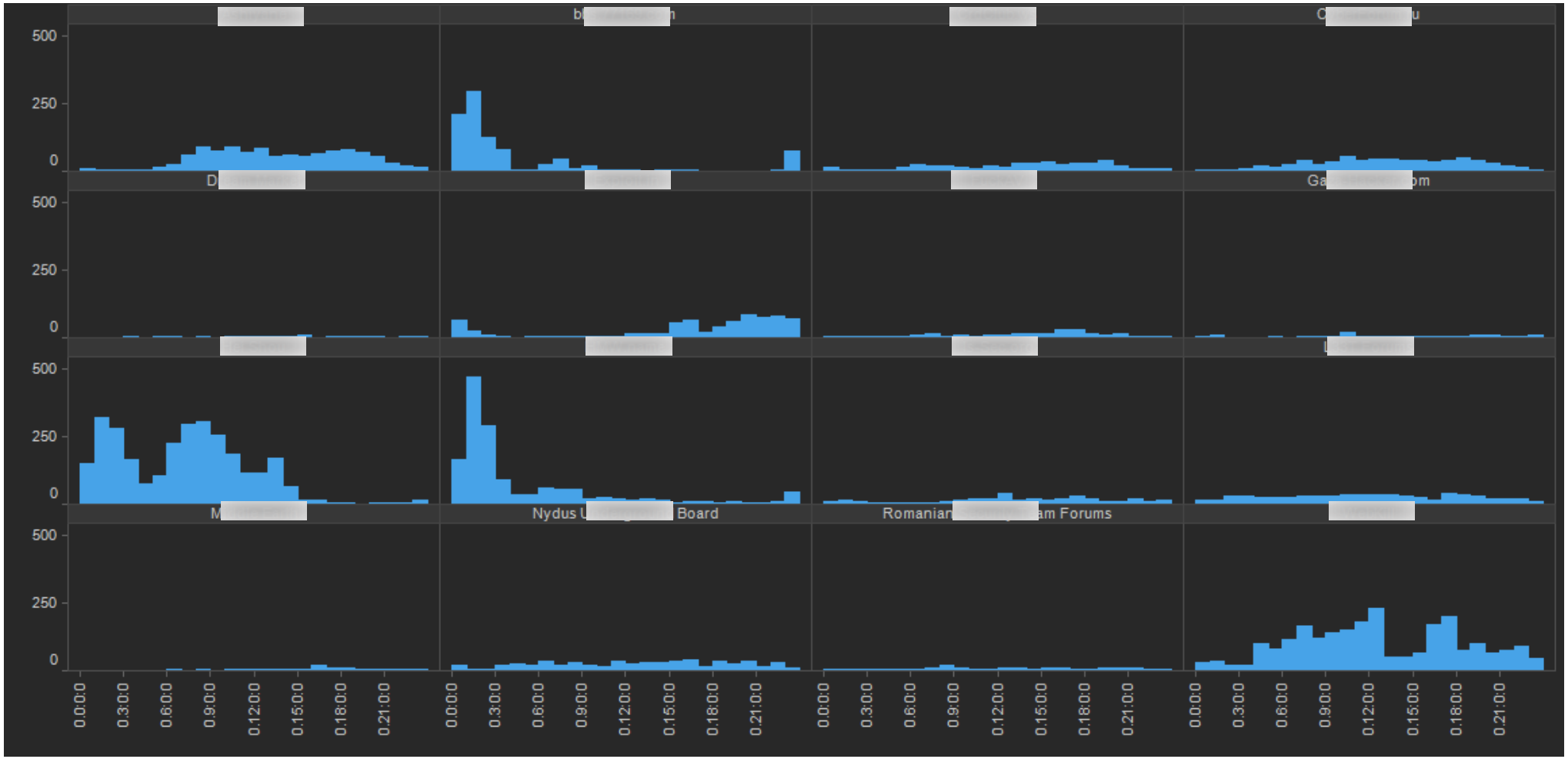
- Automated collection
- 750 hacker/criminal forums from Deep/Dark web
- 4 year collection
- 7 languages
- Indexed
- NLP/Feature extraction for cyber attacks, vulnerabilities, exploits, technical indicators, etc.

