

Production de renseignement criminel à partir de messages électroniques frauduleux

Nicolas Seriot

XIIe colloque de l'Association Internationale des
Criminologues de Langue Française

14 mai 2010
Université de Fribourg

Présentation



- Nicolas Seriot, Lausanne
- Ingénieur HES en informatique logiciel
- Développeur iPhone / iPad, Swissquote Bank
- MAS en Lutte contre la criminalité économique, 2008 ILCE Neuchâtel

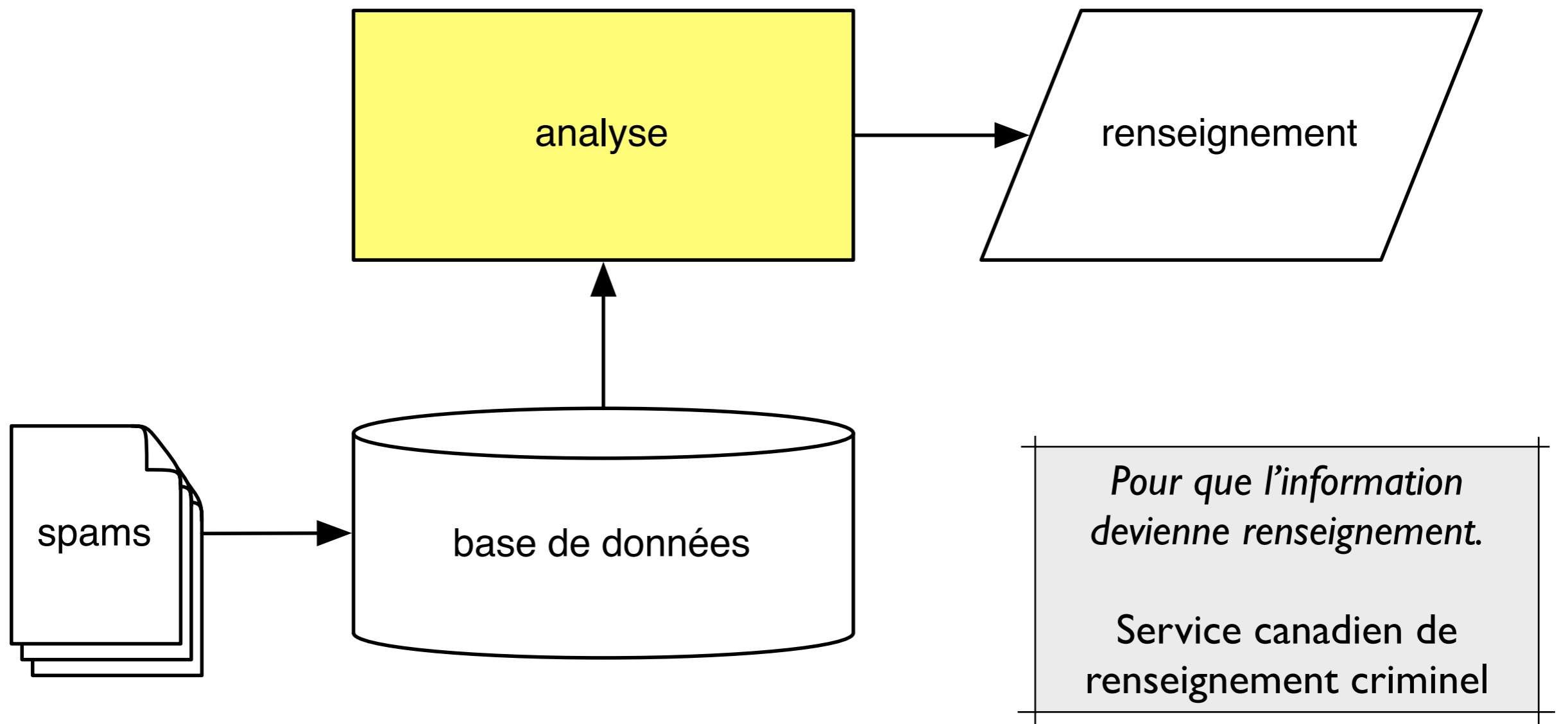
Contexte et motivation

- Projet de recherche appliquée : *Internet Surveillance for Criminal Intelligence Analysis*
- Problématique : comment extraire du **renseignement criminel** à partir d'une base de messages frauduleux ?
- Appliquer à la lutte contre la fraude des techniques issues d'autres disciplines

Hypothèses de recherche

- Les spams contiennent des **traces**, qui permettent de faire des **recoupements entre les cas**, comme les traces de semelle dans les cambriolages (Schiffer / Birrer / Cartier / Capt / Ribaux, 2004, 2007)
- Comment ? H: la technique du clustering hiérarchique ascendant est utile à l'analyste pour produire du renseignement criminel.

Mini ISCIA



I. Le spam et les traces

La vérité est que nul ne peut agir avec l'intensité que suppose l'action criminelle sans laisser des marques multiples de son passage

principe de Locard

La trace et un élément matériel, un témoin silencieux qui ne ment pas.

Alphonse Bertillon

```
Return-Path: <xxx@xxx.com>
Received: (qmail 25793 invoked by uid 64); 6 Oct 2008 02:16:12 -0000
Received: from xxx@xxx.com by alinto.net; 06 Oct 2008 02:16:12 -0000
Received: from unknown (85.218.45.86)
  by smtp.alinto.citycable.ch with SMTP; 6 Oct 2008 02:16:04 -0000
From: Capitaine Haddock <haddock@gmail.com>
To: Tintin <tintin@gmail.com>
Return-Path: <milou@gmail.com>
Subject: Bonjour
Date: Fri, 10 May 1968 16:56:05 +0200
X-aduser: 85.218.45.86
X-Qmail-Scanner-Message-ID: <122325936967525773@alinto.net>
```

Bonjour, moussaillon !



2. Le spam et le droit

- Un message est-il lié à une activité criminelle ?
- Escroquerie : les auteurs ne sont pas d'accord entre eux (débat sur l'astuce)
- LTC et LCD plus affirmatives
- Encore d'autres biens protégés par la loi...



