MFG : BILAN ET PERSPECTIVES

Pierre-Louis LIONS

Collège de France, Paris

(Projet MFG en collaboration avec Jean-Michel LASRY)

. . . et beaucoup d'autres . . .

SOMMAIRE

1. INTRODUCTION

• 10 ans de FDD (de Monaco aux Vaux de Cernay)

• Une “chaire” innovante devenue une constellation avec des IdR (inventées par la chaire) tournées vers des projets applicatifs concrets (FIME, “indices et responsabilité carbone”, assurance récolte, microstructure des marchés. . .)

• Une “marque” reconnue par la communauté scientifique. . .

• Un réseau unique de chercheurs de haut niveau (maths, économie, finance, data. . .)

• Parmi les découvertes scientifiques dans le cadre de la FDD les MFG (presque 10 ans !)

⋆ croissance extrêmement rapide de la communauté MFG et des publications (plusieurs congrès chaque année)

⋆ des applications qui se multiplient (économie, foules, data, ingénierie, réseaux, pas la voile !)

• Objectif de l'exposé : faire le point et évoquer quelques perspectives. . .).

2. LES MODELES MFG

I. INTRODUCTION

• Une nouvelle classe de modèles pour le comportement moyen (Mean Field) d'un grand nombre de "petits" agents (Games) introduit au début des années 2000 par J-M. Lasry and P-L. Lions.

• Requiert de nouvelles théories mathématiques.

• Très nombreuses applications : économie, finance, réseaux sociaux, mouvements de foules. . .

• Introduction indépendante d'une classe particulière de modèles MFG par M. Huang, P.E. Caines et R.P. Malhamé en 2006.

• Références et une grande partie des résultats mathématiques dans les vidéos du cours (et quelques sminaires) du Collège de France (4×18h) téléchargeables

• Combinaison des théories de Champ Moyen (classiques en Physique et Mécanique) et de la notion d'équilibres de Nash en théorie des Jeux.

• Equilibres de Nash pour des continua de "petits" joueurs (adaptations aisées à plusieurs groupes. . .).

• Interprétation dans des cas particuliers (mais déjà assez généraux!) comme le contrôle de processus de McKean-Vlasov. . .

• Chaque joueur générique est “rationnel” i.e. essaie d’optimiser un critère (contrôle) qui dépend des autres et la décision optimale modifie le comportement du groupe (noter que cette interprétation est limitée à une sous-classe de modèles. . .).

• Classe “énorme” de modèles : “les agents sont des particules”, “pas de dépendance des autres” sont deux cas particuliers extrêmes.

• Des antécédents en Economie (jeux anonymes et surtout Krussell-Smith)

II. UN EXEMPLE VRAIMENT SIMPLE

• Un exemple simple, déjà connu mais qui donne une idée de la classe générale de modèles (d'autres exemples "simples" plus loin. . .)

• E espace métrique, N joueurs ($1 \leqslant i \leqslant N$) choisissent une position $x_i \in E$ suivant un critère $F_i(X)$

où $X = (x_1, \ldots, x_N) \in E^N$.

• Equilibre de Nash : $\bar{X} = (\bar{x}_1, \ldots, \bar{x}_N)$ si $\forall\ 1 \leqslant i \leqslant N$ $\bar{x}_i$ min sur E de $F_i(\bar{x}_1, \ldots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \ldots \bar{x}_N)$.

• Difficultés habituelles avec la notion

• $N \to \infty$? plus simple ?

• Joueurs indistinguables :

$$F_i(X) = F(x_i, \{x_j\}_{j \neq i}),\ F \text{ sym. en } (x_j)_{j \neq i}$$

• Un (petit) morceau des théories mathématiques concernant la limite $N \to \infty$:

$$F_i = F(x, m) \quad x \in E \,, \quad m \in \mathcal{P}(E)$$

$$\text{où} \;\; x = x_i \,, \;\; m = \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}$$

• "Thm" : Les équilibres de Nash convergent, si $N \to \infty$, vers les solutions de

$$\text{(MFG)} \qquad \forall x \in \text{ Supp } m, F(x, m) = \inf_{y \in E} F(y, m)$$

• Faits : i) résultats généraux d'existence et de stabilité

ii) unicité si $(m \to F(\bullet, m))$ monotone

iii) Si $F = \Phi'(m)$, alors $(\min_{\mathcal{P}(E)} \Phi)$ permet d'obtenir une solution de MFG.

