

# Équipe STA

Évaluation AERES – 15 mars 2009

## Présentation de l'équipe

Composition

Objectifs

Rayonnement

Partenariats

Forces et faiblesses

Projet

Eclairage sur un nouveau thème

## Composition

**Télécom ParisTech (3 Pr.; 4 MC ; 1 DE)** M. Charbit, E. Moulines, F. Roueff; K. Abed Meraim , P. Bianchi, S. Cléménçon, J. Jakubowicz; G. Blanchet.

**CNRS (2 DR; 4 CR)** J-F. Cardoso, O. Cappé; G. Fort, A. Garivier, C. Lévy-Leduc, J. Najim.

**Doctorants (28)** A. Alaya, S. Barembruch, T. Ben Jabeur, H. Benoudnine, L. Berriche, M. Boulé, H. Bousbia-Slaah, N. Castaneda, J. Cornebise, M. Depecker, S. Filippi, J.F. Germain, F. Guilloux, Z. Harchaoui, I. Kacha, M. Karray, M. Kharouf, O. Kouamo, D. Lahat, A. Lung-Yut-Fong, N. Mahler, B. Mouhouche, G. Picard, T. Rebafka, L. Rigouste, N. Sokolovska, W. Soudene, T. Trigano.

**Post-Doc (4); Sabbatiques (5)** B. Benmammam, P. Etoré, J. Olsson, M. Zetlaoui, M. Taqqu, L. White, V. Reisen, S. Attalah.

# Mouvements entre 2005 et 2009

## Séjours sabbatiques externes

Juin 2007-Juin 2009 : K. Abed Meraim à l'Univ. de Sharjah.

## Recrutements de permanents [thème] (Origine)

<u>Oct 2005</u>	[Séries temporelles] C. Lévy Leduc (Paris 11)
<u>Oct 2007</u>	[Apprentissage statistique] S. Cléménçon , A. Garivier (Paris 10/ INRA, Paris 11/ CMLA)
<u>Dec 2008</u>	[Statistiques distribuées] P. Bianchi , J. Jakubowicz (Sup-elec, CMLA)

# Objectifs (1/3)

## Modélisation et inférence statistique

Fortes interactions avec les mathématiques

## pour le traitement de l'information

Une recherche méthodologique en prise avec les applications

## Thèmes de recherche

1. Apprentissage statistique : 4 ANR (KERNSIG, MGA, TAMIS, BEMOL).
2. Méthodes statistiques pour l'astrophysique: 2 ANR (COSMOSTAT, ECOSTAT).
3. Traitement statistique du signal: 2 ANR (MalCom, SESAME).
4. Méthodes de Monte Carlo: 2 ANR (ADAP'MC, BigMC).
5. Séries temporelles: 2 ANR (OSCAR, SARAH).

## Objectifs (2/3)

### Domaines d'application [thème(s)]

- Fouille de données
- Communications numériques
- Poursuite, localisation
- Analyse de données en cosmologie
- Traitement statistique de mesures pour la physique expérimentale
- Détection d'anomalies
- Modélisation aléatoire de réseaux quantiques

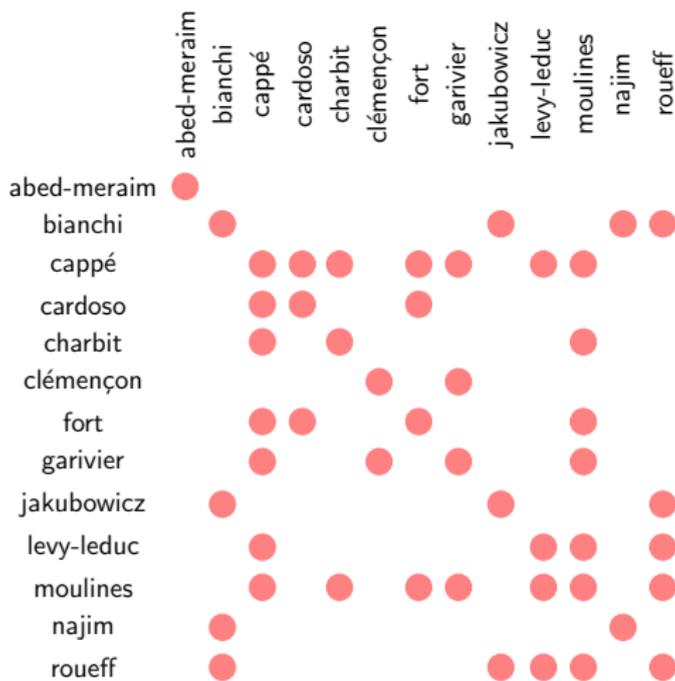
En particulier: 3 contrats bilatéraux (CSA, CS I2E), 8 thèses (FT R&D, Renault, CEA, DGA, Wavecom).