3. Le renseignement criminel

- **Intelligence led policing**

- approche préventive plutôt que réactive, basée sur le renseignement

- **Analyse criminelle**

- collecte d'information, analyse, dissémination

- **Renseignement**

- utile à la compréhension et l'action



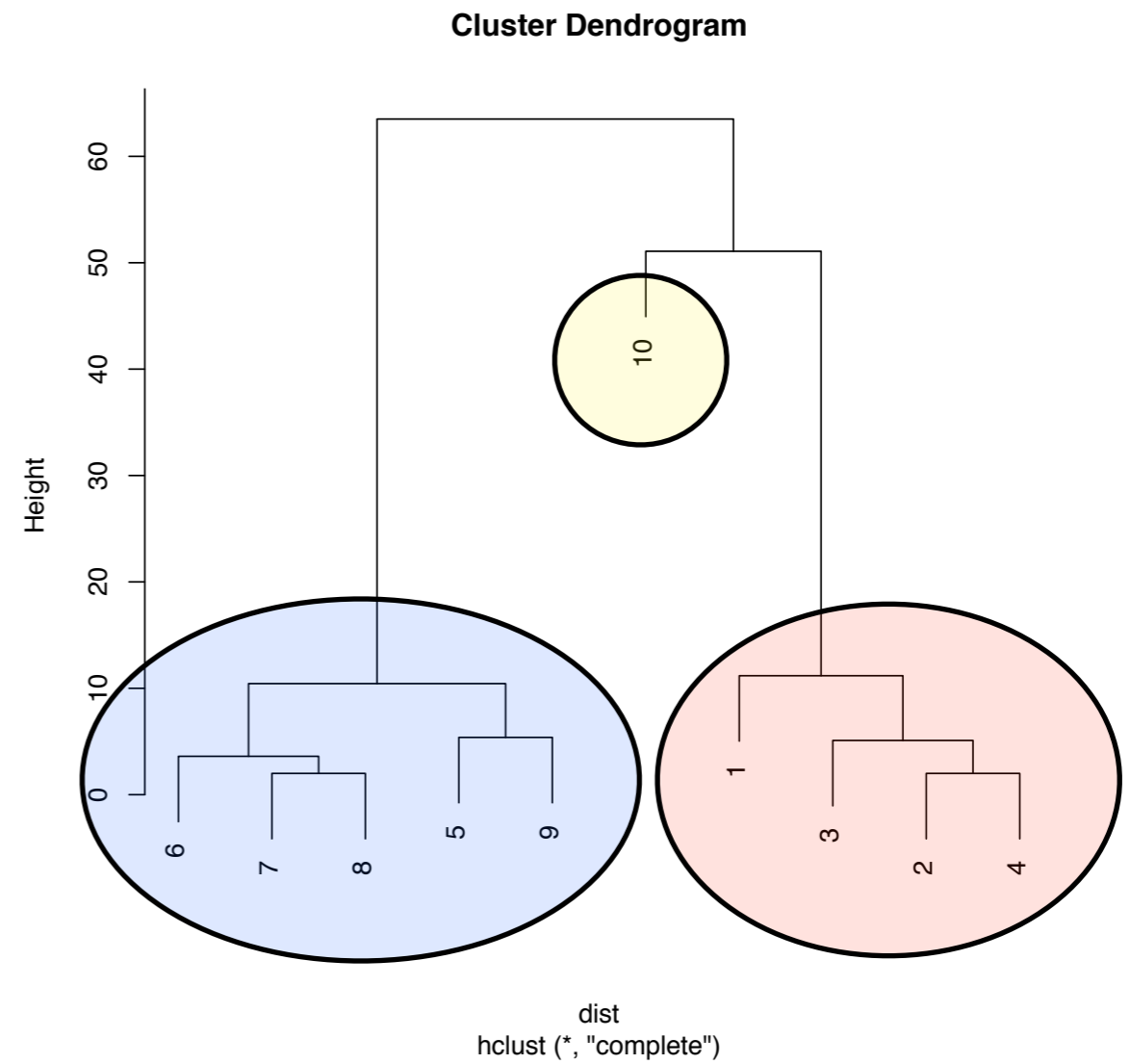
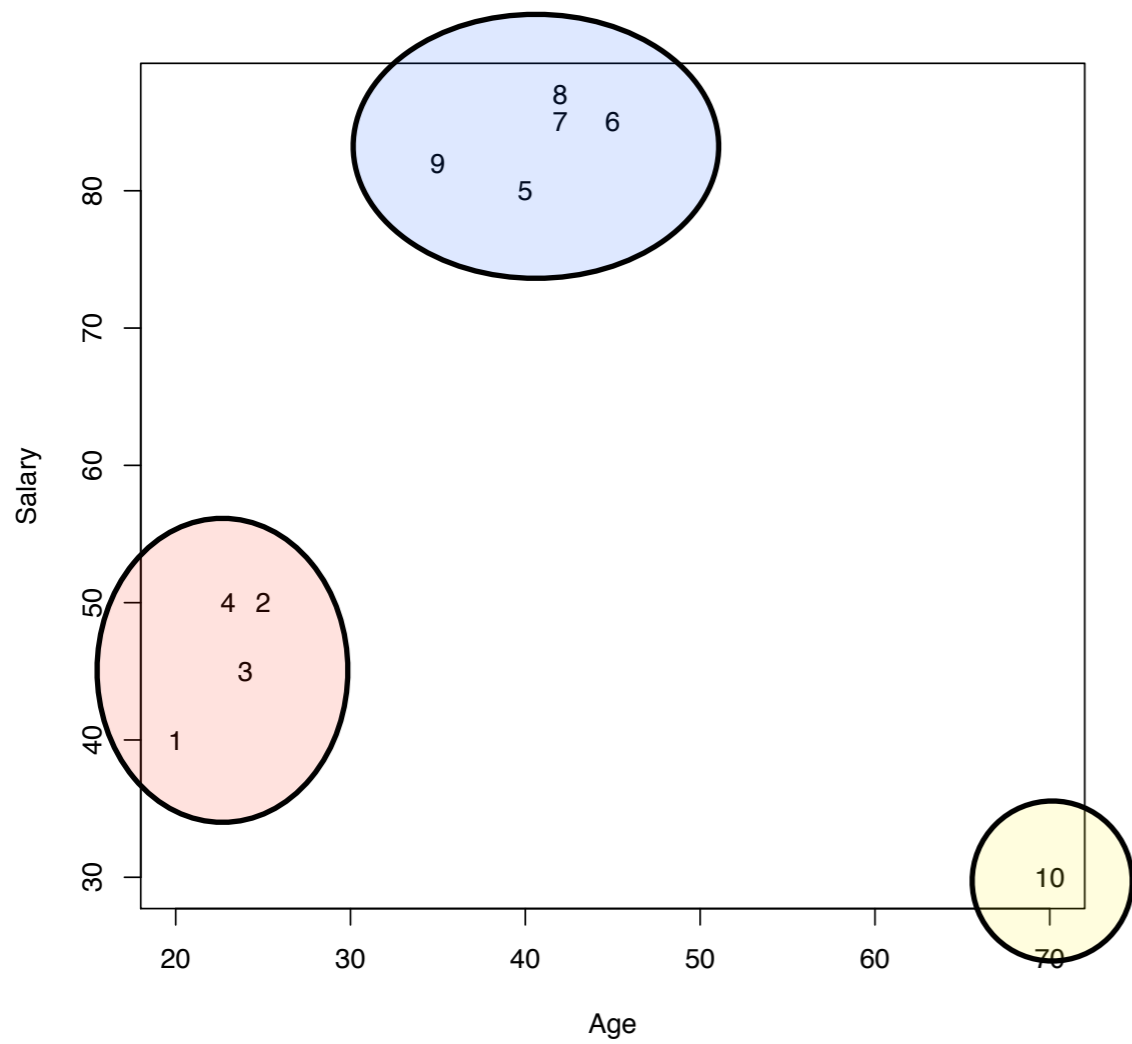
4. Data Mining

We are drowning in information, but starving for knowledge.

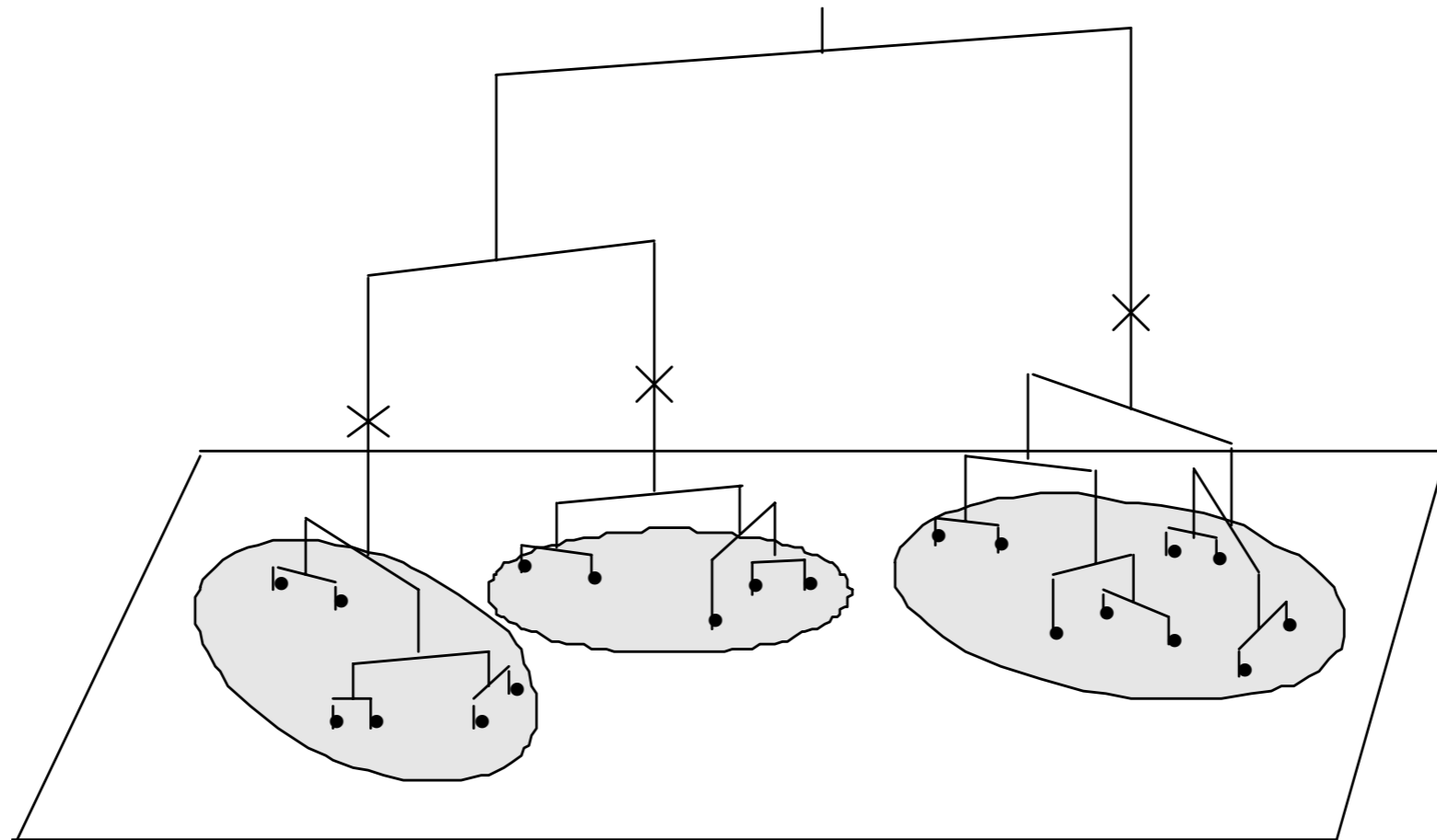
John Naisbett, dans son livre Megatrends

