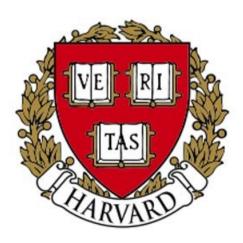
Knowing Where and How Criminal Organizations Operate Using Web Content

Michele Coscia¹ and Viridiana Rios²

1 KddLab - ISTI CNR - michele.coscia@isti.cnr.it 2 Department of Government - Harvard University - vrios@fas.harvard.edu





The Intelligence Problem

Information complexity critically affects the ability of security agencies to collect intelligence information by making it more costly



But they are often blurred

Criminal organizations usually have high interest in hiding their activities from the public

However...



drive. He had not gone its when he rain out and went up to a loory.

He space to the driver of 'Herica' which had a space to have a space to the property of the space of the property of the space of t

and told me not to go down the drive as he had alone a man and told me not to go down the drive as he had alone a man and told me not to go down the drive as he had alone a man and told me not to go down the drive as he had alone a man and told me not go down the drive as the had alone a man and told me not go down the drive as the had alone a man and told me not go down the drive as the had alone and told me not go down the drive as the had alone and told me not go down the drive as the had alone and told me not go down the drive as the had alone and told me not go down the drive as the had alone and told me not go down the drive as the had alone and told me not go down the drive as the had alone and told me not go down the drive as the had alone and to go down the drive as the had alone and to go down the drive as the had alone and to go down the drive as the had alone and to go down the drive as the had alone and to go down the drive as the had alone as the down the drive as the had alone as the down the drive as the had alone as the down the down the drive as the had alone as the down the drive as the had alone as the down the drive as the had alone as the down the three the down the down the three the down the down the three the down the three three the down the down the three three

wContinued from page 21 bourse-street, Iffichin, and the trigger of the beard revolver shous about 12 20 there were between four and it is about in rapid succession. Looking out minds with the property of the page of the p

to run concurrently.

POLICE SAW

(Conti

BELLOND their

would be surry

Ne also it me bought
lase Watters.

Mr. Birkett

Mr. Birkett

Tour mind at
body any harm

see him, and
was going to d

at, the nouse,

ion at ametion.

nd S(n. With a 15 perprice was \$1.84

to two life terms plus additional time. The federal prison system offers no chance of parole. He was convicted of planning what he claimed would have been a massive attack on a Texas restaurant filled with troops from Fort Hood.

In court, Abdo referred to Mai, Nidal Hasan - the Army psychiatrist soon to be tried in a deadly shooting rampage

cally involved with her ness."

Robicheaux laments sents families like the Abelards, who are part of what he describes as a bution

Arguit for a life sentence. Frazier had said bdo still presented a threat. Abdo's nuth was covered in court, Frazier ad, because he had earlier spat his on blood at agents believing he was jected with HIV. That belief turned outo be wrong.

Abdo is AWOL from Fort Campbell, Ky., whene was arrested with bombmaking nterials last summer at a Fort

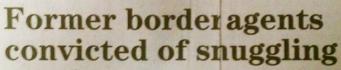
mentor as she tries to tertwines with a case of major corn

the loss of Louisiana's al sensibilities (which he share natural beauty and he re-with author Burke), he ultimately confronts the cast of evil villain

Louisiana gulf coast and has worked "The availability of as an oil company landman, news skid row social worker — a breadtl of experience that no doubt helped most fertile in the him create the array of character

> But his conscienge was put on hold af with possessing cha May 2011. Abdo torke the pornography acid only because he haak

"I just can't ima being placed on a that charge.



Associated Press

SAN DIEGO - Two former Border Patrol agents were found guilty Friday of smuggling hundreds of people into the U.S. in Border Patrol vehicles.

Raul and Fidel Villarreal were convicted of charges that they brought illegal immigrants into the U.S. for money and received bribes by public

poles, plated undercover recording devices, ut tracking instruments on Border Prol vehicles and followed a smugglin oad by airplane.

The pisecution also relied on accounts of leged accomplices and migrants whentered the country illegally. includinsome who identified Fidel Villarrean photographs. One 24-yearold Brazian woman said she paid



Former Border P

Moreover...

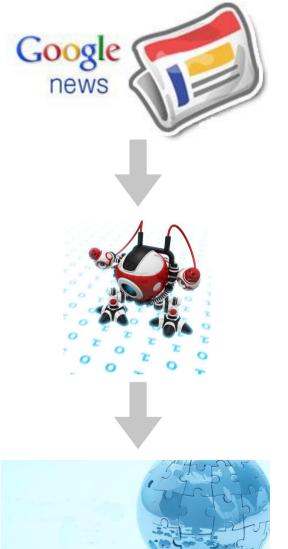




Problems

- The Web's knowledge is:
 - Too large to be analyzed as a whole;
 - Subject to reliability concerns.