Exemple : $E = \mathbb{R}^d, F_i(X) = f(x_i) + g\left(\frac{\#\{j/|x_i - x_j| < \varepsilon\}}{(N-1)|B_\varepsilon|}\right)$

$g \uparrow$ aversion à la foule, $g \downarrow$ appétance pour la foule

$$F(x,m) = f(x) + g(m * 1_{B_\varepsilon}(x)(|B_\varepsilon|^{-1})$$

$$\varepsilon \longrightarrow 0 \qquad F(x,m) = f(x) + g(m(x))$$

(MFG) $\qquad \mathrm{supp}\ \ m \subset \text{Arg min}\left(f(x) + g(m(x))\right)$

– $g \uparrow$ unicité, $g \downarrow$ non unicité

$$\min\left\{\int fm + \int G(m)/m \in \mathcal{P}(E)\right\},\ G = \int_0^z f(s)ds$$

– solution explicite si $g \uparrow$: $m = g^{-1}(\lambda - f), \lambda \in \mathbb{R}\ \text{t.q.} \int m = 1$

III. LA STRUCTURE GÉNÉRALE

• Un cas particulier : problème dynamique, horizon fini T, temps et "espace" continus, bruits Brownien (à la fois indépendants et communs), pas de taux de préférence intertemporelle, contrôle sur le drift (Hamiltonien H), critère ne dépendant que de m

• $U(x, m, t)$ $(x \in \mathbb{R}^d, m \in \mathcal{P}(\mathbb{R}^d)$ ou $\mathcal{M}_+(\mathbb{R}^d), t \in [0, T]$ et $H(x, p, m)$ (convexe en $p \in \mathbb{R}^d$)

• MFG master equation

$$\begin{cases} \frac{\partial U}{\partial t} - (\nu + \alpha)\Delta_x U + H(x, \nabla_x U, m) + \\ + \langle \frac{\partial U}{\partial m}, -(\nu + \alpha)\Delta m + \operatorname{div}\left(\frac{\partial H}{\partial p} m\right)\rangle + \\ -\alpha \frac{\partial U}{\partial m^2}(\nabla m, \nabla m) + 2\alpha \langle \frac{\partial}{\partial m} \nabla_x U, \nabla m \rangle = 0 \end{cases}$$

et $U\,|_{t=0} = U_0(x, m)$ (coût final)

• ν "quantité d'aléa indépendant" , α "quantité d'aléa commun"

• Problème en dimension infinie !

• Si $\nu = 0$ (ind) : les équilibres de Nash à N joueurs sont un cas particulier

en prenant $x = x_i, m = \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}$

• Agrégation/décentralisation : SI $H(x,p,m) = H(x,p) + F'(m)$ et $U_0 = \Phi_0'(m)$, alors $U = \frac{\partial \Phi}{\partial m}$ résout MFG si Φ résout HJB sur $\mathcal{P}(E)$ pour le contrôle optimal d'une EDPS

• Cas particulier : beaucoup d'extensions et de variantes. . .

IV. DEUX CAS PARTICULIERS

• Problème ∞ d en général mais réductions à la dimension finie dans deux cas

1. Bruits indépendants ($\alpha = 0$)

intégration le long des caractéristiques en m donne

$$(\mathrm{MFGi}) \begin{cases} \frac{\partial u}{\partial t} - \nu \Delta u + H(x, \nabla u, m) = 0 \\ u \mid_{t=0} = U_0(x, m(0)), m \mid_{t=T} = \bar{m} \\ \frac{\partial m}{\partial t} + \nu \Delta m + \operatorname{div} \left(\frac{\partial H}{\partial p} m\right) = 0 \end{cases}$$

où $\bar{m}$ est donnée

Système FORWARD — BACKWARD !

contient comme cas particuliers : HJB, chaleur, milieux poreux, FP, V, B, Hartree, elliptique semilinéaire, barotrope Euler ...

