

L'utilisation des flux Internet dans l'analyse économique

Fabrice Le Guel, Pr. Thierry Pénard, Dr. Raphaël Suire

CREREG

UMR CNRS 6585

Centre de REcherche Rennais en Economie
et Gestion

GRANITIC

GRoupe d'ANalyse de l'Internet et des TIC
[HTTP://GRANITIC.UNIV-RENNES1.FR](http://granitic.univ-rennes1.fr)

Déroulement de la présentation

L'intérêt des flux Internet pour les économistes

1. Que mesure-t-on ?
2. Comment le mesure-t-on ?
3. Pourquoi le mesure-t-on ?

Perspectives

Références bibliographiques

Que mesure-t-on ?

Un état des lieux de la littérature

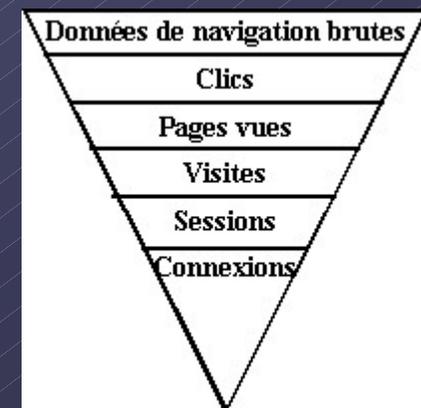
De l'analyse des données de navigation des internautes ...

Une littérature d'abord issue de l'informatique (gestion du réseau Internet) :

- Les 'précurseurs': Glassman (1994), Catledge & Pitkow (1995), Cunha et al. (1995),
- Les 'suiveurs' : Tausher (1996), Abdulla (1998)...

Permettant de construire une série d'indicateurs propres à l'analyse des comportements de navigation :

- Les clics, pages vues, visites, sessions, connexions,...



Que mesure-t-on ?

Un état des lieux de la littérature

... aux comportements **de choix** des agents économiques



Huberman, Pirolli, Pitkow, Lukose, (1998), *Science*

« [...] there is **value in each page a user visits**, and that clicking on the next page assumes that it will be valuable as well [...] an individual will continue to surf until the expected cost of continuing is perceived to be larger than the discounted expected **value of the information** to be found in the future. This can be thought of as a **real option** in financial economics, for which it is well known that there is a threshold value for exercising the option to continue » p.95

Hypothèse : les comportements de navigation des internautes sont des comportements de choix sous contraintes (temps, attention, revenu)

Choix entre différents sites concurrents:

‘Comportements de choix inter-sites’
(users centric)

Choix à l’intérieur d’un site:

‘Comportements de choix intra-site’
(site centric)

Que mesure-t-on ?

Un état des lieux de la littérature

Une littérature désormais économique et 'managériale'

Les quatre opportunités de recherches :

Usages (choix) / Niveau d'utilisation (inter-intra)	Comportements intra-site	Comportements inter-sites
Recherche d'informations	Choix entre différentes pages Bucklin & Sismerio (2003) Huberman, Pirolli, Pitkow, Lukose (1998) Johnson, Bellman, Lohse (2002)	Choix entre différents sites Goldfarb (2002) Lee, Zufryden, Drèze (2003) Johnson, Moe, Fader, Bellman, Loshe (2003) Montgomery & Faloutsos (2000)
Achats en ligne	Choix entre différents produits (commerce électronique) Moe & Fader (2003) Mandel & Johnson (2002) Fader & Hardie (2001)	Choix entre des produits 'homogènes' (<i>Shopbots</i>) Smith & Brynjolfsson (2001) Montgomery, Hosanagar, Krishnan, Clay (2003) Smith, Dick, Brynjolfsson (2003)

Comment le mesure-t-on ?