- techniques permettant l'**extraction d'informations** pertinentes dans de grandes bases de données
- marketing, biologie, analyse de risques, ...
- **clustering** détection de groupes homogènes bien distincts sur la base de différences entre les observations

Le clustering



Le clustering

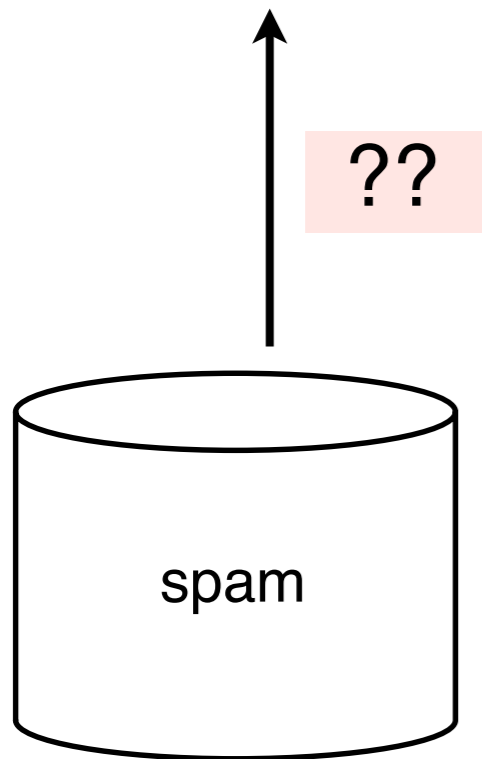
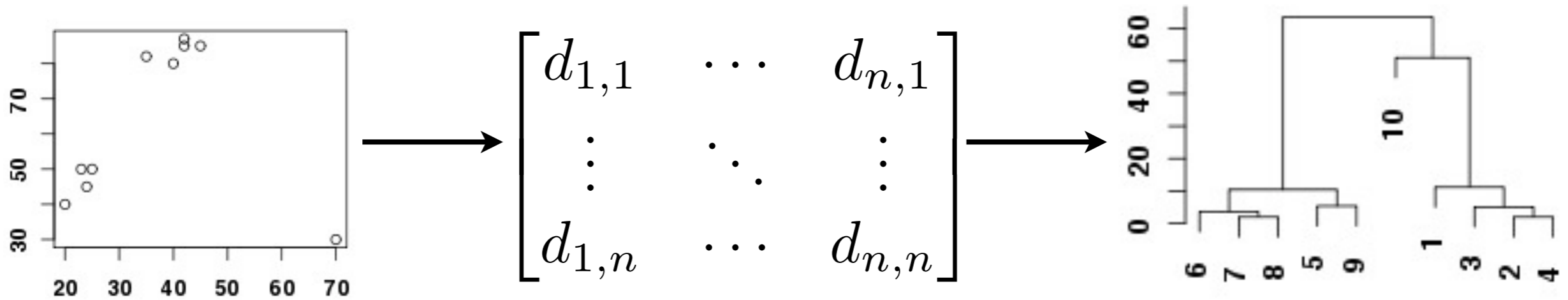


Le dendrogramme peut être vu comme un mobile
(Lebert et al. 1995)

Application du clustering

- **cadre conceptuel** de l'analyse criminel
- on ne cherche pas une recette de cuisine
- on cherche une **méthodologie** :
 - compréhensible
 - reproductible
 - communicable

Application du clustering



1. analyse multivariée

- plusieurs descripteurs
- types de variables
- normalisation
- poids relatifs
- peu de variables "sûres"
- subjectivité de l'analyste

Texte I

*** New Watches!! Year 2008 models ***

Many Year 2008 Latest Arrival new models Added

- : 100% Finest Quality Watches
- : Japanese & Swiss movement Replicas
 - : Same day shipping
 - : Ship by DHL, USPS, FedEx & UPS

View over 300 A-Z Brands on Our site

<http://ruuc.dyhigh.cn>

Texte 2

*** NEW WATCHES!! Year 2008 Models ***
170 year 2008 Latest Arrival new models Added

- : 100% Finest quality watches
- : Japanese & Swiss Movement replicas
 - : Fast Shipping
 - : Ship by FedEx, UPS, DHL and USPS

View A-Z Brands on Our site

<http://rfl.putrig.cn>

Texte 3

Hi,

We Sell exactly replicated SwissWatches with over 1 million
Units
sold!

Hurry before Inventory all Gone!

You can View our large selection of BreitlingRolexes, Gucci,
Cartier, Tag Heuer etc

<http://rkbl.fabring.cn>

Best Regards,

Darren

CEO/President

Director of SwissWatches Company

Texte 4

It cost You Nothing (Yes! \$0) to Give us a call, we will Contact you back

No Exams/Books/Tests/Classes/interview
100% No Pre-School Qualification Required!

Inside USA: 1-718-989-5740
Outside USA: +1-718-989-5740

Degree, bacheelor, MasteerMBA, PhDD Available in the field of your choice So you can even become a doctor and Receive all the benefits that Comes with it!

Please leave below 3 info in voicemail:

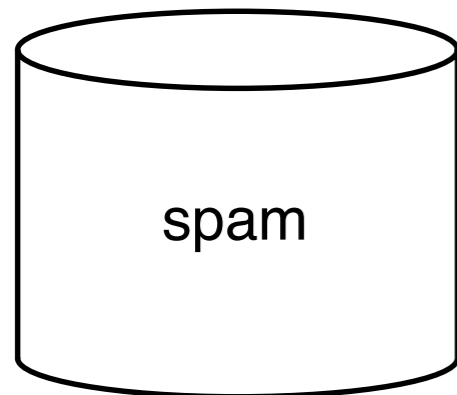
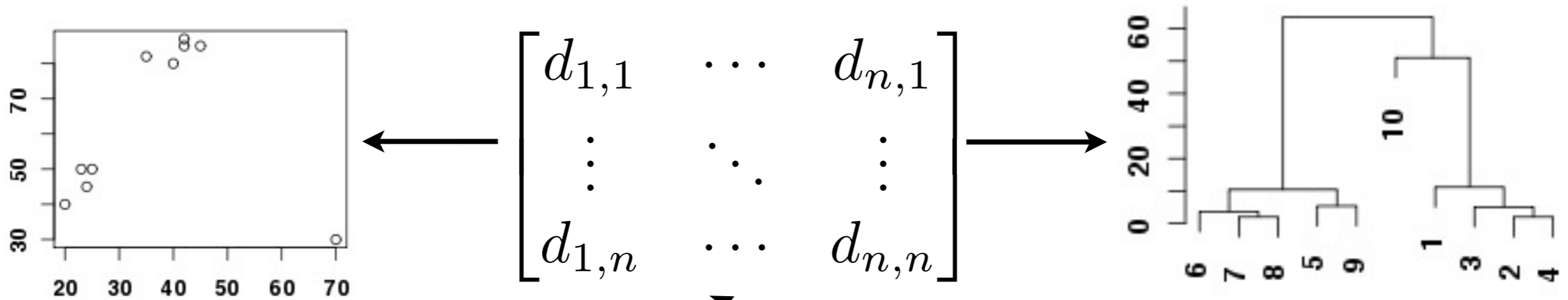
- 1) Your name
- 2) Your country
- 3) Your phone no. [Please Include Countrycode]