2. Espace d'états fini ($i \leqslant i \leqslant k$)

(MFGf) $\frac{\partial U}{\partial t} + (F(x, U) \, . \, \nabla) \, U = G(x, U), U \,|_{t=0} = U_0$

(pas de bruit commun pour simplifier ...)

$$x \in \mathbb{R}^k, \; U \to \mathbb{R}^k, \; F \; \text{et} \; G : \mathbb{R}^{2k} \to \mathbb{R}^k$$

système hyperbolique non conservatif

Exemple : Si $F = F(U) = H'(U), G \equiv 0$

et si $U_0 = \nabla \varphi_0$ $(\varphi_0 \to \mathbb{R})$ alors

– on résout HJ

$$\frac{\partial \varphi}{\partial t} + H(\nabla \varphi) = 0 \; , \; \varphi \,|_{t=0} = \varphi_0$$

– on pose $U = \nabla \varphi$, U "résout" (MFGf) dans ce cas

3. MATHS

De très nombreuses questions, des résultats partiels et de nombreux problèmes ouverts

- Existence/régularité :
  - (MFGi) "simple" si H "régulier" en m (ou si H au plus linéaire...), OK si monotone (ZOOM 1) (Cardaliaguet-Delarue-Lasry-L)
  - (MFGf) OK si (G, F) monotone sur $\mathbb{R}^{2k}$ ou temps petit (ZOOM 2)
- Unicité : OK si "monotone" ou T petit...
- Non existence, non unicité, pas de régularité (!)
- Propriétés qualitatives, états stationnaires et stabilité, comparaison, cycles...
- $N \to \infty$ (voir ci-dessus) (Cardaliaguet-Delarue-Lasry-L)
- Méthodes numériques (3 méthodes assez "générales" et des cas particuliers...)

- Variantes : autres bruits, plusieurs populations...
- Hétérogénéité aléatoire, information partielle...
- Applications (MFG Labs...)
- Espaces d'état de dimension infinie (mesures, historique, bayésien, Vlasov-McKean, états conditionnels)
- Point de vue Hilbertien pour des équations sur l'espace des mesures
- Taux de préférence intertemporelle (et $\lambda \to \infty$ modèles effectifs)

- Limites macroscopiques
- ? Au delà des MFG ? (fluctuations, grandes déviations, transitions)
- MFG et temps optimal (liens avec l'optimal design) – thèse Ch. Bertucci
- Deux exemples S. :
  - à quelle heure commence la réunion ?
  - la holla
- Remarque : les MFG nous ont permis de faire des progrès dans la compréhension de questions classiques (Hamilton-Jacobi, solutions stratifiées, contrôle stochastique. . .) et dans la manipulation d'équations sur l'espace des mesures (Vlasov-McKean, contrôle Bayésien et info partielle. . .)

ZOOM 1

$$
\text{(MFGi)} \begin{cases} \frac{\partial u}{\partial t} - \nu \Delta u + H(x, \nabla u) = f(x, m) \\ u \mid_{t=0} = U_0(x), m \mid_{t=T} = \bar{m} \\ \frac{\partial m}{\partial t} + \nu \Delta m + \operatorname{div}\left(\frac{\partial H}{\partial p} m\right) = 0 \end{cases}
$$

- $m \mapsto f(\bullet, m)$ opérateur régularisant
  $\exists$ solution régulière
- unicité si opérateur monotone ou si T petit
- $f(m(x)) \uparrow$: $\exists$! solution régulière $\nu > 0$
- $f(m(x)) \uparrow$: si $\nu = 0$ $m = f^{-1}(\frac{\partial u}{\partial t} + H(x, \nabla u))$

  l'équation en m devient alors une équation quasilinéaire elliptique du second ordre $(x \in Q, t \in [0, T])$ avec les conditions aux limites "elliptiques"

$$
u \mid_{t=0} = U_0(x), \ \frac{\partial u}{\partial t} + H(\nabla u) = f(\bar{m}) \ \text{si} \ t = T
$$

ZOOM 2

$$(\text{MFGf}) \begin{cases} \frac{\partial u}{\partial t} + (F(x,U)\,.\,\nabla)\,U = G(x,U) \;\; x \in \mathbb{R}^d \\ U \to \mathbb{R}^d \;,\; U\,|_{t=0} = U_0(x) \end{cases}$$