Le traitement des données de navigation et les outils d'analyse

1. Des données brutes à une matrice de choix individuels

Données de navigation

```

csl7 787125748 56906 "http://www.ludvigsen.dhhalden.no/webdoc/ahlan.gif" 1700 2.691705
csl7 787125764 828867 "http://www.ludvigsen.dhhalden.no/webdoc/ahlan.gif" 0 0.0
csl7 787125785 497644 "ftp://ftp.hiof.no/pub/levant/lebanon/fjeil/fjeil01.jpg" 21839 15.229143
csl7 787126872 204994 "http://www.ludvigsen.dhhalden.no/webdoc/levant_servers.htm#Lebanon" 0 0.0
csl7 787126872 389659 "http://www.ludvigsen.dhhalden.no/webdoc/ahlan.gif" 0 0.0
csl7 787128239 578430 "ftp://ftp.hiof.no/pub/levant/music_and_song/music.list" 40418 8.752783
csl7 787128301 795825 "http://www.ludvigsen.dhhalden.no/webdoc/levant_servers.htm#Lebanon" 0 0.0
csl7 787128302 100752 "http://www.ludvigsen.dhhalden.no/webdoc/ahlan.gif" 0 0.0
csl7 787128330 9639 "ftp://ftp.hiof.no/pub/levant/music_and_song/singers/AbdelHAlimHafez/abdh.txt" 1385 3.227795
csl7 787128350 941071 "http://www.ludvigsen.dhhalden.no/webdoc/levant_servers.htm#Lebanon" 0 0.0
csl7 787128351 203370 "http://www.ludvigsen.dhhalden.no/webdoc/ahlan.gif" 0 0.0
csl7 787132565 117453 "http://www.cpsr.org/dox/home.html" 3232 3.803559
csl7 787132569 78375 "http://www.cpsr.org/dox/cpsr.gif" 1644 8.892870
csl7 787132588 528202 "ht

```

Traitements des informations

Individu	Session	Site	Début	Durée	Bytes Echangées
156	1	http://home.mcom.com/	27/01/1995 18:20:01	58 s	147
156	1	http://lycos.cs.cmu.edu/	27/01/1995 18:20:59	2 m 33 s	1 325
156	2	http://www.mit.edu:8001/	21/02/1995 18:11:36	16 m 28 s	40 256
156	3	http://cathouse.org/	21/02/1995 19:12:35	15 m 29 s	6 895
156	4	http://cathouse.org/	22/02/1995 19:45:50	1 m 2 s	1 002
156	4	http://bvp.wdp.com/	22/02/1995 19:46:52	1 h 49 m 43 s	440 296

- 1: Google.fr
- 2: Voila.fr
- 3: Amazon.fr
- 4:FNAC.fr
- 5: Kelkoo.fr
- 0: Autre site

Choix y_{it}

Occasion de choix i sur une session ($t=1$ à T)

Individu ($i=1$ à N)

1	2	4	0	1	3	4	3	5	4	3	1	3
2	1	4	1	5	1	2	0	0	0	0	0	0
1	4	1	0	0	0	0	0	0	0	0	0	0

'Matrice des choix Individuels'

(3 individus, 13 occasions de choix)

Comment le mesure-t-on ?

Le traitement des données de navigation et les outils d'analyse

2. Les outils de l'analyse économétrique pour identifier les déterminants des choix

- Les modèles de choix discrets en données transversales

(test sur les CSP, revenu, âge...)

- Les modèles de choix discrets en données de panel

(Prise en compte de l'histoire : variation du revenu... + test de l'effet des choix passés sur les choix présents)

Occasions de choix (visites)

1	2	5	0	1	3	5	3	5	4	3	1	4
2	1	3	1	5	1	2	0	0	0	0	0	0
1	5	1	0	0	0	0	0	0	0	4	0	4

Individus

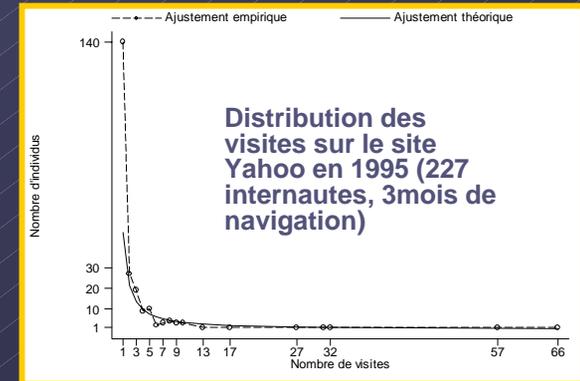


- Les modèles de durées

(Le temps comme variable explicative)

- Les modèles de comptage

[Distribution du nombre de visites d'un site et/ou d'achats en ligne, sur une période donnée et en fonction d'un ensemble de caractéristiques individuelles (exemple: expérience de navigation, âge, revenu...)]



Comment le mesure-t-on ?