Our Approach



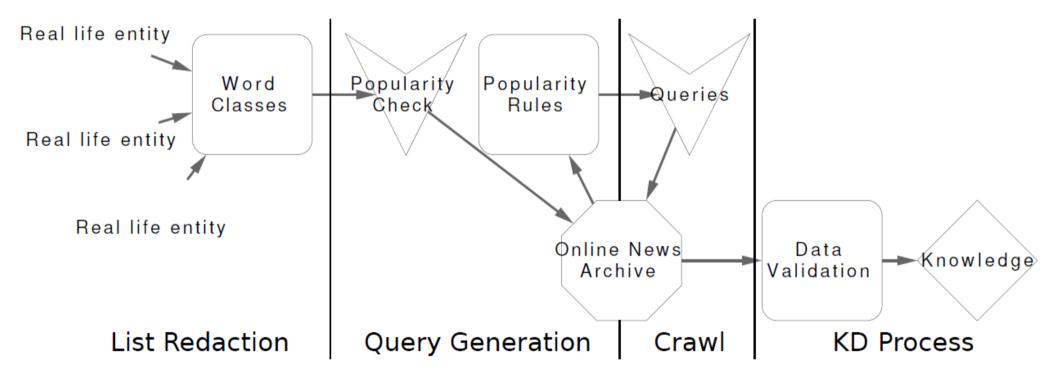
Already deals with data selection and cleaning to index only reliable and relevant news

We design an intelligence-oriented, topic-specific automatic query and data cleaning process



We derive an operative picture of the world and we test it against ground truth

MOGO: Making Order using Google as an Oracle



Entity Definition Rules

- Firstly define the "actor lists"
 - In our case: drug traffickers' names and municipalities
- Classify words into:
 - Generic
 - Common
 - Unique
 - State Unique

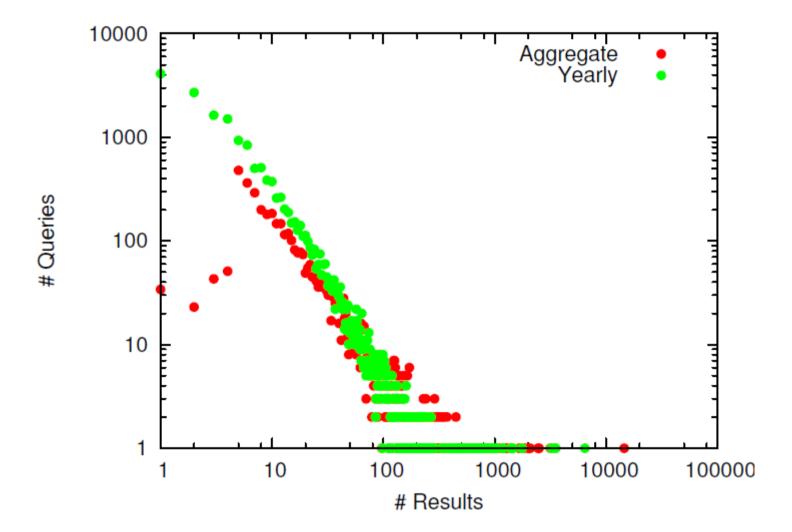
First Query Generation Step

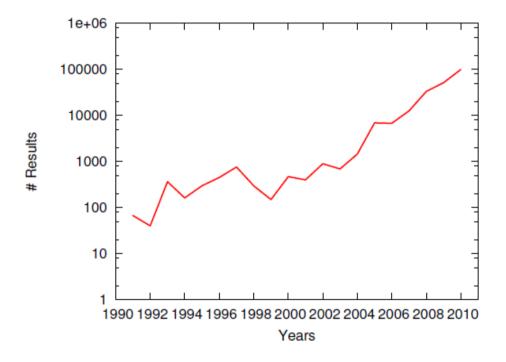
- A query is either:
 - Unique word by itself (because unambiguous)
 - State Unique + State Name (it makes it unambiguous)
 - Prune out Generic and Commons
- Other rules to avoid ambiguity (i.e. a municipality with the same name of a state)

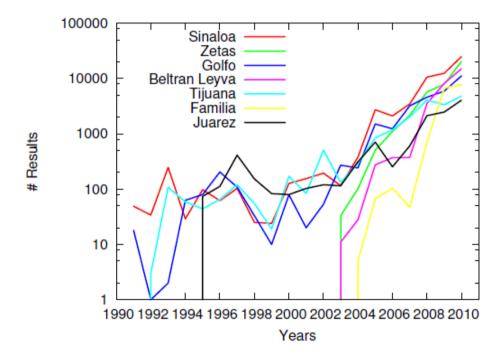
The Popularity Check

- Pruning out clear outliers that escaped the query generation step
 - Exaggerate popularity
 - Extremely high deviation