- chocs (discontinuités de U) en temps fini en général
- Problème bien posé sur $[0, T_{\max})$ ($T_{\max} \leqslant +\infty$)
- $\exists$! solution régulière monotone en x si U_0 monotone et (G, F) monotone de $\mathbb{R}^{2,k}$ dans $\mathbb{R}^{2k}(+\ldots)$
- $+$ changement de fonctions inconnues :

  ex. : $\frac{\partial U}{\partial t} + (F(U).\nabla)U = 0$

alors $V = F(U)$ résout

$$\frac{\partial V}{\partial t} + (V.\nabla)V = 0$$

classe max de régularité

$$\forall \delta > 0, \ \inf_{x \in \mathbb{R}^d} \operatorname{dist}(Sp(DV_0(x)), (-\infty, \delta]) > 0$$

($V_0 = F(U_0)$ donne la classe maximale de régularité $\approx$ composée de 2 applications monotones)

Remarque : donne des résultats nouveaux de régularité pour les équations de Hamilton-Jacobi du 1er ordre.

ZOOM 3 : Le point de vue Hilbertien

(Ω, F, P) espace de proba suffisamment “riche”, $\mathcal{H}$ Hilbert $L^2(\Omega, P)$

$$\Phi(m) = \Phi(X) \text{ si } \mathcal{L}(X) = m$$

Les MFG deviennent alors

$$\frac{\partial U}{\partial t} + (\mathcal{F}(X,U).D)U = \mathcal{G}(X,U) + \alpha \Delta_d U$$

$$(+\nu\ D^2U(G,G) \quad G \perp F_X)$$

$$\Delta_d U = \Delta_Z\ U(. + Z)|_Z = 0(Z \in \mathbb{R}^d)\ ,$$

où $U : \mathcal{H} \to \mathcal{H}$

Remarques : 1) MFG $\quad U(X) \in F_X, \mathcal{L}(U(X)) = \mathcal{L}(U(Y))$
si $\mathcal{L}(X) = \mathcal{L}(Y)$
2) $U(X) = \nabla_x U(x, \mathcal{L}(X))|_{x=X}$

Permet de prouver que le problème est bien posé si T petit ou dans le cas monotone (sur $\mathcal{H}, \neq$ monotone en m ...)

4. APPLICATIONS

I. ECONOMIE/FINANCE

• Jeux anonymes en temps continu ! (modèle de Ayagari, hétérogénéité des agents)

Achdou-Lasry-Moll-L

• Comprendre Krussell-Smith et simulations numériques

Achdou-Lasry-Moll-L

• Compétences et carnets d'adresses

Lucas-Moll

• Modèles "d'industries minières" (ressources non ou partiellement renouvelables)

Achdou-Giraud-Lasry-Scheinkman-L

– situation multi-agents, offre/demande à 2 niveaux

– réserves/stocks (quantité de biens ou clients...)

– coûts de production (ou de mise sur le marché)

– investissements (augmenter les stocks ou acquisitions)

et la possibilité de prendre en compte de nouveaux entrants

• modèles de communautés : communautés "libres" (incitations collectives, bien commun) ou services collaboratifs avec plate-forme propriétaire). Valeur des "biens gratuits"

Lasry-L + ...

II. DIVERS

- Mouvements de foule
- Réseaux de communication, flotte de véhicules. . .
- 2 exemples S.