Call Now!! 24-hours a Day, 7-Days a week waiting For your call

Inside USA: 1-718-989-5740
Outside USA: +1-718-989-5740

Our staff Will get back To you in 1-3 working days

Application du clustering



2. distance entre les textes

- analyse univariée !
- neutralité
- facilité d'application
- proche de l'intuition humaine
- mieux adapté à l'analyse criminelle

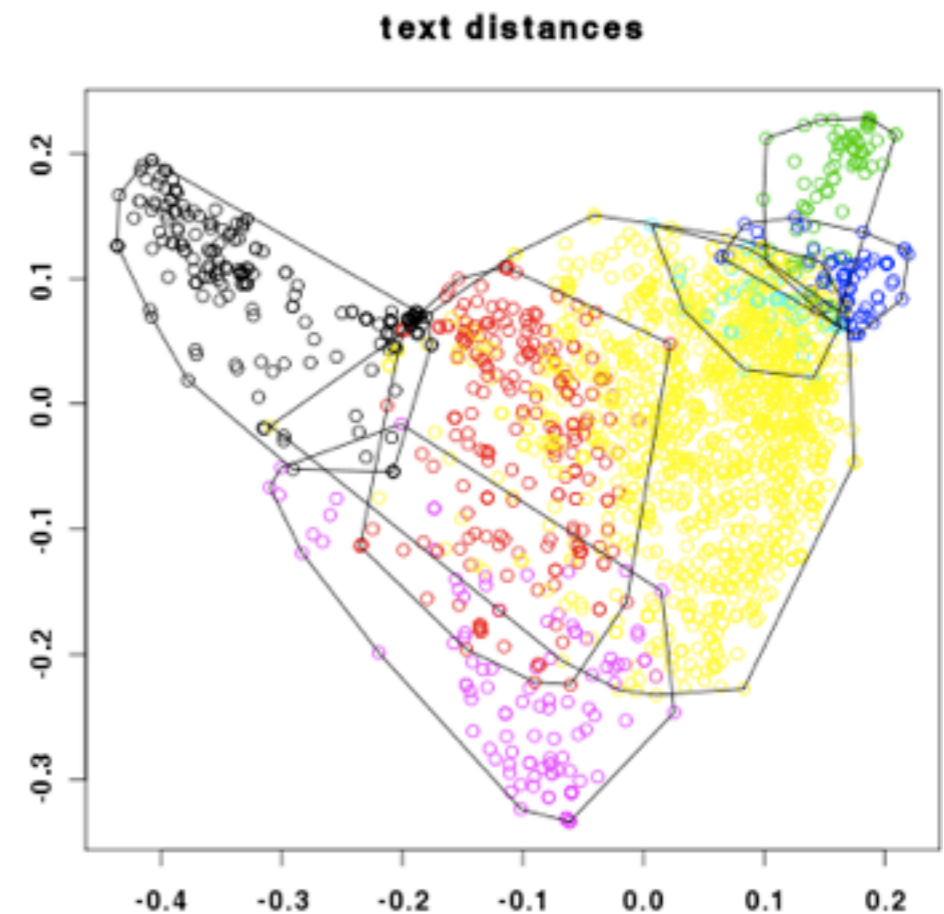
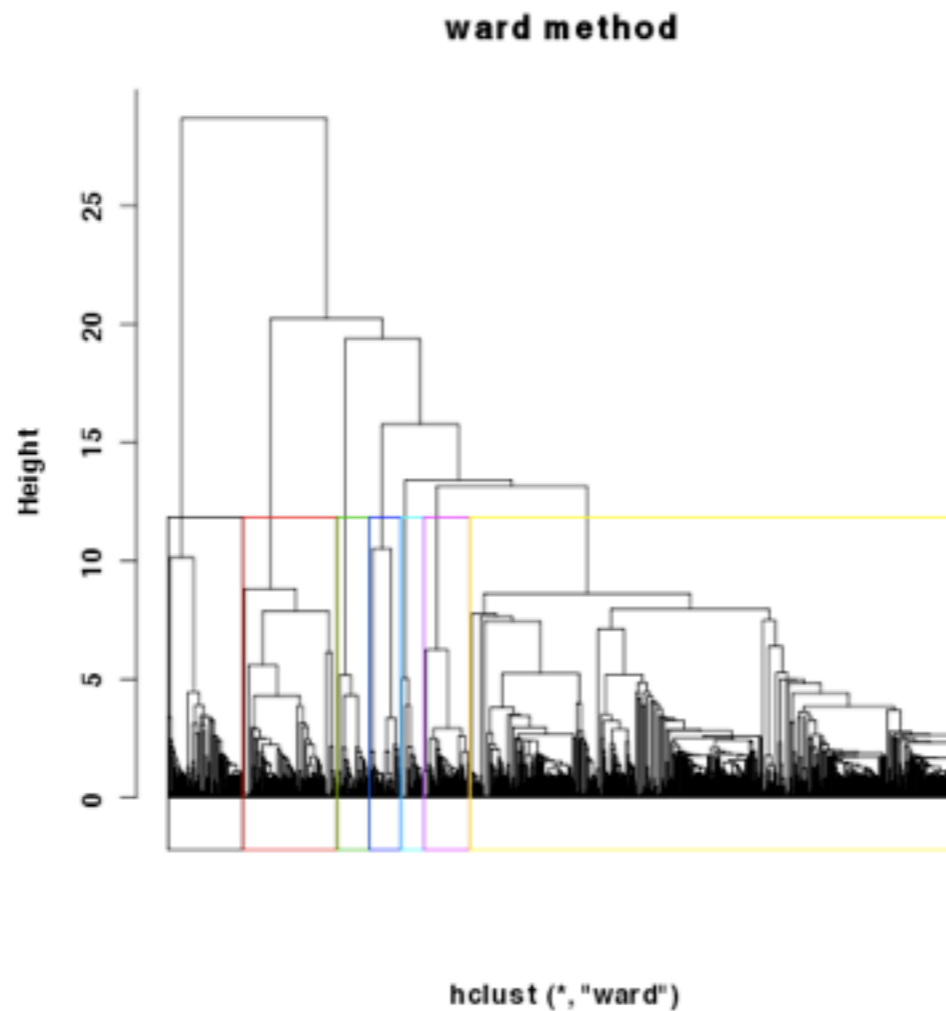
Prototype logiciel

Email Scams Intelligence Tool

[home](#)
[emails](#)
[addresses](#)
[countries](#)
[search](#)
[clusters](#)

[importer](#)
[admin](#)

1950 emails, 7 clusters.



[Cluster 1](#), 185 emails
[Cluster 2](#), 232 emails
[Cluster 3](#), 80 emails
[Cluster 4](#), 77 emails
[Cluster 5](#), 57 emails
[Cluster 6](#), 112 emails
[Cluster 7](#), 1207 emails

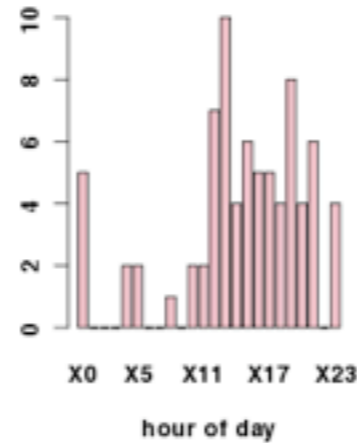
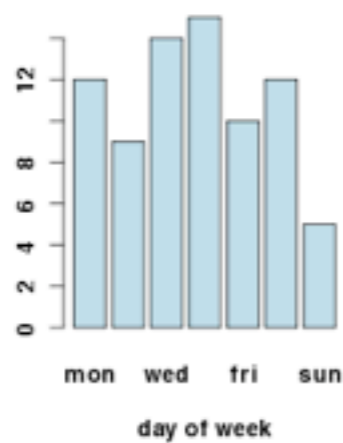
Prototype logiciel

Email Scams Intelligence Tool

- [home](#)
- [emails](#)
- [addresses](#)
- [countries](#)
- [search](#)
- [clusters](#)
- [importer](#)
- [admin](#)