Surprisingly little handle re-use
96.6%+ used once
1,446,815 total

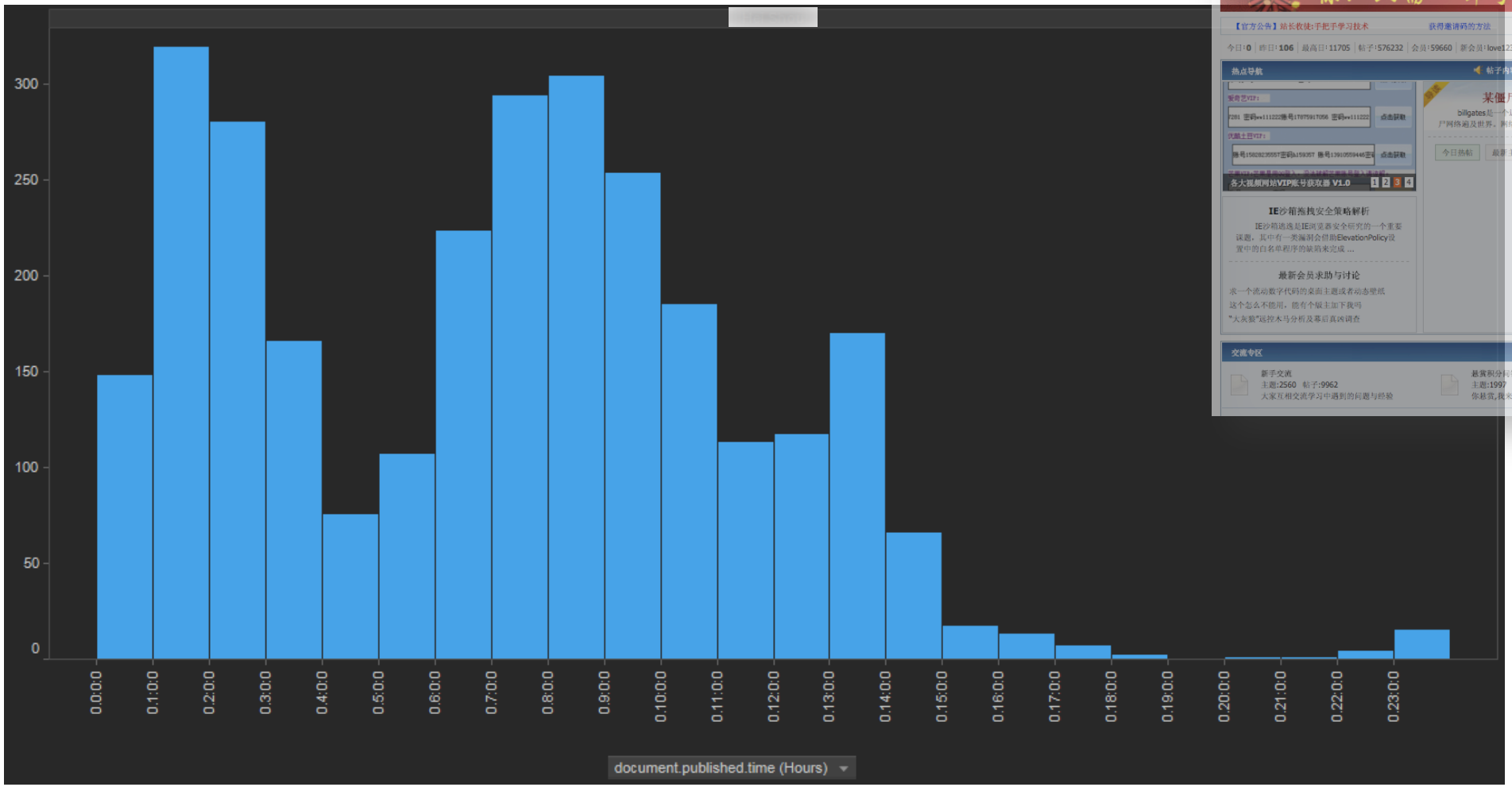
First Pattern: Hackers Have Hours

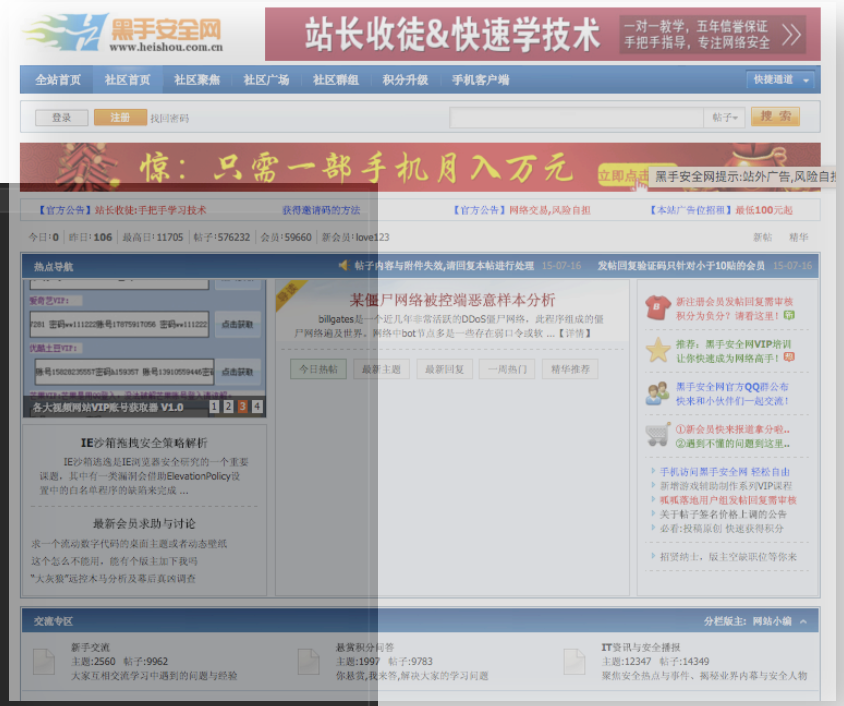


They Just Differ

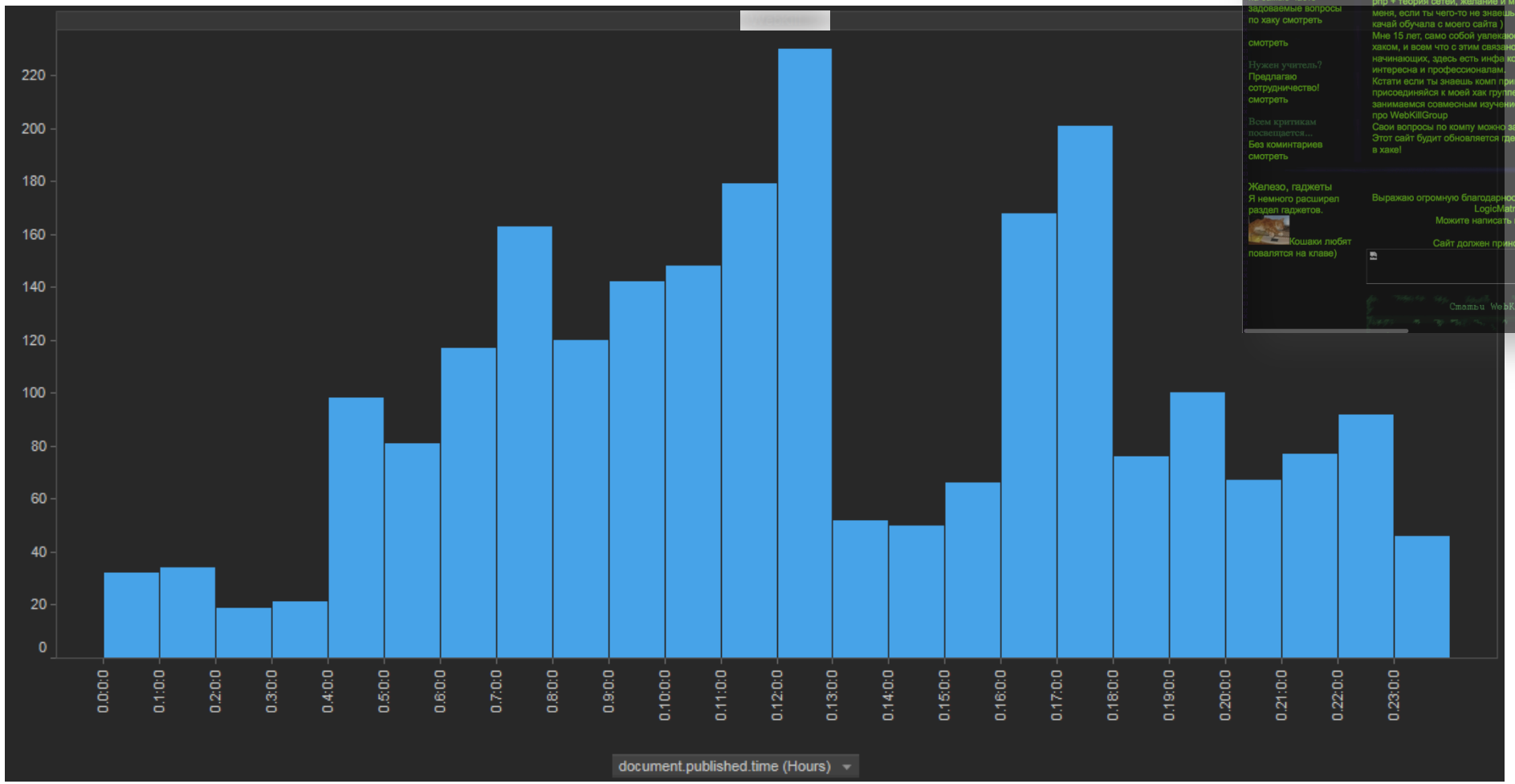


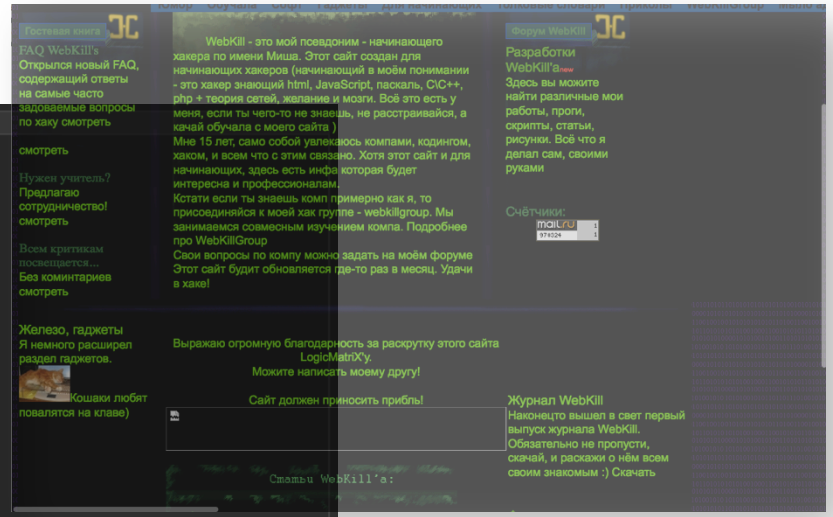
Distinct Patterns

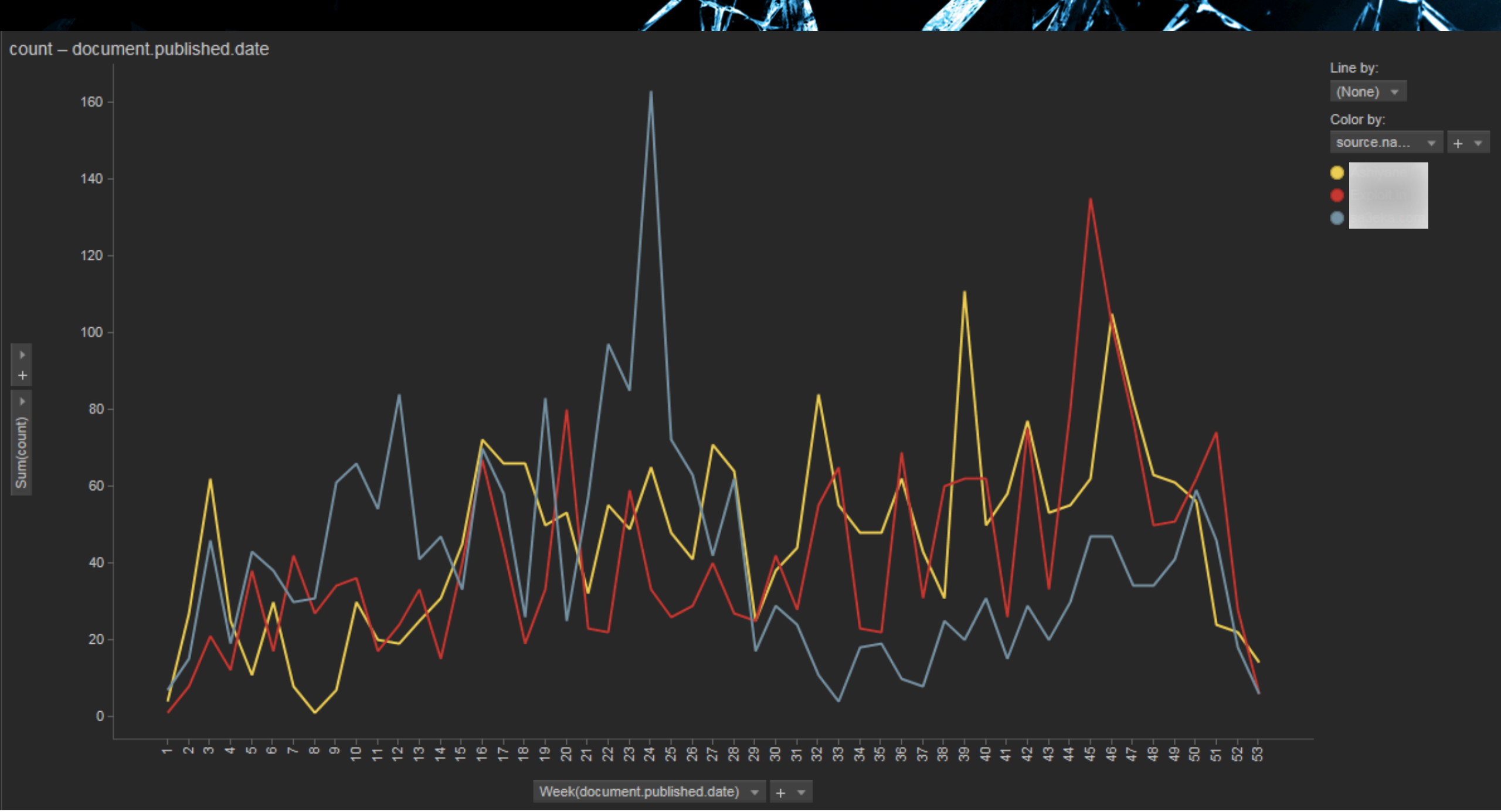




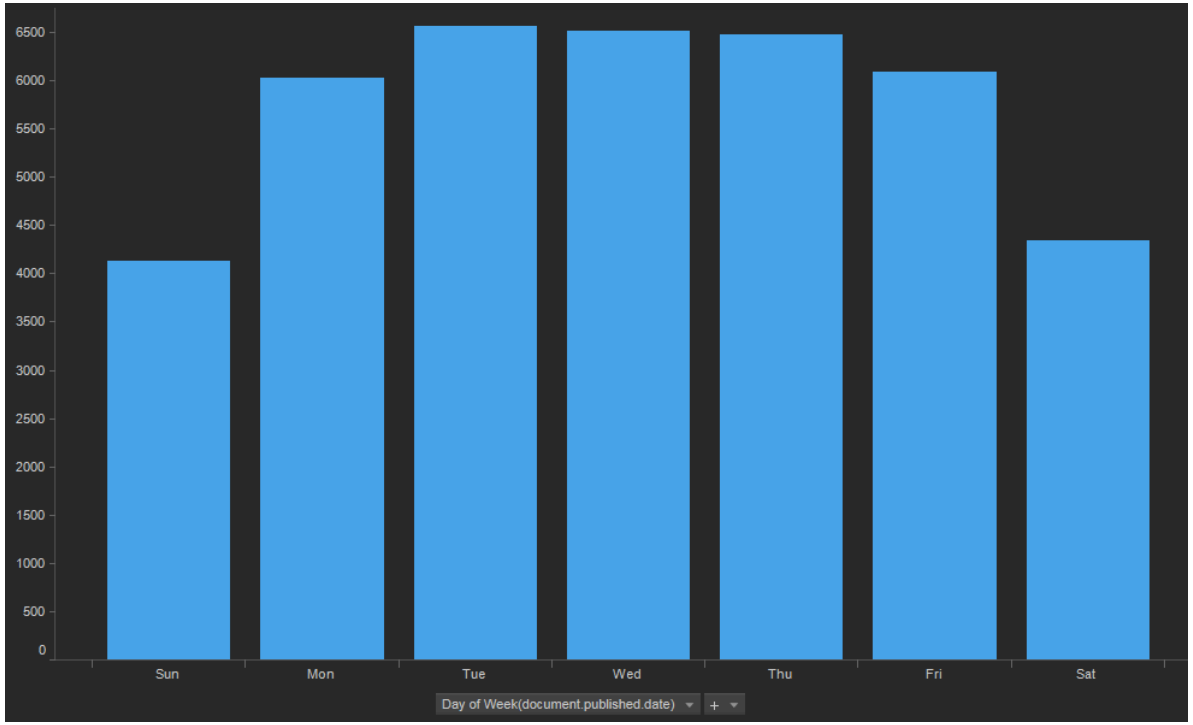
Multiple Geos or Crews?