Le traitement des données de navigation et les outils d'analyse

Un résultat sur modèle de choix discrets en données transversales

APPLICATION A UNE ENQUETE SUR LES USAGES INDIVIDUELS (2 000 ménages bretons)

FACTEURS	PROBABILITE D'ACHETER EN LIGNE
IMPACT POSITIF	<ul style="list-style-type: none"> • L'entourage achète aussi sur Internet • Utilisateur expérimenté de l'informatique et d'Internet
IMPACT NEGATIF	<ul style="list-style-type: none"> • L'entourage n'achète pas sur Internet • CSP cadre supérieur, profession intellectuelle supérieure
NEUTRES	<ul style="list-style-type: none"> • Age • Niveau d'étude • Sexe • Localisation • Qualité de la connexion Internet (haut débit ou bas débit) • Possession d'autres équipements high tech

Source : Le Guel, Pénard, Suire (2003) « Adoption et usage marchand de l'Internet : une étude économétrique sur données françaises », Document de travail CREREG.

Pourquoi le mesure-t-on ?

Internet: un univers 'sans friction' ?

Hypothèses de la littérature

INTERNET (Bakos, 1997., Shapiro & Varian, 1999) =

- disparition des coûts de déplacement =
- une demande plus volatile (la 'concurrence n'est qu'à un seul clic') =
- dispersion des prix plus faible en ligne que dans le monde tangible =
- marché plus efficient et augmentation du bien-être social

- des agents économiques mieux informés =
- une diminution des inégalités sociales

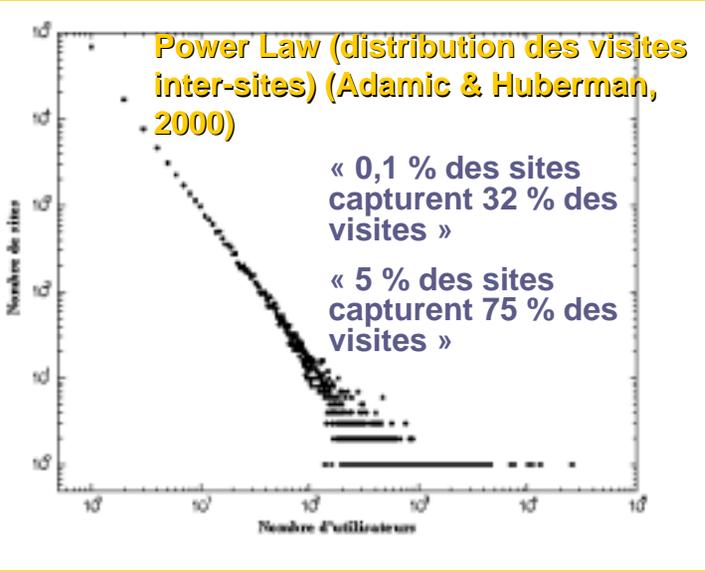
Pourquoi le mesure-t-on ?

Le paradoxe des usages en ligne : persistance des choix dans un univers supposé 'sans friction'

Résultats de la littérature empirique :

- (1) : Une dispersion des prix en ligne (Lee, 1997., Bailey, 1998, Brynjolfsson & Smith, 2000., Pénard, 2001, Pénard-Larribeau 2002...)
- (2) : Une persistance des choix sur Internet (Adamic & Huberman, 1998, Montgomery, 2003) :

La majorité des internautes visitent une minorité de sites/ de pages



Matrice de Transition intra-site (Montgomery, 2003)