Data Statistics







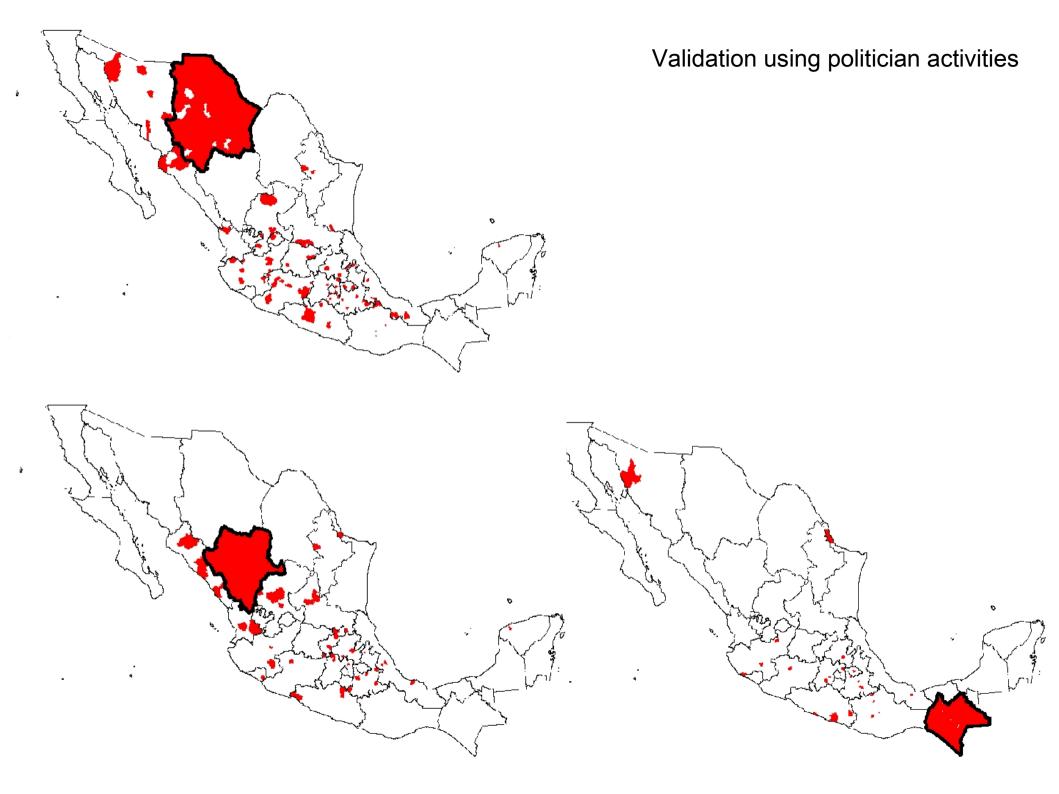
Data Cleaning & Validation

Cleaning

- Just because a trafficker co-appears a lot with a municipality, it does not mean that the connection is strong
- We validate the results against a null model

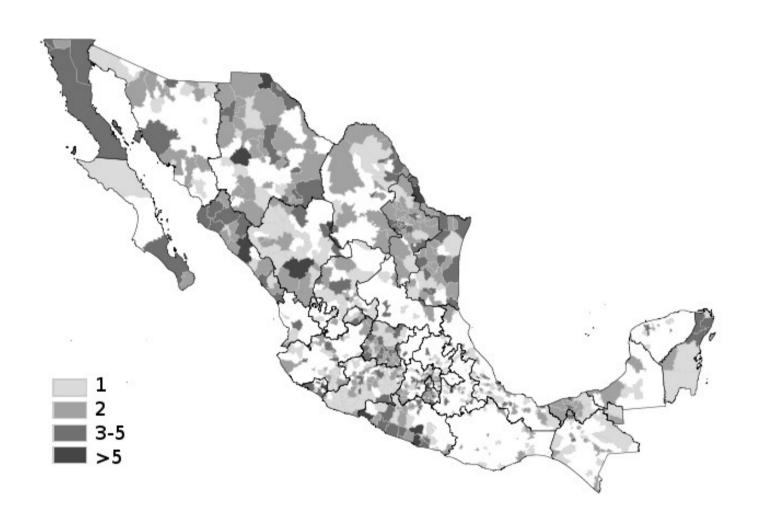
$$PMF(t_{i} = \bar{t_{i}}) = \frac{\begin{pmatrix} \mathcal{T}_{i} \\ \bar{t_{i}} \end{pmatrix} \begin{pmatrix} \mathcal{M} - \mathcal{T}_{i} \\ \bar{m_{i}} - \bar{t_{i}} \end{pmatrix}}{\begin{pmatrix} \mathcal{M} \\ \bar{m_{i}} \end{pmatrix}}$$