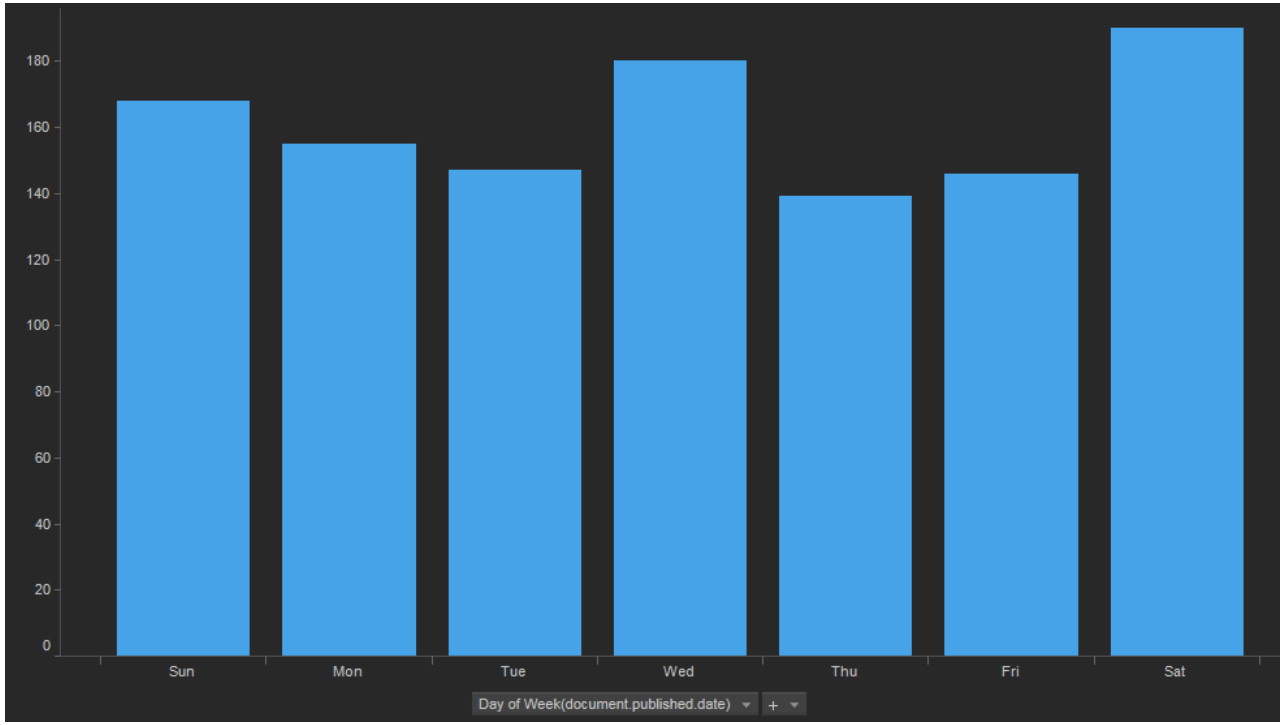





count – document.published.date
Line by:
(None)
Color by:
source.na...
Sum(count)
160
140
120
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53
Week(document.published.date)