III.2 MEANINGFUL DATA

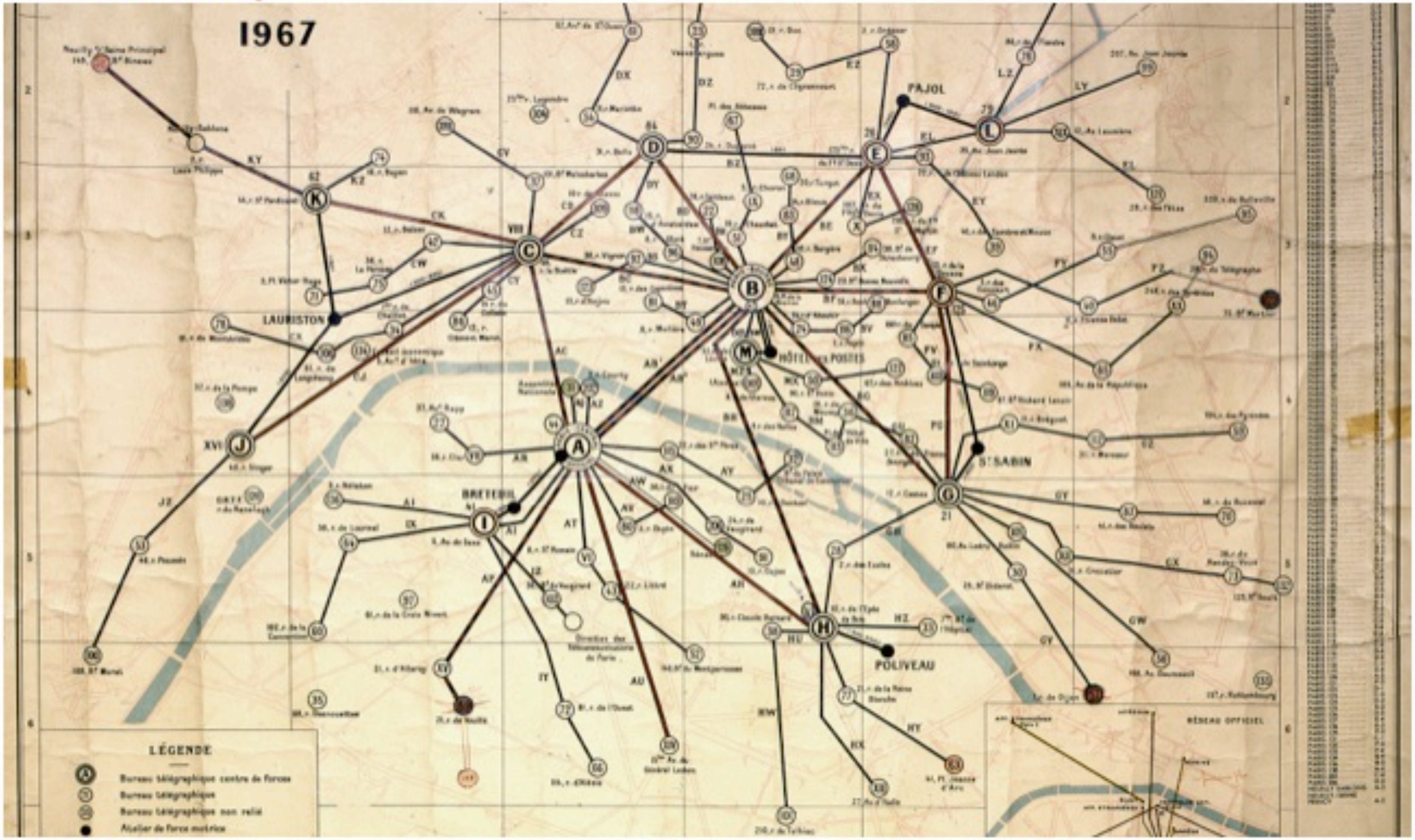
- Meaningful Data / Big Data
- Sujet ancien (statistiques, actuariat, micro…) très à la mode depuis quelques années pourtant…
- Fusion des modes historiques de stockage, de transmission et d'utilisation des informations (nombres, calcul, écriture, livres, enregistrements sonores, photos, films…) dans un monde numérisé massivement connecté
- Les capacités de traitement essentiellement constantes et réduites pendant des siècles ont explosé (presque infinies!)
- Tendance: intelligence/travail/service partagé/collaboratif/ coopératif!