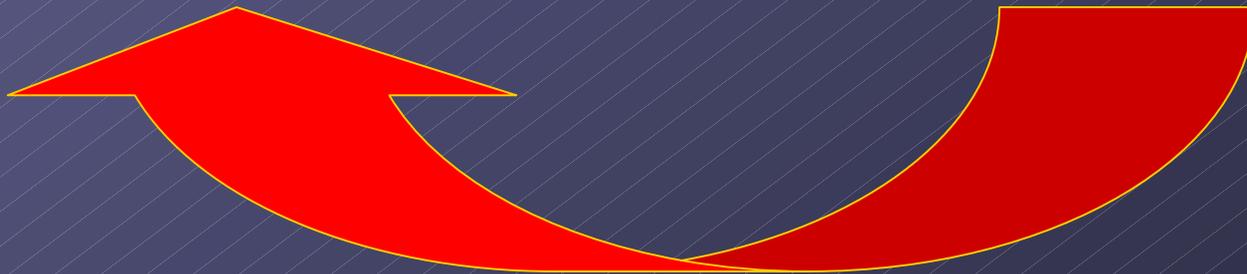
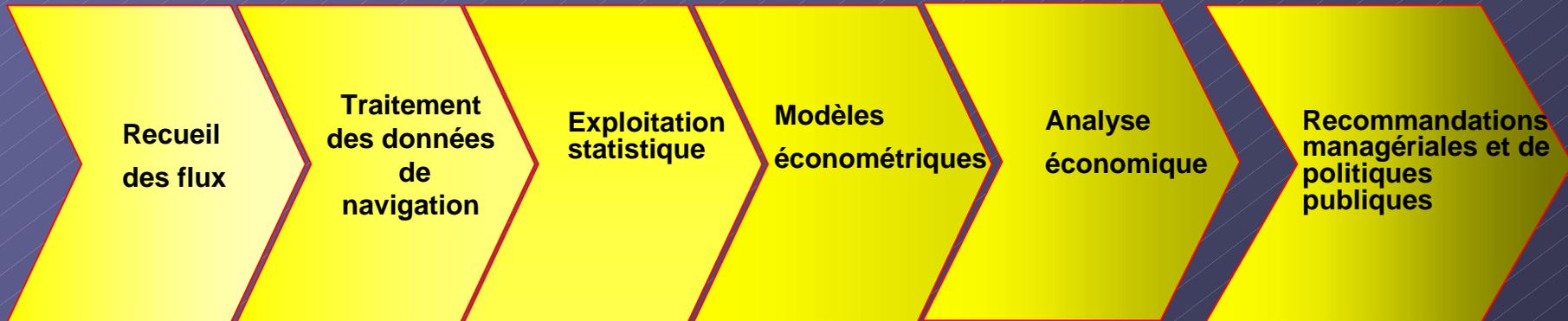
		Category of Previous Viewing							
		Home	Account	Category	Product	Inform.	ShopCart	Order	Exit
Category of Current Viewing	Home	.23	.01	.01	.01	.10	.02	0	.16
	Account	.01	.69	.01	.01	.02	.15	0	.01
	Category	.17	.02	.60	.31	.15	.05	0	.16
	Product	.01	0	.20	.43	.10	.05	0	.05
	Information	.25	.06	.08	.12	.46	.15	.87	.61
	Shop. Cart	.01	.16	.01	.03	.02	.45	.13	.01
	Order	0	0	0	0	0	.10	0	0
	Exit	.32	.06	.09	.09	.14	.02	0	0
Marginal		.06	.05	.32	.17	.23	.05	.01	.11
Initial Prob.		.16	.02	.16	.06	.60	.01	0	0

- (3) : L'existence d'une fracture numérique de 'second niveau' ? Celle des usages en ligne (Hargittai, 2002)

Perspectives

A la recherche de partenariats

Chaîne d'expertise



Quelques Références:

- Moe, W., Fader, P. S., (2003), 'Dynamic conversion behavior at e-commerce sites', Under review at *Management Science*.
- Bucklin, R. E., Lattin, J. M., Ansari, A., Bell, D., Coupey, E., Gupta, S., Little, J. D. C., Mela, C., Montgomery, A., Steckel, J., (2002), 'Choice and the Internet : from clickstream to research stream', *Marketing Letters*, Vol. 13, n° 3.
- Smith M.D., Brynjolfsson E., (2001), 'Consumer Decision-making at an Internet Shopbots ', *The Journal of Industrial Economics*, Vol. XLIX, N°4.
- Montgomery, A. L., Faloutsos, C., (2000), 'Trends and patterns of WWW browsing behavior', *Working Paper*, Carnegie Mellon University.
- Huberman, B. A., (2001), 'The laws of the web. Patterns in the ecology information', MIT Press.
- Huberman, B. A., Pirolli, P. L. T., Pitkow, J. E., Lukose, R. M., (1998), 'Strong regularities in the World Wide Web surfing', *Science*, April 3, Vol 280, n° 5360.
- Lee, S., Zufryden, F., Dreze, X., (2003), 'A study of consumer switching behaviour across Internet Portal websites', *International journal of Electronic Commerce*, Spring, Vol. 7 (3).
- Goldfarb, A., (2002), 'Analyzing Website Choice using Clickstream Data', in *Advances in Applied Microeconomics*, Vol. 11, The Economics of the Internet and E-commerce, Michael R. Baye. (Eds), Elsevier Science Ltd.

MERCI

fabrice.lequel@univ-rennes1.fr

thierry.penard@univ-rennes1.fr

raphael.suire@univ-rennes1.fr