Email Group

77 emails in cluster number 4



<u>id</u>	<u>date</u>	<u>country</u>	<u>parts</u>	<u>subject</u>
1068	2008-03-21 14:45:23	Cote D'Ivoire	2	From Keven and Joy Kwasi,
607	2008-05-11 13:35:51		2	Re: From Synthia Anthony
624	2008-04-24 21:56:24	Nigeria	1	RE: FROM YOUR SISTER
674	2008-04-11 13:33:01	Cote D'Ivoire	2	Good Day
677	2008-04-25 14:05:28	Cote D'Ivoire	2	Good day
690	2008-05-04 16:58:33	Cote D'Ivoire	2	Greeting From Miss Souzane Kouassi
682	2008-05-03 12:48:29	Cote D'Ivoire	2	Good day
701	2008-05-01 13:25:26	Cote D'Ivoire	2	greetings to you
810	2008-05-10 23:07:27	Cote D'Ivoire	2	I am 22 years old deaf girl,
797	2008-05-01 17:59:36	Cote D'Ivoire	2	HOLA,
798	2008-05-01 17:59:01	Cote D'Ivoire	2	HOLA,
799	2008-05-03 13:31:08	Cote D'Ivoire	2	HOLA,
1393	2008-04-24 00:51:29	Cote D'Ivoire	2	Very urgent please.

Cas d'analyse

- le clustering peut s'appliquer aux spams...
- mais permet-il de produire du renseignement ?

Lorsqu'une personne dépose plainte à Marseille, il faut mettre en évidence qu'elle a été victime de la même escroquerie que d'autres personnes situées dans d'autres villes de France.

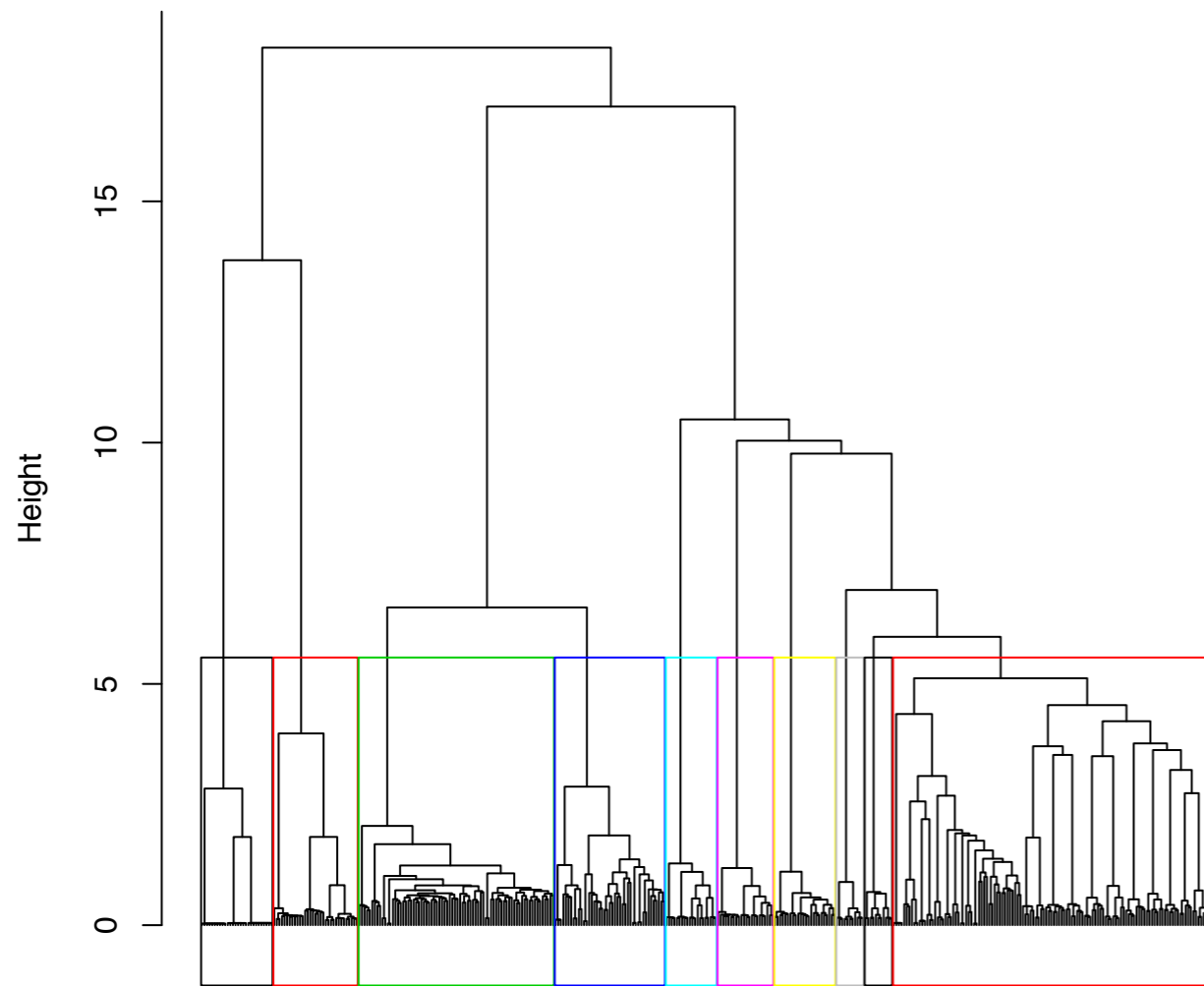
Fabien Lang, adjoint au chef de l'Office contre la cybercriminalité liée aux technologies de l'information et de la communication (OCLCTIC), France

Cas stratégique

- structure et origine des spams reçus ?
- clustering dans la phase exploratoire
- on cherche ensuite des liens entre les clusters
 - adresses IP
 - similarités dans les titres
 - encodage, ...
- Données : messagerie GMail