=10 vraiment

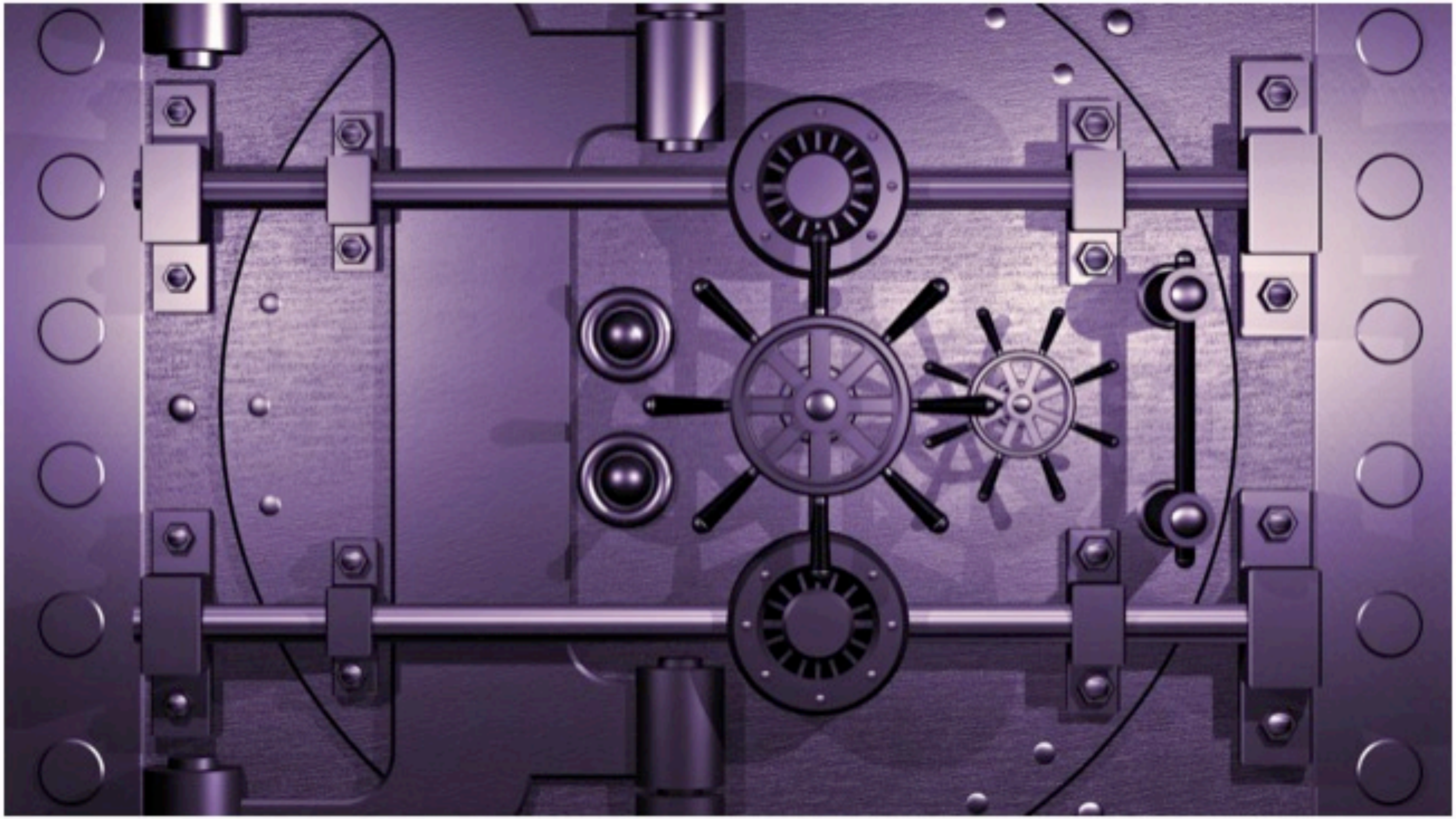
- Vraiment « big »: Avogadro
- Vraiment rapide: 50 ms pour estimer la proba d'un click sur un ad
- Vraiment varié: structuré ou non...
- Vraiment connecté: Réseaux Sociaux, Forums…
- Vraiment « creux » (sparse): FB 1 B users, 10 M pages, fréquence 1/1000000
- Vraiment contextuel (Tweets)
- Vraiment difficile à visualiser
- Vraiment bruité/sale
- Vraiment important

Biblio.Trinity College x 1 milliard de milliards

Un petit réseau 2D / FB (1milliard de nœuds)