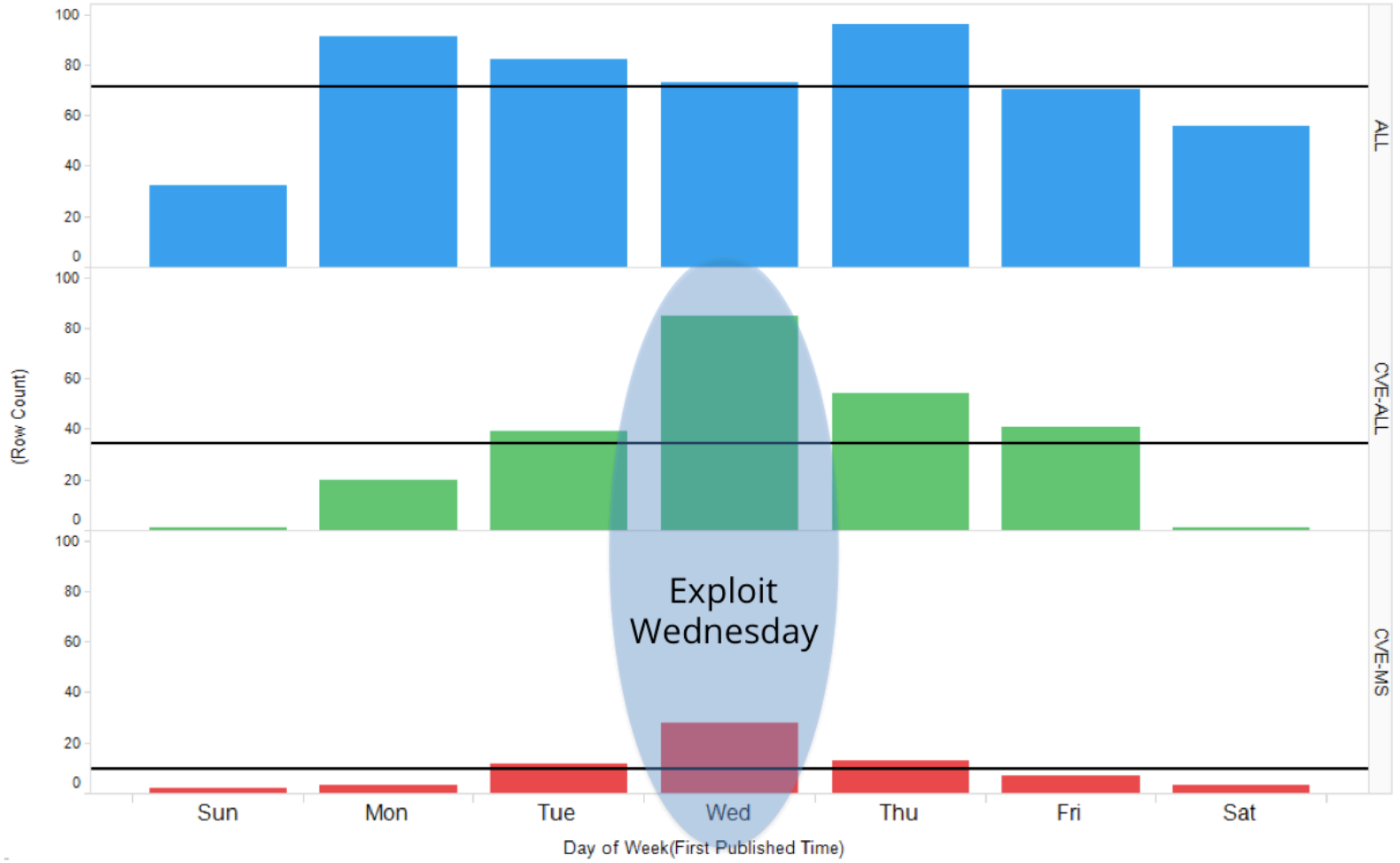
Distinct Weekly Patterns



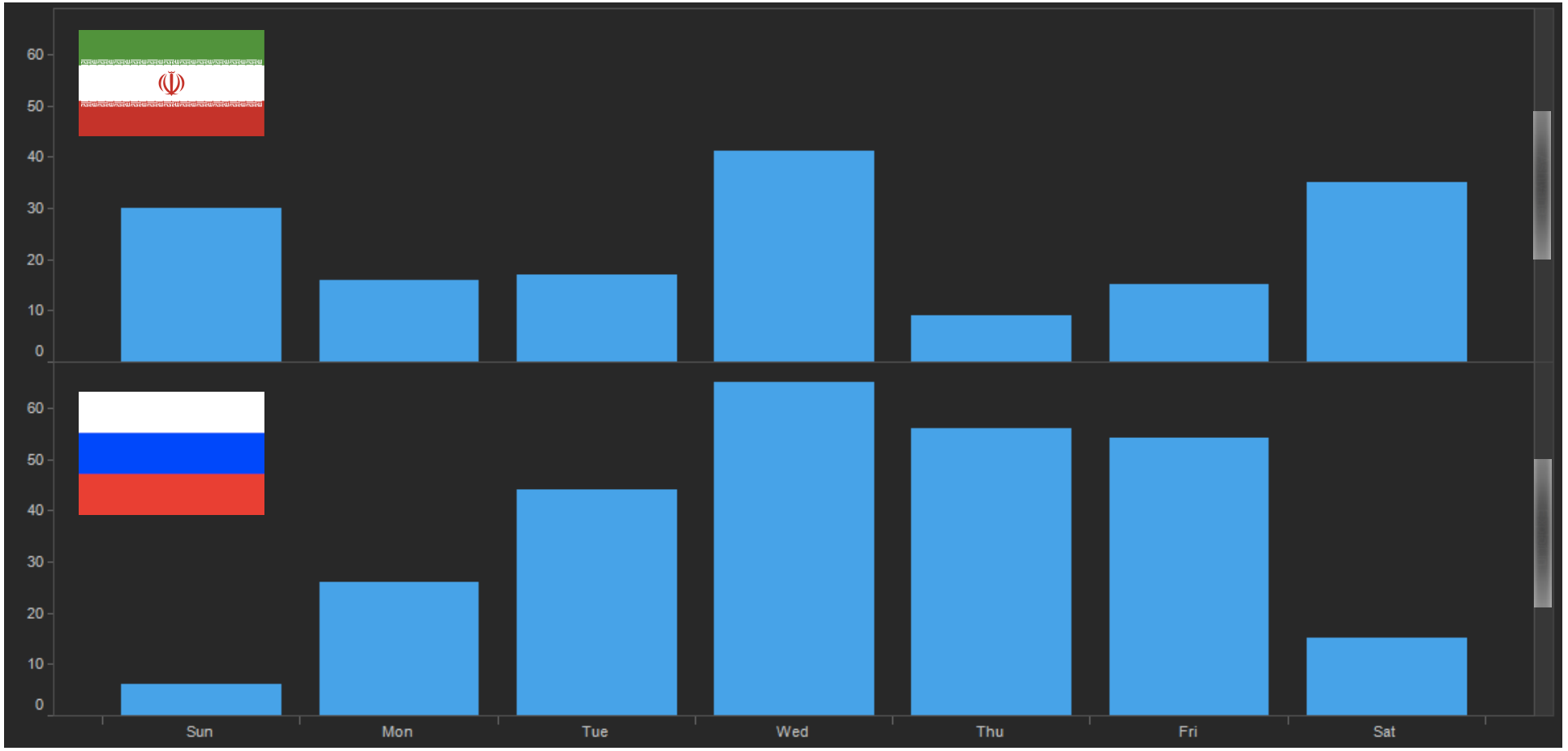


Smoking out the Rats with External Events

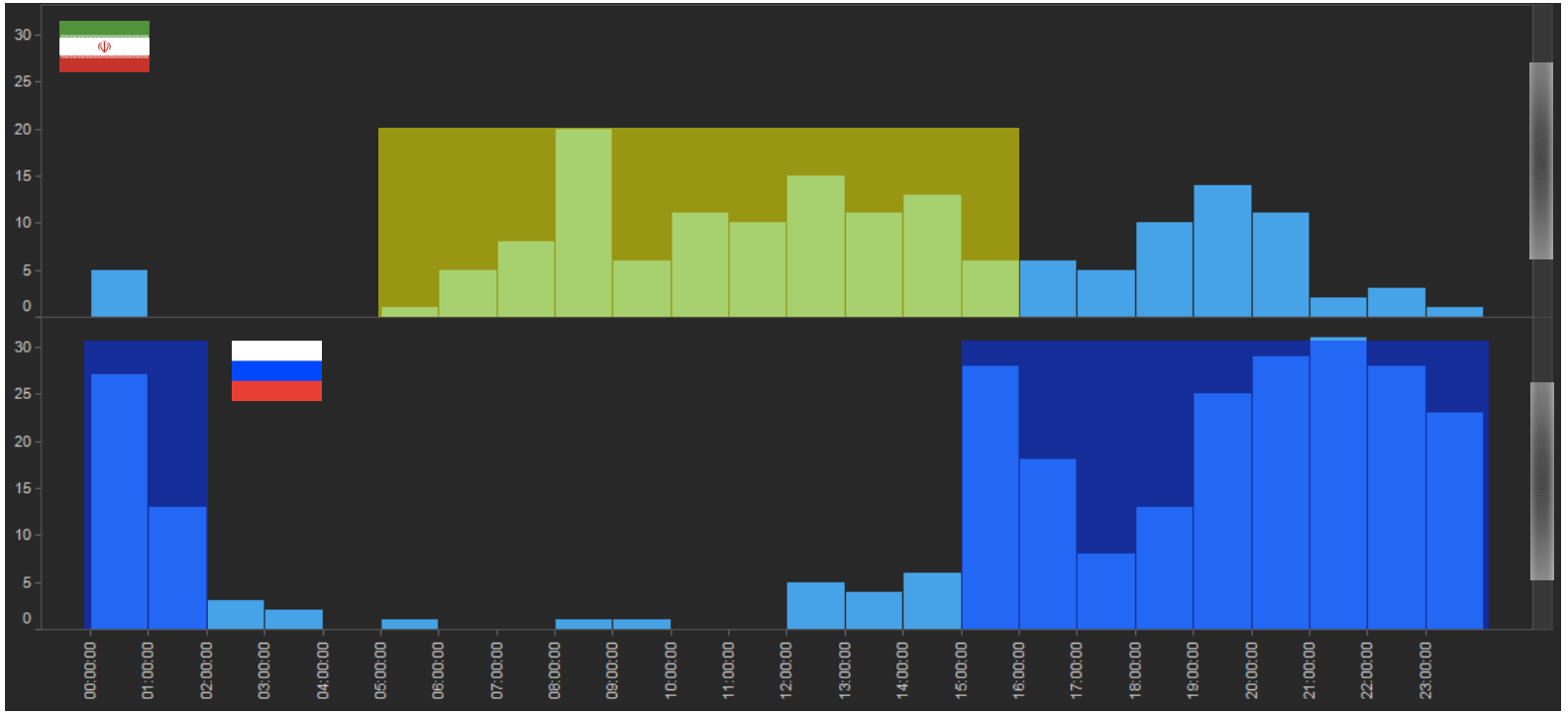
Patch Tuesday Driving Exploit Wednesday



Patch Tuesday Drives Russian and Iranians

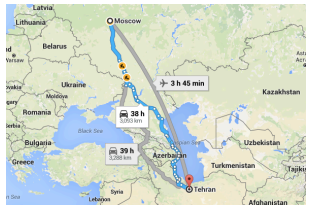


Even When Work Hours Are Different

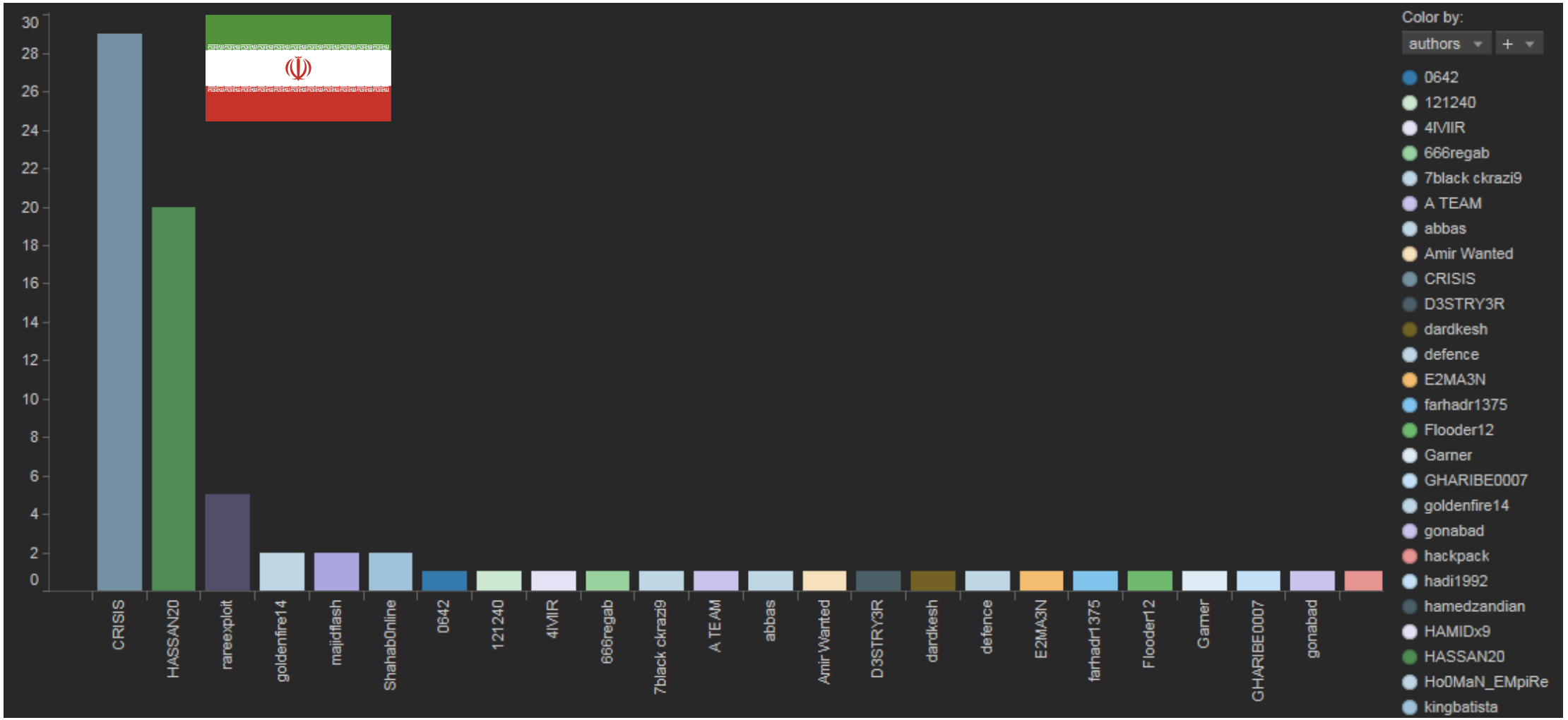


Tehran, Iran is 30 minutes ahead of Moscow, Russia

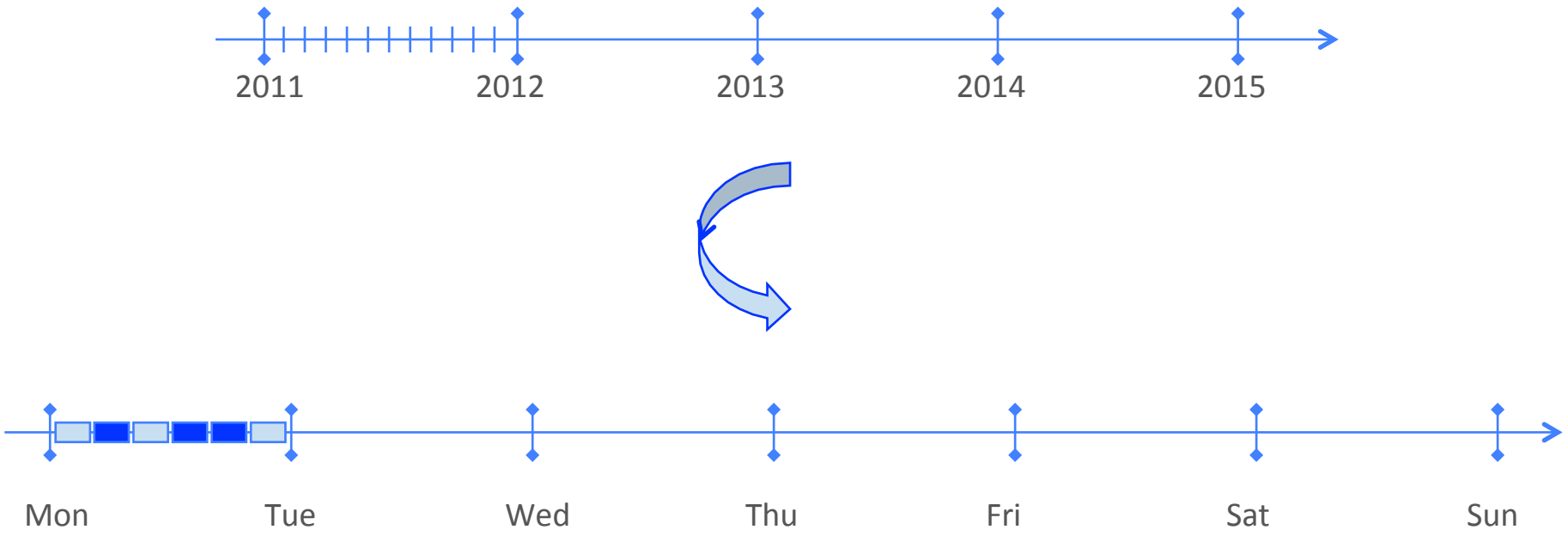
5:54 AM Saturday, in Moscow, Russia is
6:24 AM Saturday, in Tehran, Iran