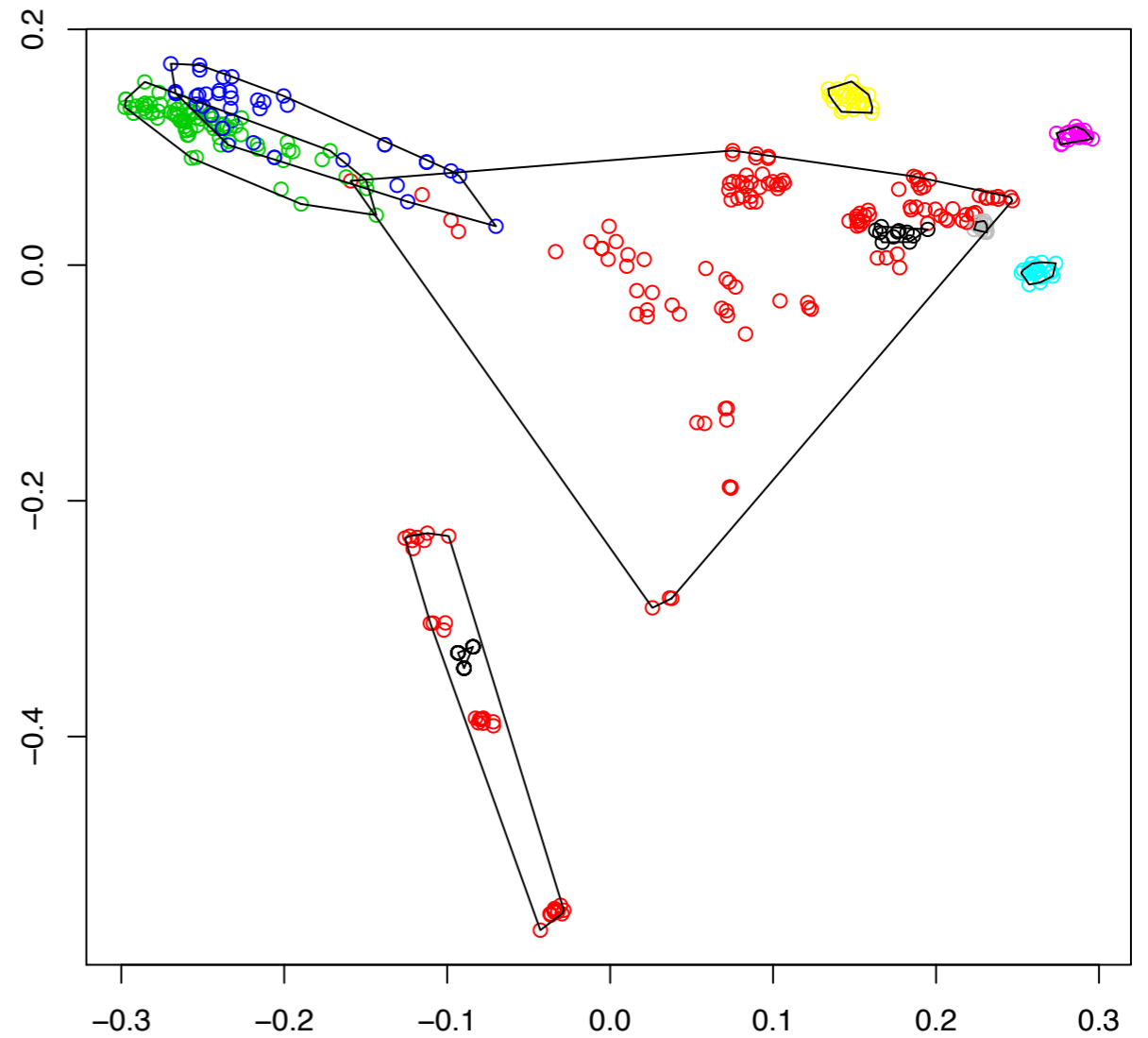
Cas stratégique

ward method

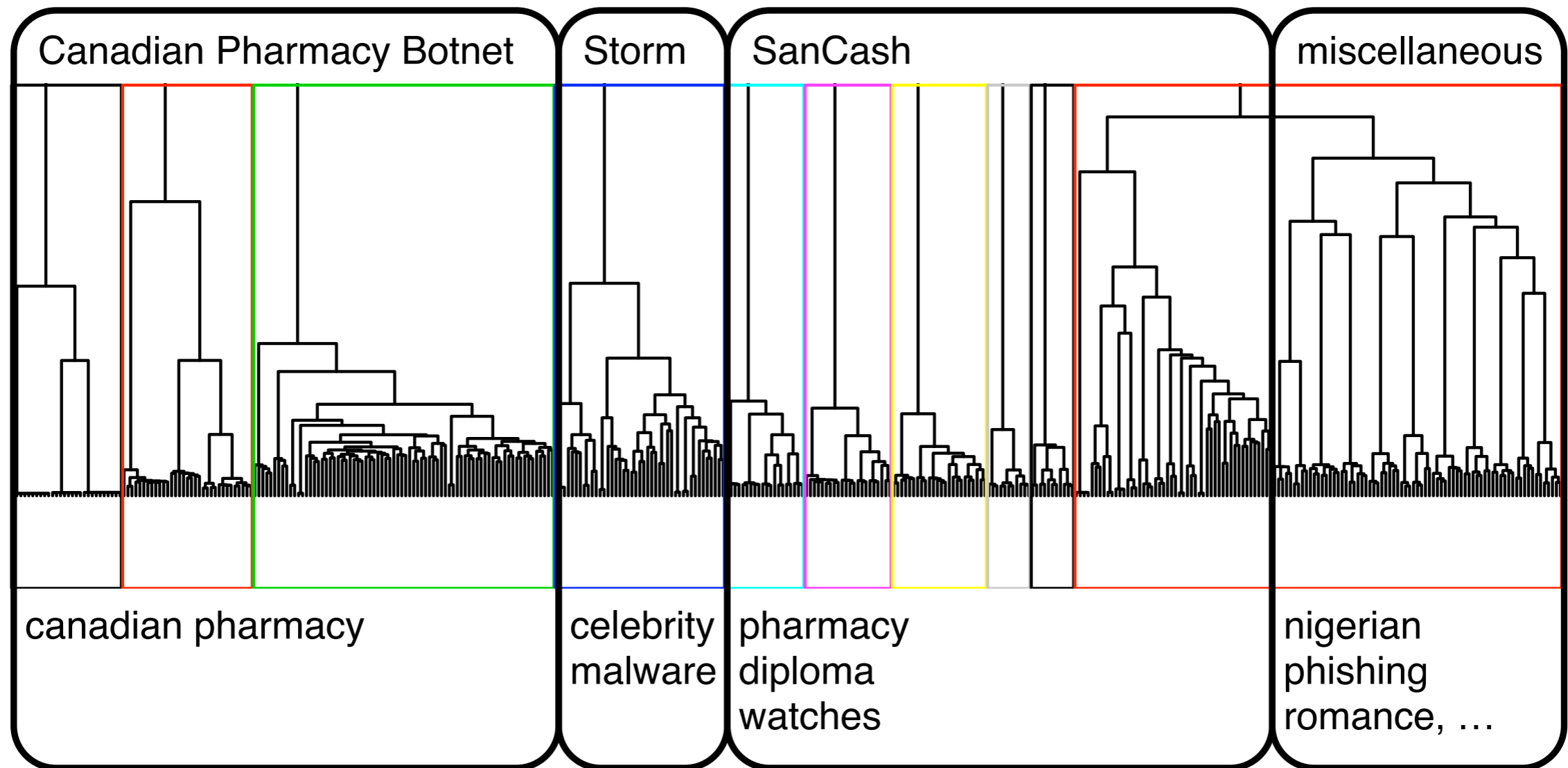


hclust (*, "ward")