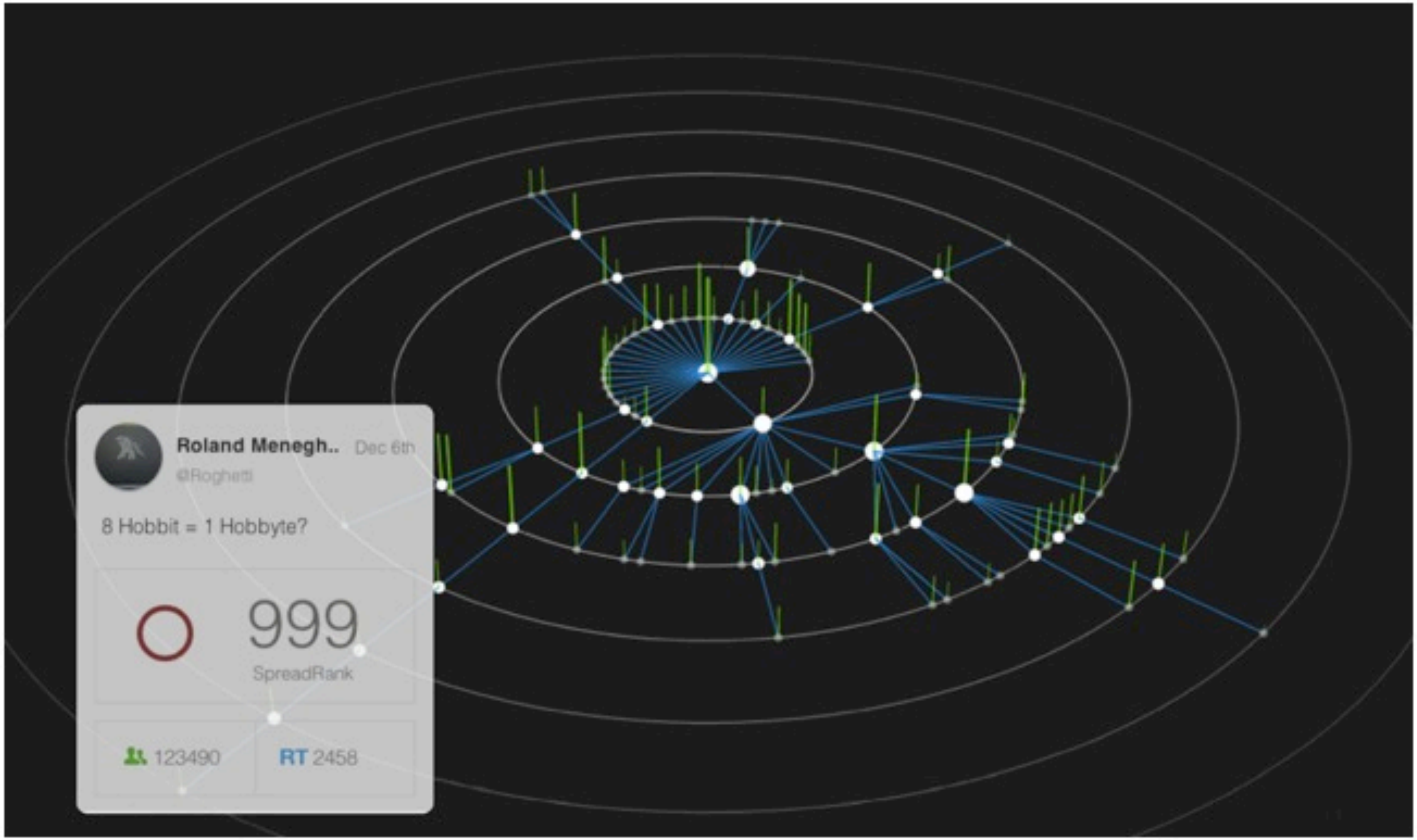
Nouvelle richesse?



Quelques faits (suite)

- Seulement 9? Vraiment difficile et
- Potentiel économique énorme (CRM, Marketing, Ads, Pricing, Marques…): 300 % de croissance par an de ce secteur pour la France
- Ruée sur les data scientists!
- Risques: Confiance/Sécurité et données personnelles, mais aussi de trop promettre sans bien penser l'application, la chaîne et les algorithmes
- Du concret à la science au concret à la…. : MFG Labs (2010, Havas Media 2013)
- Where does my tweet go (freeware), CINEMUR (spin-off), MyWarner pour WB

TWITTOSPHERE



Tourisme en France (photos partagées sur le Web)

Quelques instants de science pour conclure!