$$CDF(t_i = \bar{t_i}) = \sum_{i} PMF(t_i = a)$$

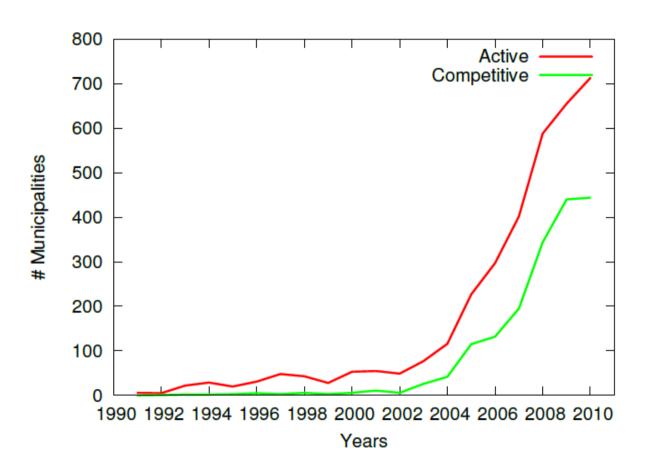


Results

Drug-Active Municipalities

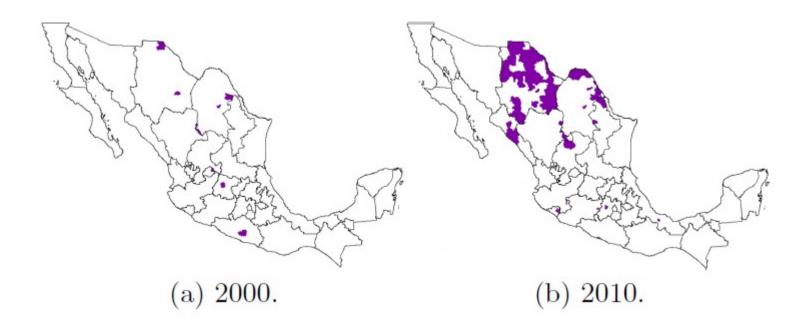


Competition Among Cartels

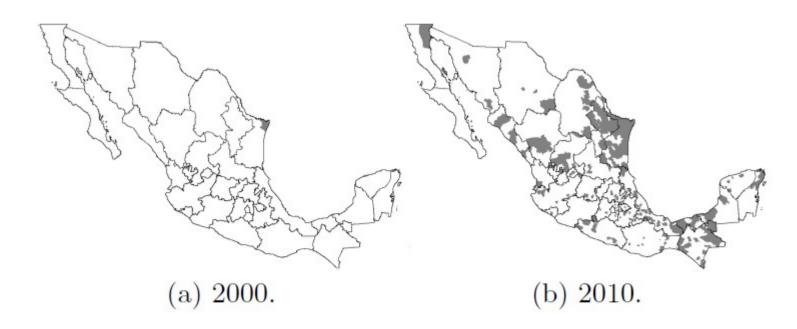


Drug Cartels Classification

Trafficking Org	2010 Territories	Start Year	Territories	Abandoned	Years operated	Competitive	Exploratory	k
Juárez	74	1997	13.85	10.15	2.78	-0.67	-0.03	1
Tijuana	39	1997	10.1	8.15	2.74	-0.96	-0.21	1
Sinaloa	176	1993	25.6	16.95	2.84	0.18	0.64	1
Barbie	66	2006	5.75	2.45	1.63	-0.48	-0.72	2
Mana	32	2006	3.8	2.2	2.15	-0.82	-0.73	2
Sinaloa faction	53	2008	5.15	2.5	1.96	-0.67	-0.70	2
BL faction	57	2008	5	2.15	1.79	-0.52	-0.75	2
Other	24	2008	2.15	0.95	1.38	-0.99	-0.75	2
BL	157	2004	18.65	10.8	2.08	0.81	-0.36	3
Fam	227	2005	18.75	7.4	2.09	0.95	0.01	3
Golfo	244	1994	35.55	23.5	3.01	1.25	1.07	4
Zetas	405	2003	42.2	21.95	2.71	1.94	2.55	4



The migration pattern for Juarez cartel.



The migration pattern for Zetas cartel.

Thank You

Questions?