text distances

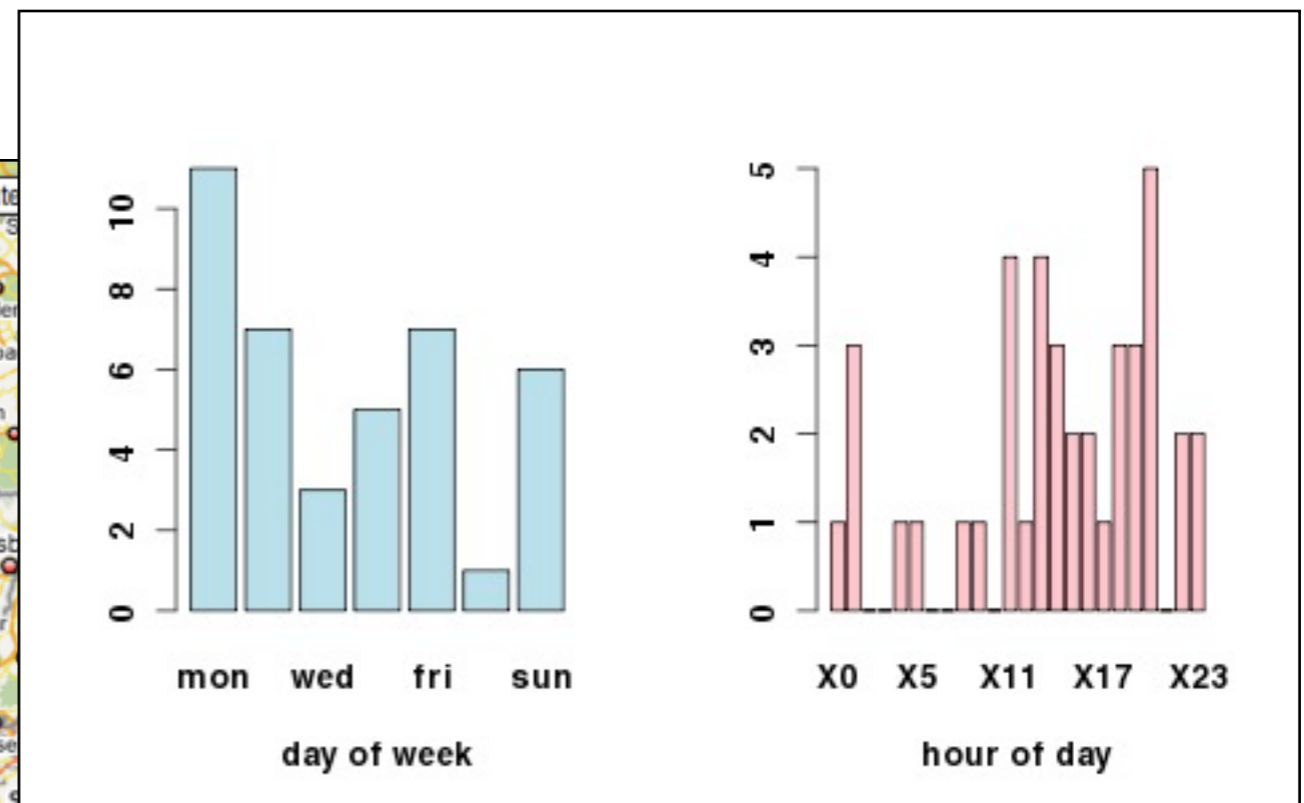
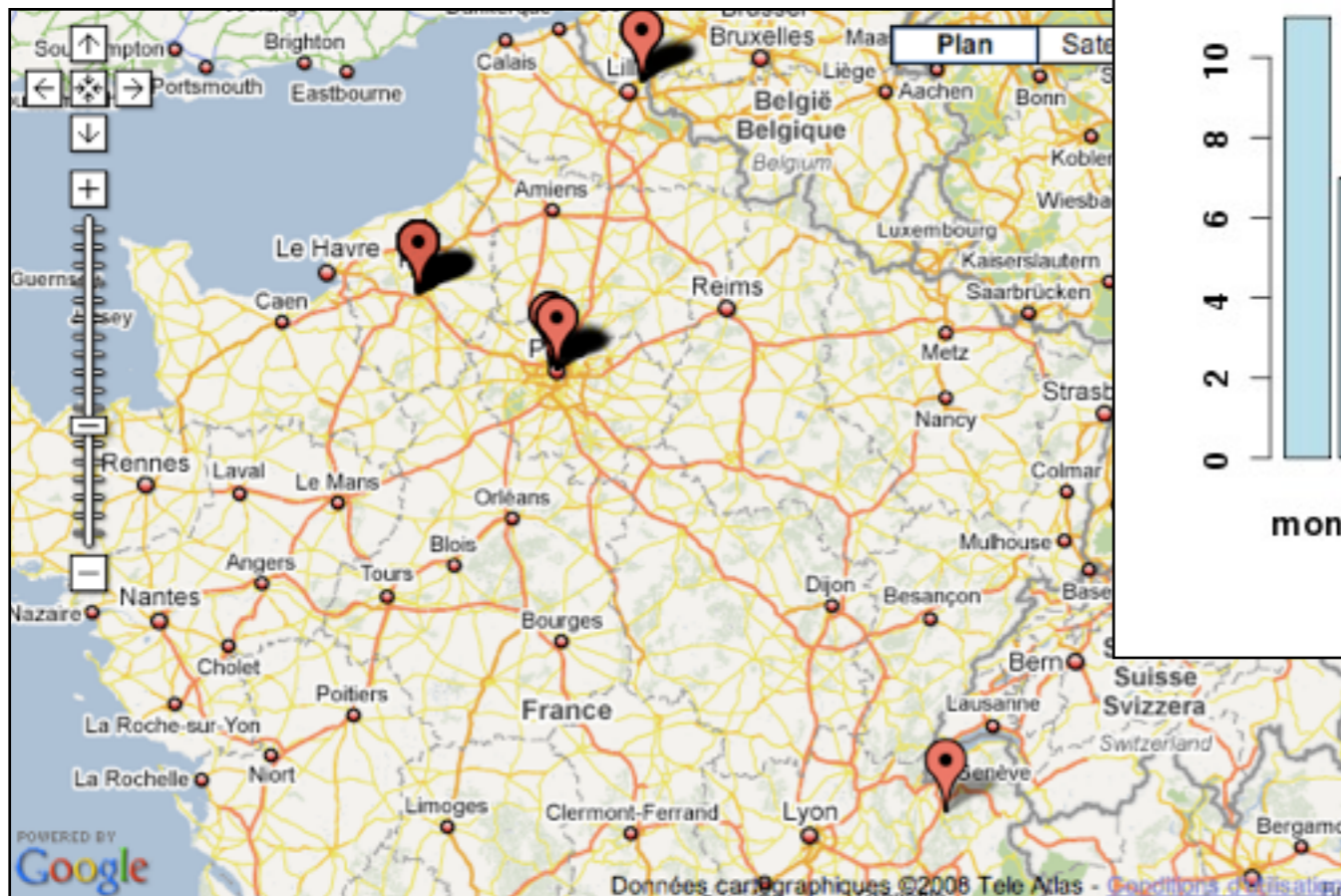


Renseignement stratégique



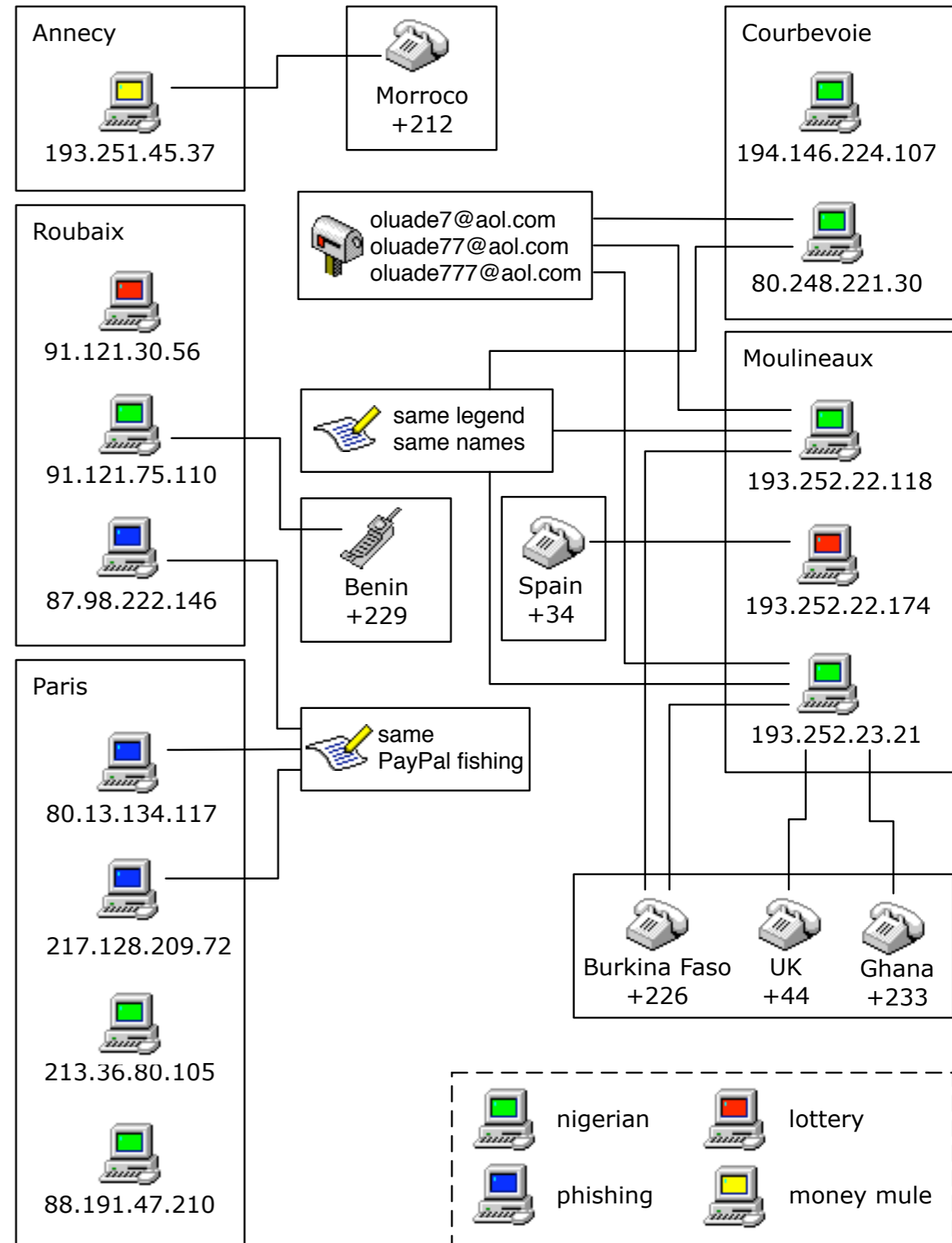
Cas tactique

- Les escroqueries envoyées depuis la France.
- 1950 messages frauduleux récoltés au printemps 2008, dont 40 en France.