Two Kings of Vulnerabilities



Clustering Pattern of Life



Understanding Behavior by Math

The *correlation* between two points, *a* and *b*, with *k* dimensions is calculated as:

$$\frac{cov(a,b)}{std(a) \times std(b)}$$

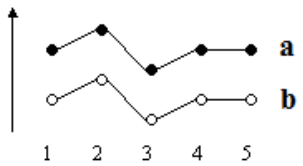
where

$$cov(a,b) = \frac{1}{k}\sum_{j=1}^{k}(a_j - \bar{a}) \times (b_j - \bar{b})$$

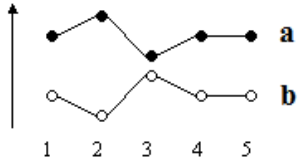
$$std(a) = \sqrt{\frac{1}{k}\sum_{j=1}^{k}(a_j - \bar{a})^2}$$

$$\bar{a} = \frac{1}{k}\sum_{j=1}^{k} a_j$$

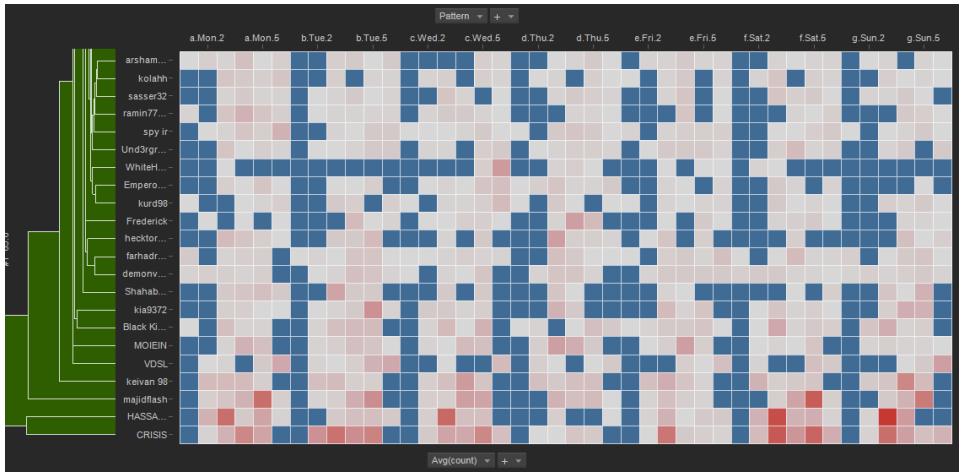
This correlation is called *Pearson Product Momentum Correlation*, simply referred to as *Pearson's correlation* or *Pearson's r*. It ranges from +1 to -1 where +1 is the highest correlation. Complete opposite points have correlation -1.

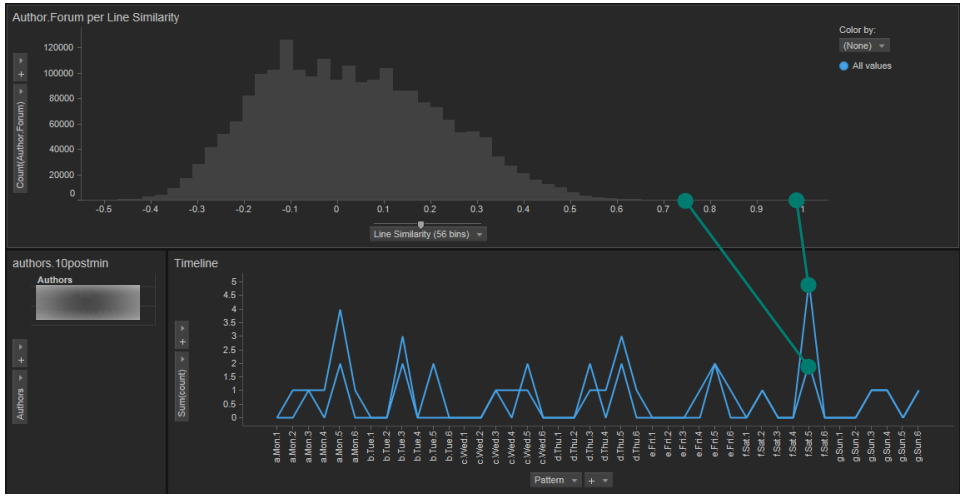


a and *b* are identical, which means they have maximum correlation.

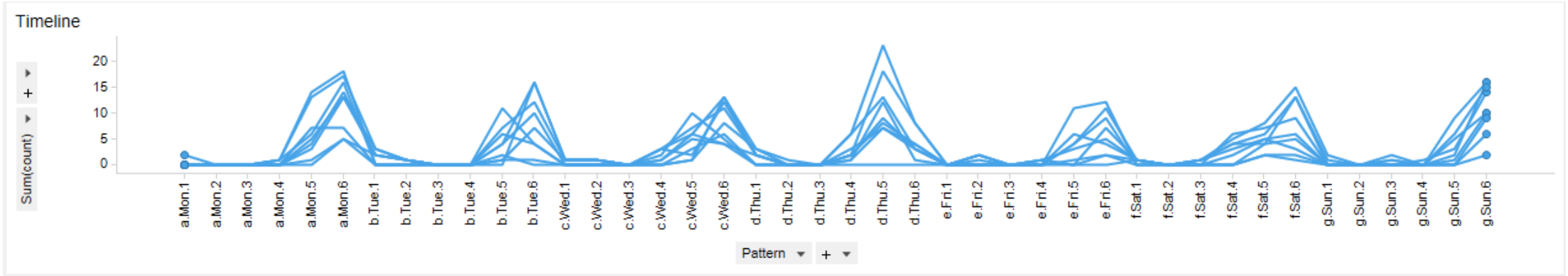


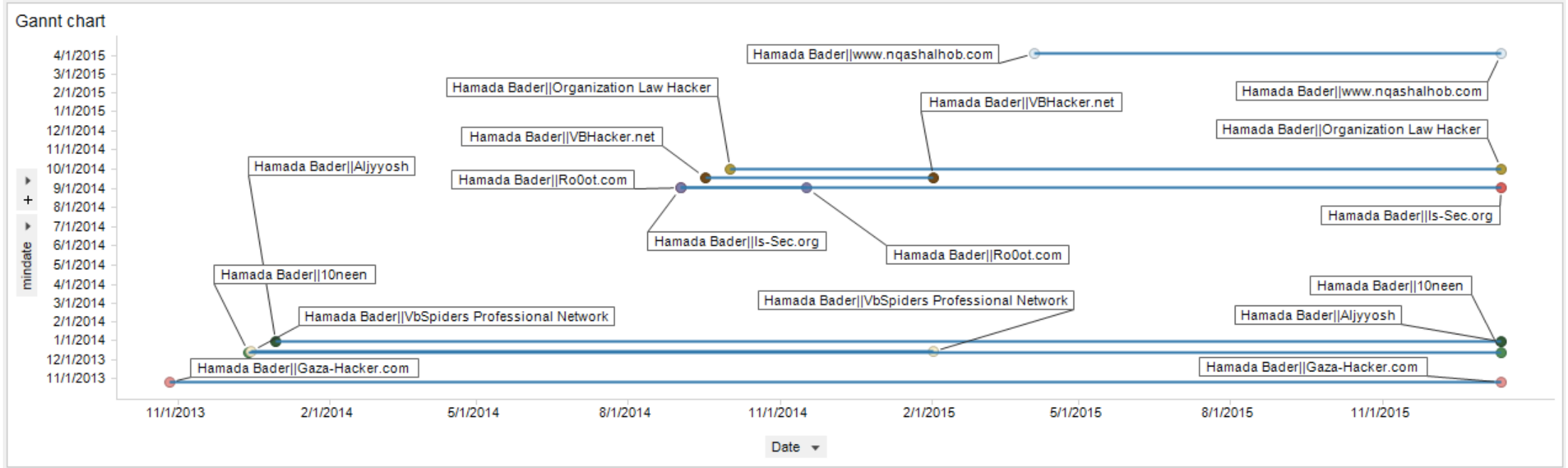
a and *b* are perfectly mirrored, which means they have the maximum negative correlation.