- Une fois l'application bien pensée et la chaîne bien concue, quid des algo? Pour « clusteriser », classifier…
- De nombreux développements importants dans la conception et l'implémentation efficace d'algorithmes (Informatique et ML, Optimisation, Statistiques, Théorie des graphes, RO, Physique Statistique…) qui permettent d'aborder de très nombreuses applications
- Mais il subsiste des difficultés cruciales: le choix des métriques, peu de prédictibilité, relations de causalité souvent inacessibles, modélisation rudimentaire, non prise en compte du facteur humain (choix/stratégies basés sur le passé, le présent et le futur par anticipation)

Une voie possible: MFG pour MFD

- Modèles MFG: une classe nouvelle de modèles mathématiques combinant des idées de Théorie Des Jeux et de Physique Statistique (J.M.Lasry….) pour comprendre et prédire le comportement d'un grand nombre d'agents/joueurs
- Permet de prendre en compte le facteur humain, de définir des métriques de manière auto-cohérente (FB et films…) et d'englober beaucoup de méthodes classiques en ML et Stats (K-Mean, EM…) en considérant chaque data point comme un agent et les clusters comme des clubs…
- BEAUCOUP RESTE A FAIRE!!!

III. DONNÉES SIGNIFIANTES

• MFG Labs

• Une expérience pratique et des modèles surtout pour les "big data actifs""

• Nouveaux modèles qui incluent les modèles classiques de "clustering" en M.L. (K-mean, EM . . .), puis des algorithmes

• Pas besoin de structures euclidiennes ou de distances "a priori"

- Données “actives”

Ex. 1 : Taxis

Ex. 2 : Films

Des personnes sont “proches” si elles aiment des films qui sont “proches” : déterminer des (sortes de) distances cohérentes sur les items/personnes

• Même pour des “données passives”, les modèles font sens : les données deviennent des agents... (déjà beaucoup de terminologie de Théorie des Jeux en M.L./Analyse de Données)

• Clustering : K-Mean classique

ensemble de points $\{x_1, \ldots, x_v\}$ dans $\mathbb{R}^d (N >> 1)$

Trouver K points $y_1, \ldots, y_k$ tels que $\exists$ partition $(A_1, \ldots, A_K)$ de $\{1, \ldots, N\}$ pour laquelle

$$\text{i) } |y_i - x_j| \leqslant |y_{i'} - x_j| \ , \ \forall_j \in A_i \ , \ \forall i' \neq i$$

$$\text{ii) } y_i = (\# A_i)^{-1} \sum_{j \in A_i} \times j$$

INTERPRETATION MFG :

INTRODUIRE • UN CRITÈRE GLOBAL

$F(u_1, \dots, u_k)$

<u>Ex</u> : $\min(u_1, \dots, u_k)$

- K fonctions valeurs $(u_1, \dots, u_k)$
- K "densités" $(m_1, \dots, m_k)$

f densité initiale des données (discrètes ou pas !)

MFG : EXEMPLE

$$\rho u_i - \nu \Delta u_i + \frac{1}{2}(\nabla u_i)^2 = F_i(x; m_i)$$

$$\rho m_i - \nu \Delta m_i - div(\nabla u_i m_i) = \rho \, \frac{\partial F}{\partial u_i} \, f$$

ex. $\quad 1_{(u_i < \min\limits_{j \neq i} u_j)}$

RETOUR SUR K-Mean

$$F_i = \frac{1+\rho}{2}|x - \frac{\int xm_i}{\int m_i}|^2 - \nu d$$

$$\text{alors} \ : \ u_i = \frac{1}{2}|x - y_i|^2, y_i = \frac{\int xm_i}{\int m_i} \ \text{ et}$$

$$\int m_i = \int f 1_{(u_i < \min_{j \neq i} u_j)}, \int xm_i = \int xf_i 1_{(u_i < \min_{j \neq i} u_j)}$$

Cela permet de

- créer de très nombreux nouveaux modèles

• régulariser les clusters, de permettre des clusters dans des clusters, des recoupements...

• pas besoin de distances, pas besoin de structures euclidiennes (choisir un critère F, des critères de groupes $F_i \to u_i \ldots$)

• transposition aux graphes est facile (EDO, massivement //)

Remarque : $\quad -\Delta u + |\nabla u|^2 = e^u(+\Delta)e^{-u}$

$$e^{u_i} \sum_j (e^{-u_j} - e^{-u_i}) = \sum_j (e^{u_i - u_j} - 1)$$

• détermination d'équilibres pour les réseaux sociaux : "distance sur les items" $\longleftrightarrow$

"distance sur les usagers" $\longleftrightarrow$ préférences $\longleftrightarrow$ "distances sur les items"...

MERCI!

pour votre attention… et votre patience!