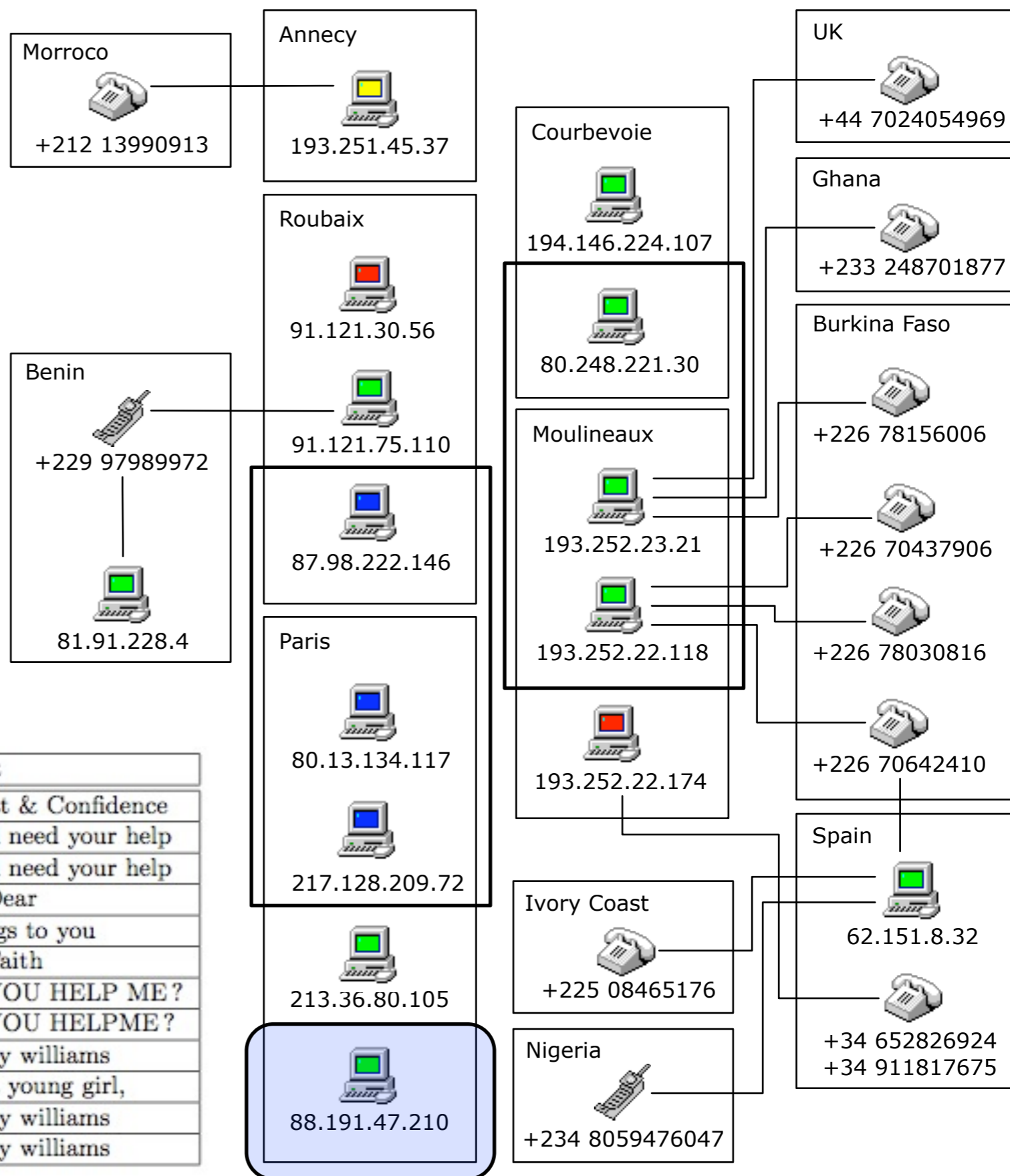
I. liens forts avec les machines en France (IP, email, ...)

2008-03-22 01:03:51	Work From Home Alert
2008-03-25 13:18:15	Your Response is needed
2008-03-25 13:21:10	Your Response is needed
2008-03-28 13:29:56	RE: SHELL PETROLEUM LOTTERY
2008-03-28 22:38:06	CONFIRMATION FOR WINNERS
2008-04-04 15:37:54	Regards
2008-04-08 17:39:01	FOR GOD CRAVE YOUR INDULGENCE
2008-04-10 13:32:11	bubaye(dme)
2008-04-14 22:21:40	CONFIDENCIAL/URGENT/ASSISTANT
2008-04-15 00:36:19	In Trust & Confidence
2008-04-20 14:26:30	Funds Clearance/Beneficiary
2008-04-20 14:26:30	Funds Clearance/Beneficiary
2008-04-21 01:11:58	Attention! Votre compte PayPal a été limité!
2008-04-21 05:57:21	Confirmer Votre Compte Orange en ligne!
2008-04-21 23:23:14	Service PayPal - Mettre à jour Votre Compte PayPal
2008-04-23 08:01:17	Account Suspension Notice - Votre compte PayPal a été limité .
2008-04-23 09:13:56	From Mr Emmanuel Please Call Me.
2008-04-23 18:02:39	Votre compte PayPal a été limite!
2008-04-24 11:00:52	Ecowas Contract Award
2008-04-24 11:04:22	Ecowas Contract Award
2008-04-24 14:00:14	timothy contacting you
2008-04-25 18:34:49	Dearest.Friend,
2008-04-25 18:35:05	Dearest.Friend,
2008-04-25 19:58:34	From Mr. Paul Smith
2008-04-25 19:59:33	From Mr. Paul Smith
2008-04-28 19:43:23	Dear Friend
2008-04-28 20:17:03	Dear Friend
2008-04-28 20:17:43	Dear Friend
2008-05-01 16:27:00	Service Alert! Attention Votre Compte PayPal a été limité .
2008-05-04 04:36:20	From Mr. John Alison.
2008-05-04 11:44:05	Great opportunity for us.
2008-05-04 12:34:22	Great opportunity for us.
2008-05-04 20:46:00	Getback to me as soon as you received this mail
2008-05-05 11:08:31	Great Opportunity for us.
2008-05-05 20:23:56	PLEASE REPLY IMMEDIATELY
2008-05-05 23:43:37	Great opportunity for us.
2008-05-06 15:53:32	Please read very carefully from
2008-05-12 01:58:31	MY DEARLY BELOVED FRIEND.
2008-05-13 16:01:39	TREAT AS URGENT PLS.
2008-05-13 20:11:55	Dearest



2. ajout des machines à l'étranger (liens forts)

3. considération des textes similaires (liens faibles)



Date	Country	Text distance	Subject
2008-04-15 00:36:19	France	0.026902	In Trust & Confidence
2008-04-16 21:03:19	Cote D'Ivoire	0.201896	please i need your help
2008-04-16 20:39:19	Cote D'Ivoire	0.201896	please i need your help
2008-04-09 19:49:21	Cote D'Ivoire	0.335960	Hello Dear
2008-05-01 13:25:26	Cote D'Ivoire	0.344109	greetings to you
2008-04-25 13:55:10	Cote D'Ivoire	0.455528	From Faith
2008-04-10 16:41:51	Cote D'Ivoire	0.459351	CAN YOU HELP ME ?
2008-03-29 15:21:28	Cote D'Ivoire	0.460737	CAN YOU HELPME ?
2008-05-01 17:06:35	Cote D'Ivoire	0.482412	from joy williams
2008-04-10 10:26:43	Cote D'Ivoire	0.482446	assist a young girl,
2008-05-02 04:11:24	Cote D'Ivoire	0.483033	from joy williams
2008-05-02 04:08:36	Cote D'Ivoire	0.483033	from joy williams

Renseignement tactique

- Un même groupe (ou un individu) envoie des spams nigériens depuis Courbevoie et Moulinaux, avec des numéros de contact en Grande-Bretagne et en Afrique de l'ouest.
- Un autre groupe envoie des phishing PayPal depuis Paris et Roubaix, rédigés en Français. Nous n'avons pas trouvé de traces de connexions de ce groupe avec l'étranger.

Conclusion

- on peut appliquer le clustering à une base de spams
- l'analyse multi-variée convient mal
- la distance entre les textes est plus intéressante
- le clustering permet de produire du renseignement
 - en complément à d'autres techniques
 - phase exploratoire pour le cas stratégique
- distance entre les textes pour le cas tactique

Questions ?

nicolas@seriot.ch

Merci