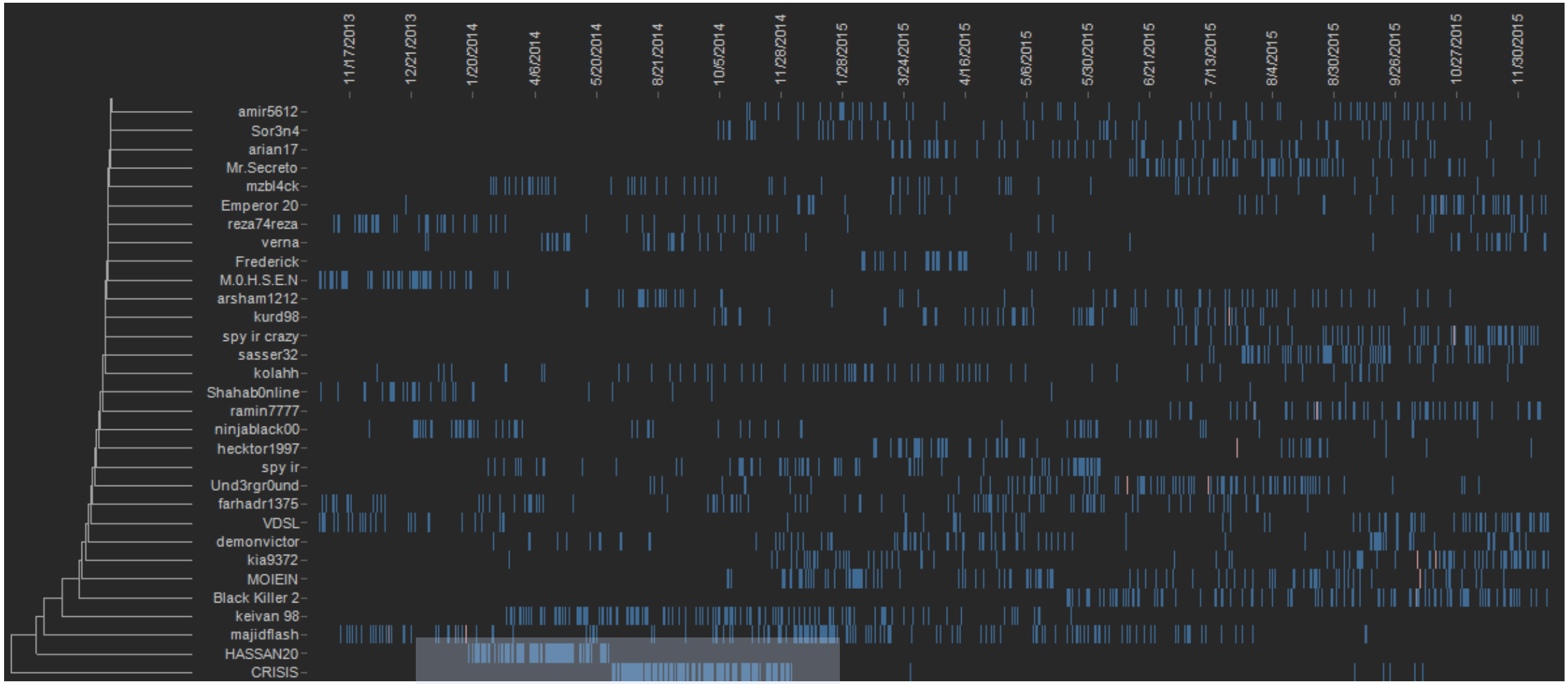


Sample: Hamada Bader (non malicious, selling IT products)

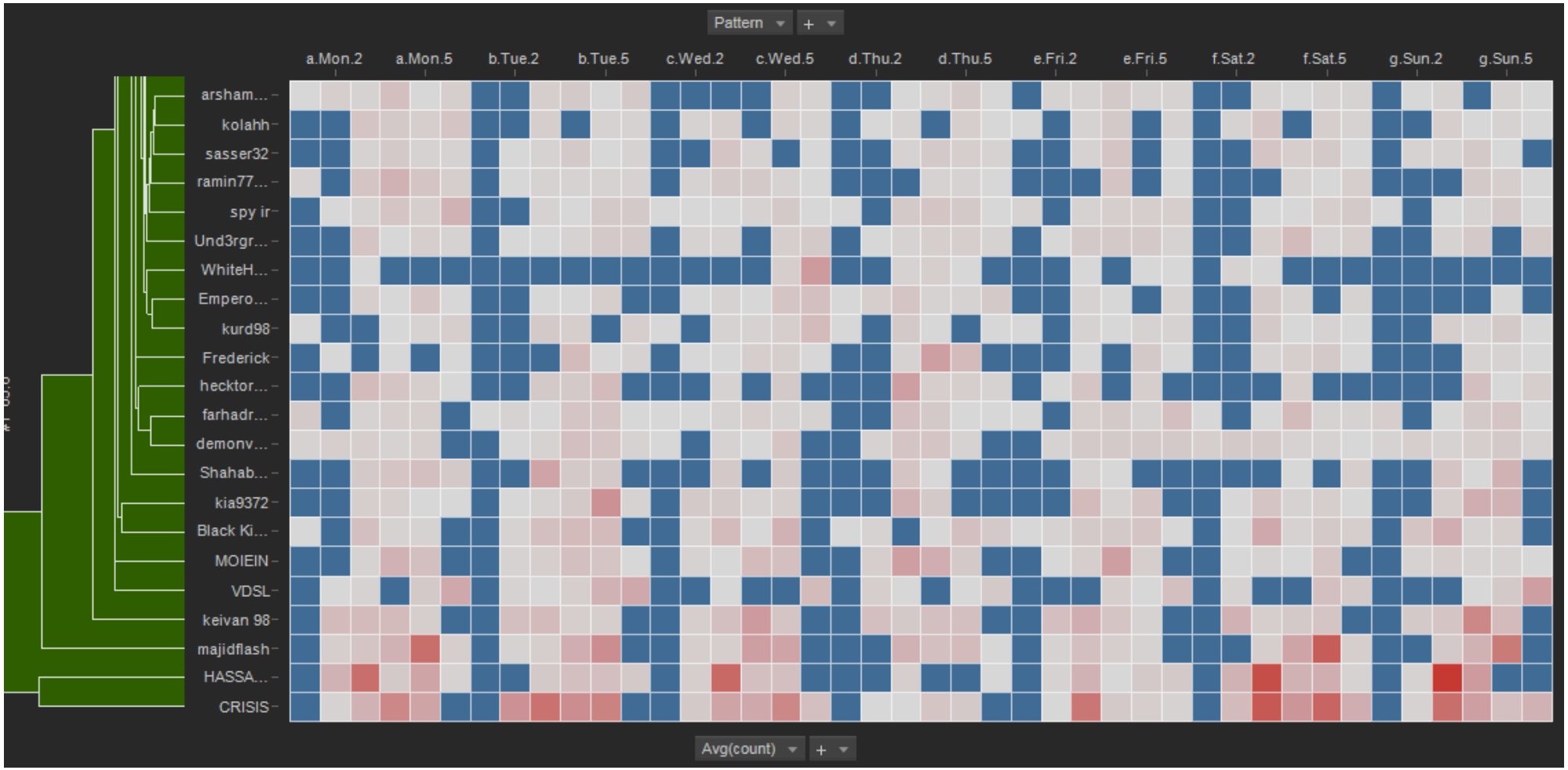




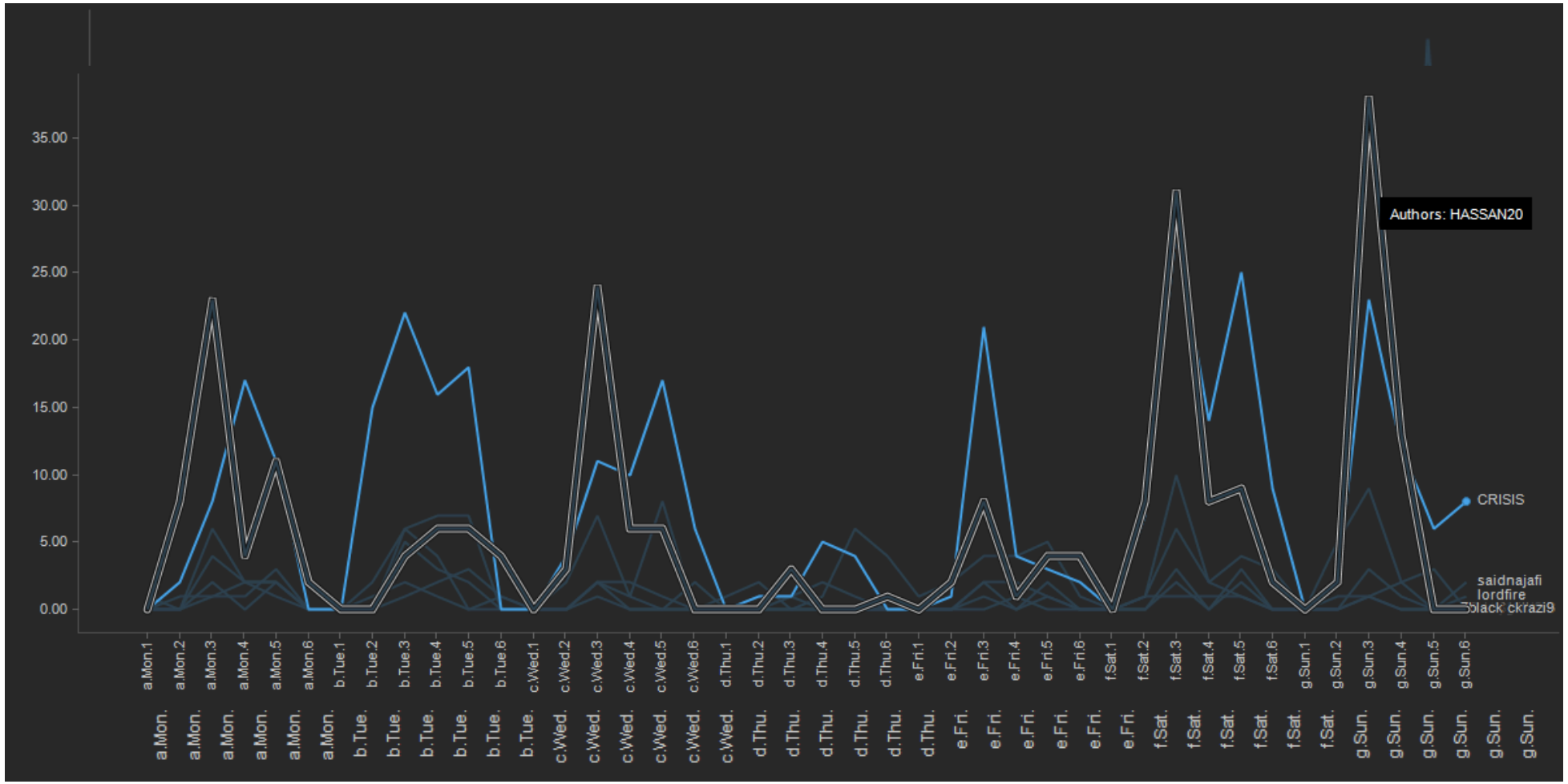
First Cut: HASSAN20 = CRISIS?



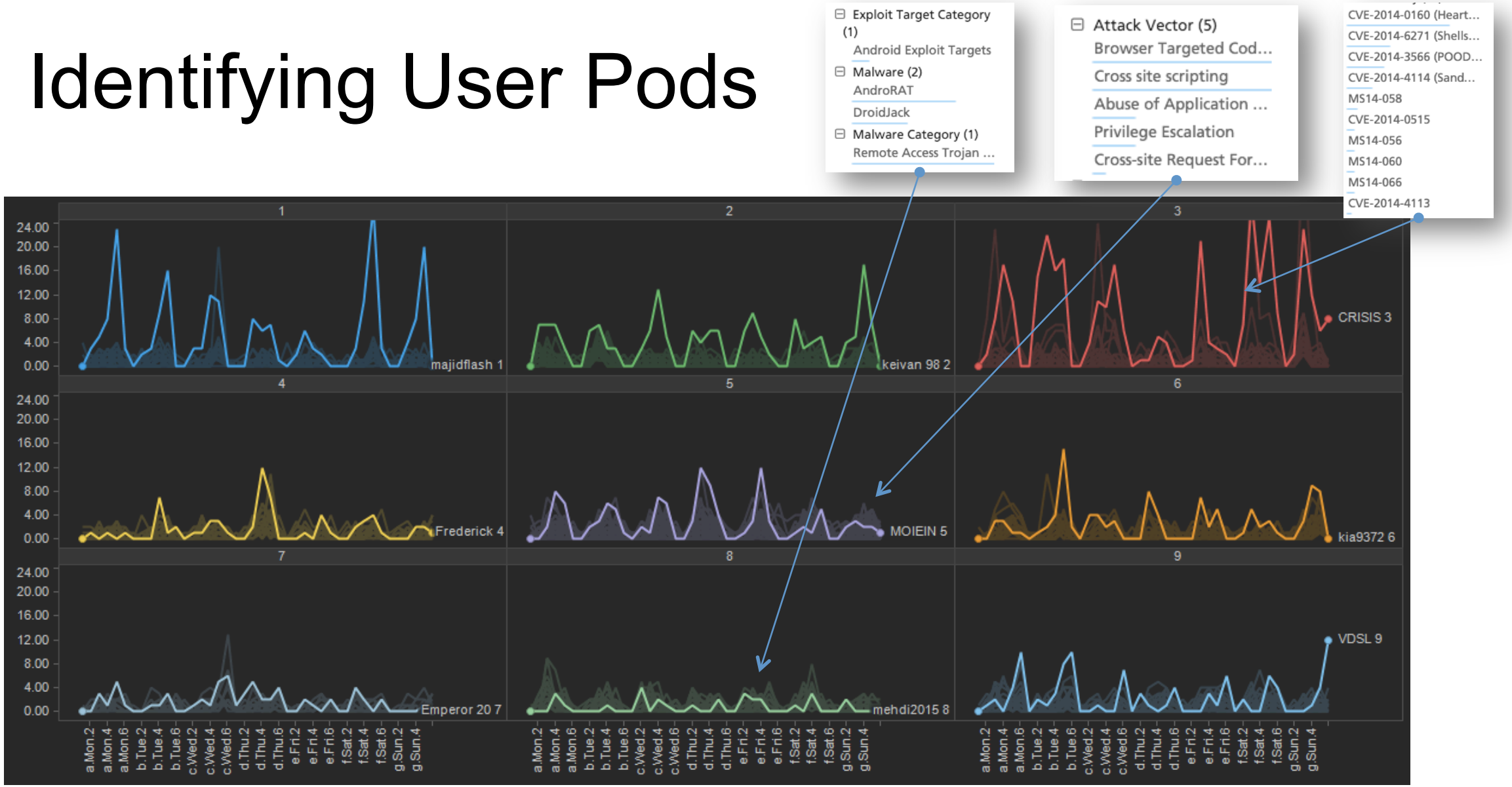
Clustering Pattern of Life



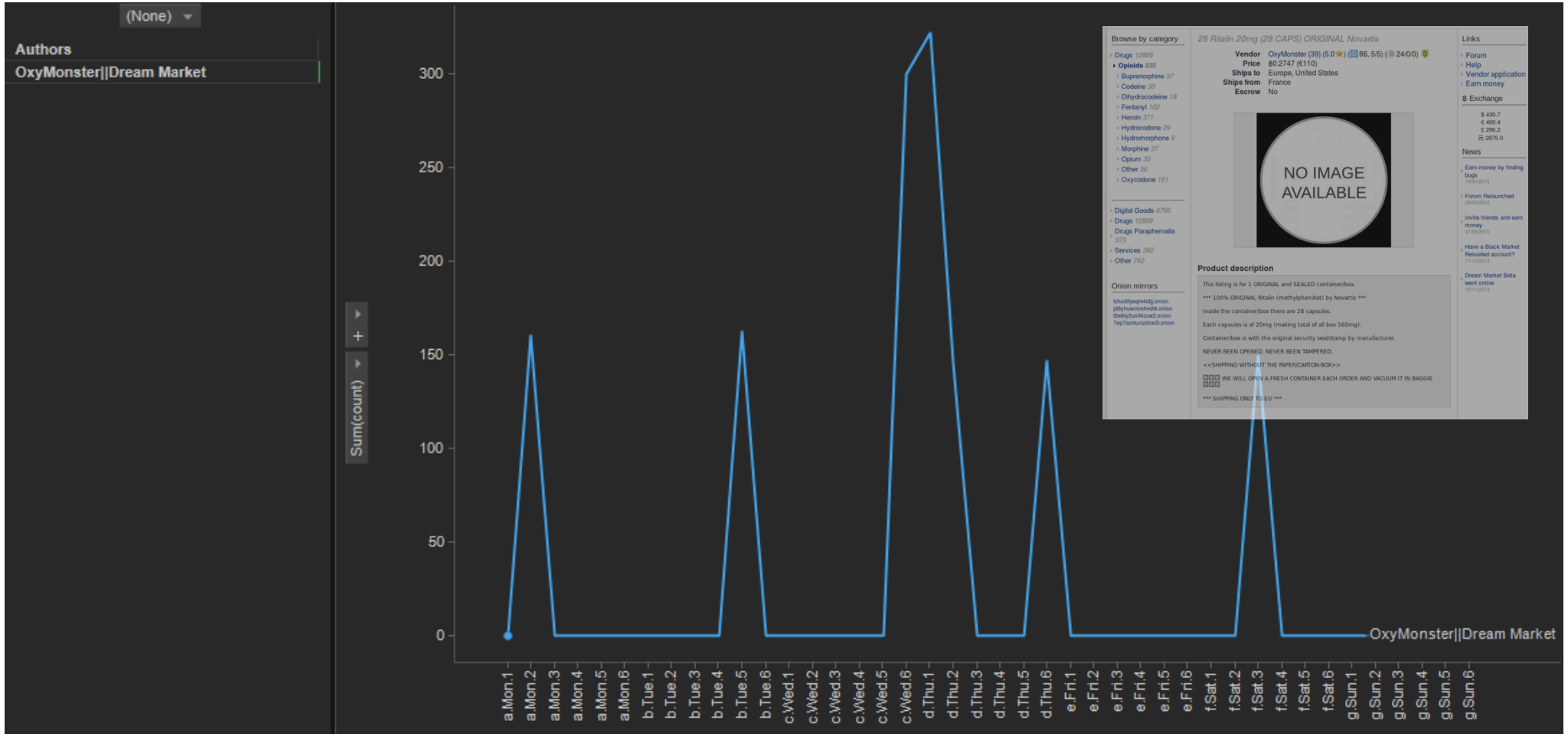
Identifying Hang Arounds



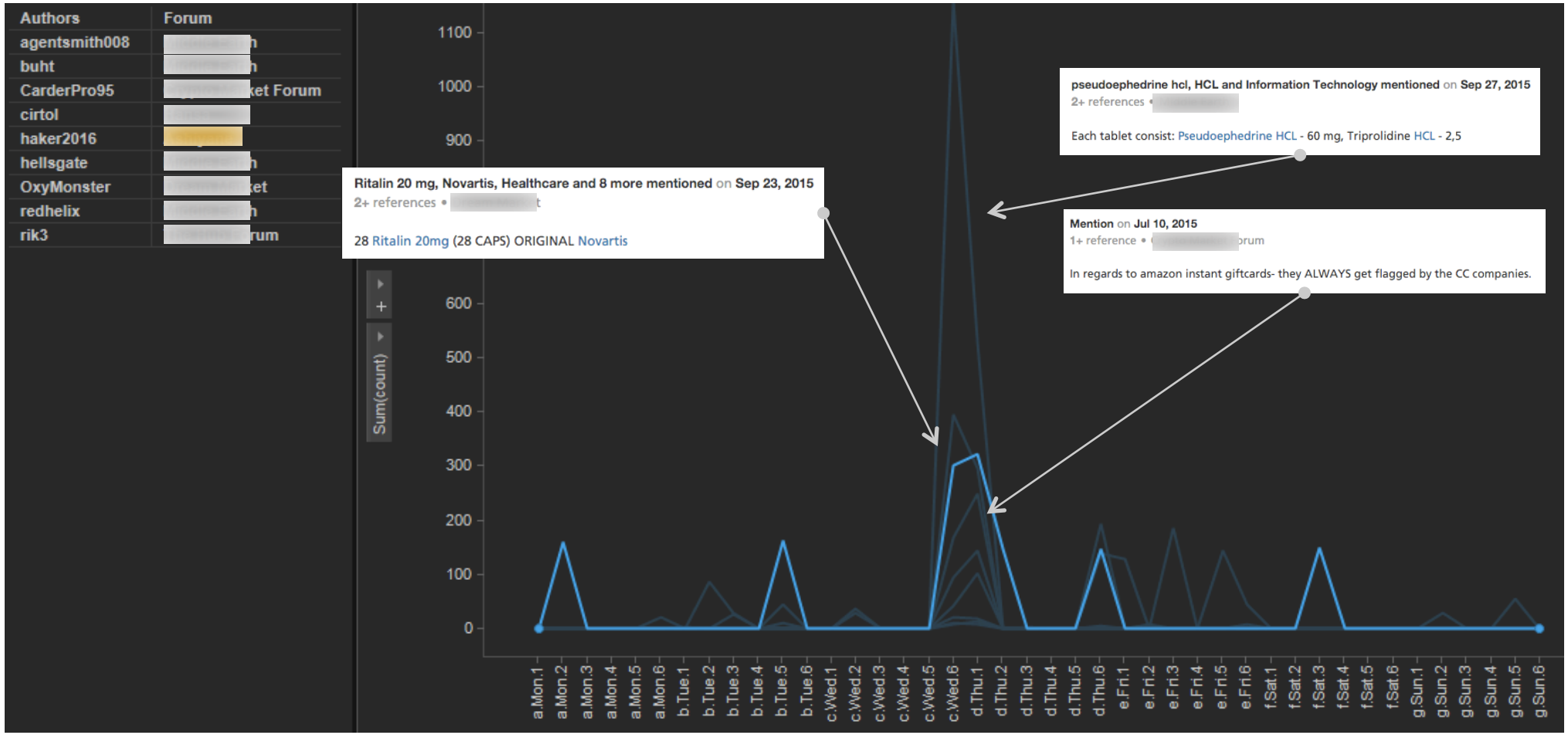
Identifying User Pods



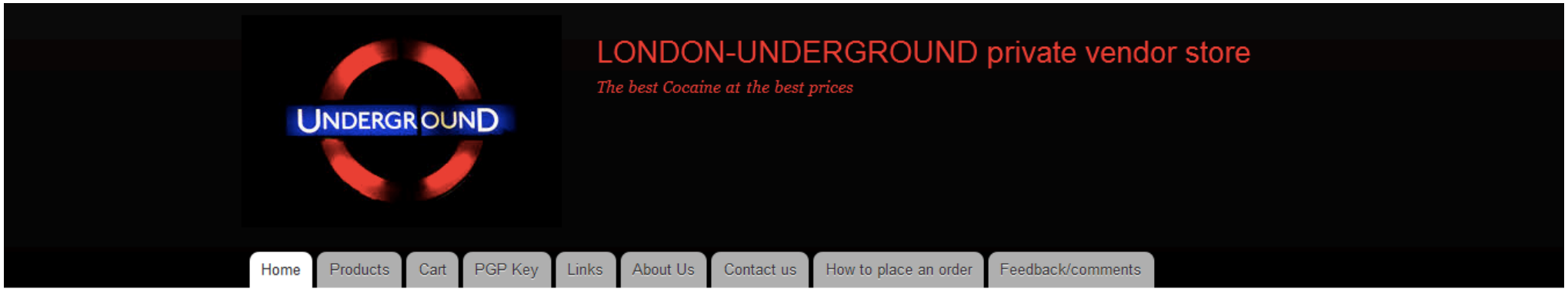
Making the Wednesday Night .Onion Rounds



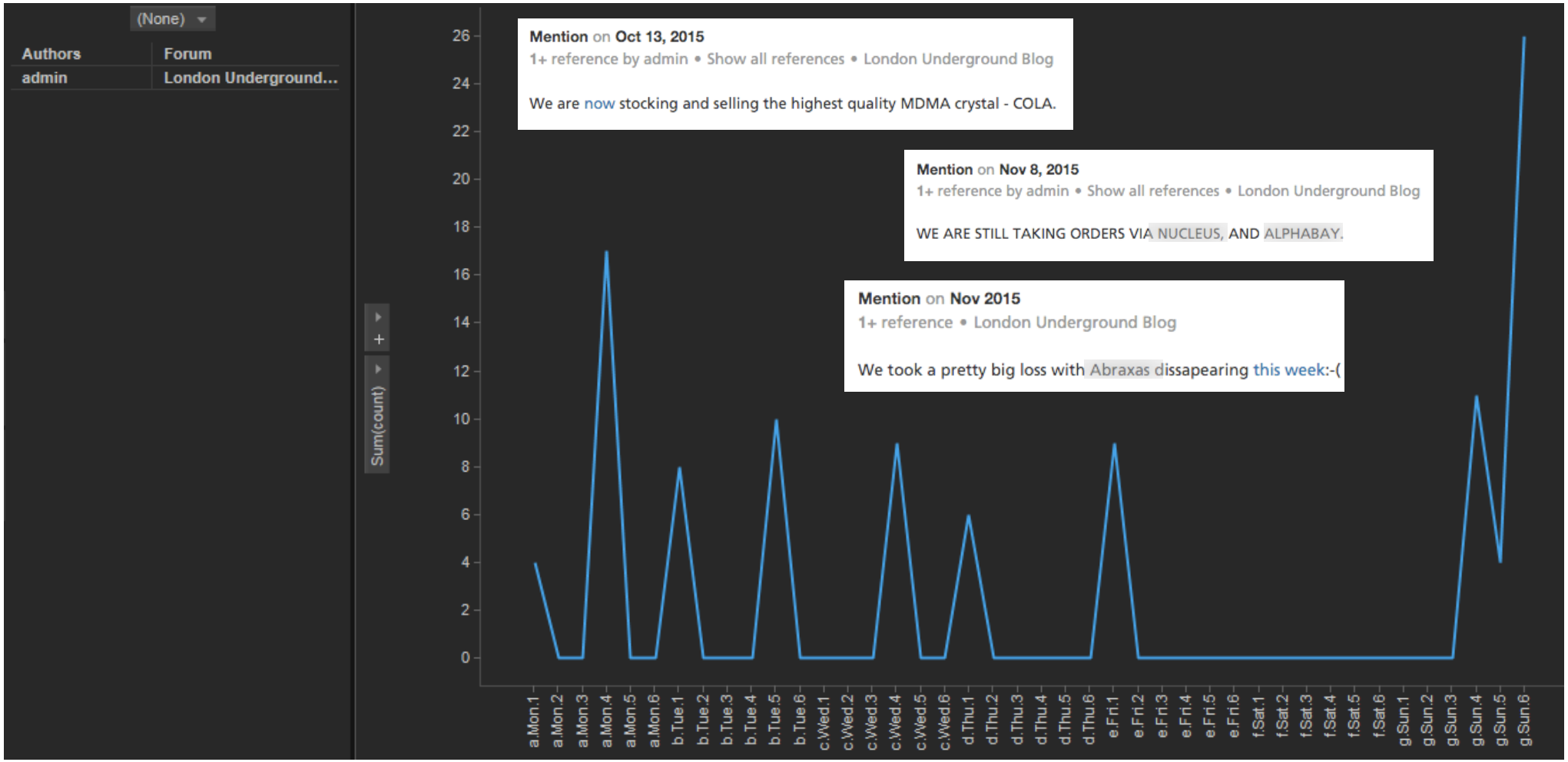
Making the Wednesday Night .Onion Rounds



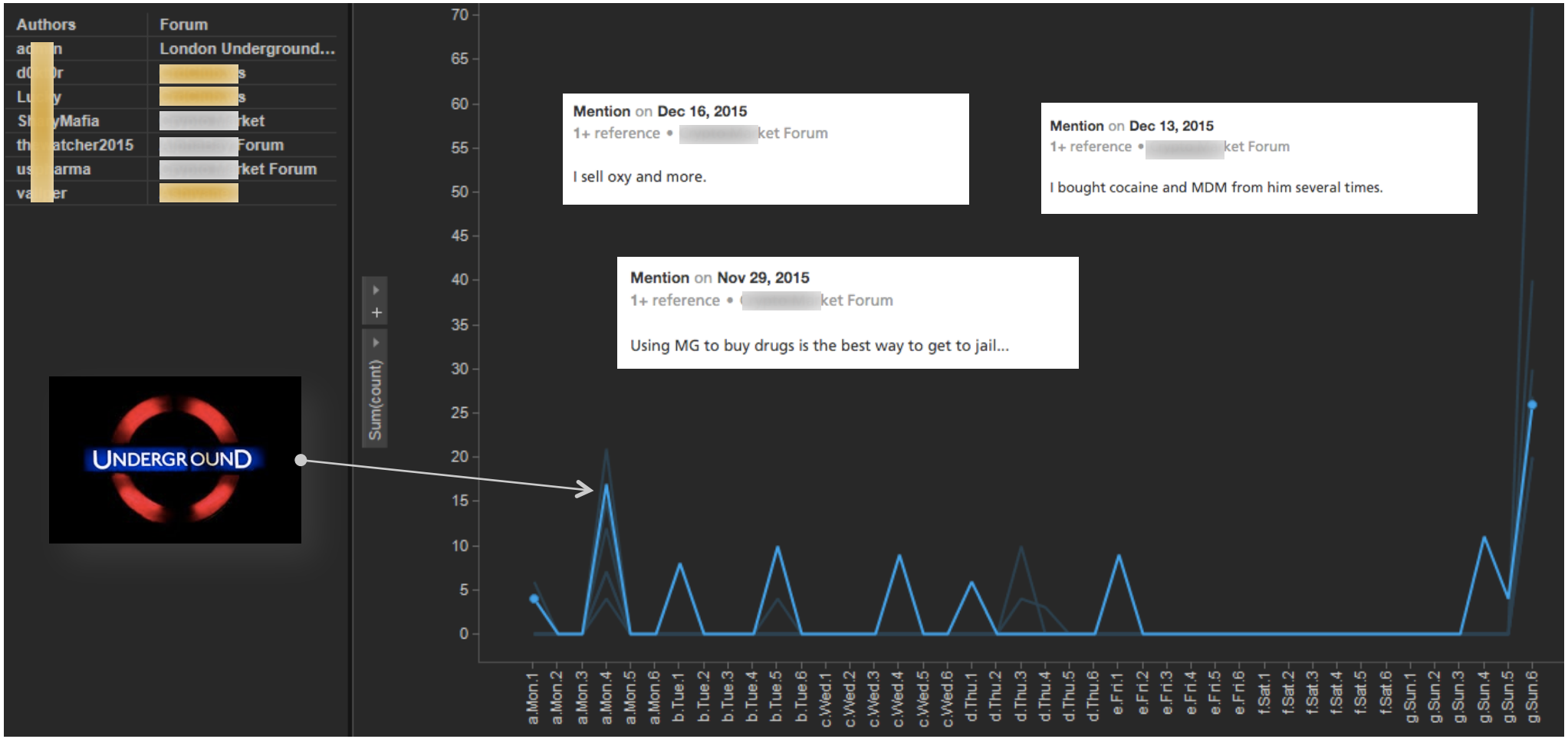
Profiling the Store Admin



Profiling the Store Admin



Profiling the Store Admin



Summary and What's Next?

- Pattern of life on metadata supports attribution
- Lead generation [duh!] requiring secondary sourcing

- Build hard analytic: "find handles like this"
- Cross validation on secondary data sources

- *Christopher Ahlberg, Co-Founder & CEO*
- *Recorded Future*
- *c@recordedfuture.com @cahlberg*