# Objectifs (3/3)

## Outils méthodologiques VS Thèmes

	<i>Appr. stat.</i>	<i>Stat. pour l'astro.</i>	<i>T. stat. du signal</i>	<i>Méthodes MC</i>	<i>Séries temp.</i>
Méthodes de Monte Carlo	●	●	●	●	●
Modèles à données latentes	●	●	●	●	●
Statistiques asymptotiques	●		●	●	●
Chaînes de Markov	●		●	●	●
Analyse en composante indépendantes		●	●		
Analyse en ondelettes		●			●
Grandes matrices aléatoires			●		
Théorie de l'apprentissage	●				
Modèles à mémoire longue					●

# Collaborations au sein de l'équipe



# Rayonnement

- **Forte production scientifique** : en moyenne **1.2** publications en revue et **1.6** en conférence par **personne** et par an sur la période
- **Un fort taux de publications en revues (43% du total)** : IEEE Trans. Signal Proc. (14); Stoch. Proc. App., Astronomy & Astrophysics (2x6); Annals Applied Probab., J. Time Series Analysis, Annals Statist. (3x5); Bernoulli (4); IEEE Signal Proc. Lett., Signal Proc., Eurasip JASP, MNRAS, Phys. Rev. D, IEEE Trans. Inform. Theory, Statist. & Comput., J. Statist. Planning Inference, Ann. IHP, IEEE Trans Image Proc. (2), ...
- Depuis 2007, présence régulière dans les conférences majeures de machine learning (6 NIPS, 2 ICML)

# Rayonnement

- Participations à
  - **Des équipes éditoriales** : Bernoulli, ESAIM P&S, Stoch. Proc. Appl. (E. Moulines), J. Royal Statist. Soc., Ser. B. (O. Cappé)
  - **Des comités scientifiques** des grandes conférences nationales et internationales en traitement du signal et en apprentissage statistique : IEEE ICASSP, IEEE SSP Workshop et ICML, NIPS, ECML, CAP.
- **Organisation de rencontres scientifiques** : IEEE Statistical Signal Processing Workshop (2005), New directions in Monte Carlo Methods (2007).
- **Exposés invités et tutoriaux** dans les workshops : ICSPC (K. Abed Meraim), IAP Colloquium (O. Cappé), New Developments in MCMC (G. Fort), SAMSI Sequential Monte Carlo Methods (E. Moulines), Physcomnet (J. Najim)...

# Partenariats (1/2)

## Partenaires académiques

- Au sein du LTCI : équipes COMNUM, AAO, TII, MIC<sup>2</sup>
- En région parisienne
  - Mathématiques appliquées : Univ. Paris 10 Nanterre (MODAL'X), Univ. Paris 7 Denis Diderot (LPMA), ENS Ulm (TREC).
  - Traitement statistique pour l'astrophysique : Univ. Paris 7 Denis Diderot (ADAMIS), Institut d'Astrophysique de Paris.
  - Communications numériques : Univ. Paris-Est (IGM).
  - Apprentissage statistique : Univ. Paris-Dauphine (Cérémade), École des Ponts (CERMICS), ENS Ulm (WILLOW) et Cachan (CMLA).
- Coorganisation des séminaires joints: *séminaire parisien de statistique, Statistical Machine Learning in Paris (SMILE), séminaire BIG'MC...*

## Partenariats (2/2)

### Principaux partenaires industriels

- France Télécom R&D (5 thèses)
- Renault (2 thèses)
- CEA (2 thèses)
- DGA (2 post-doc)
- Dans la période récente, également banques (1 postdoc, 1 thèse) et PME (3 thèses)

### Fort taux de succès à l'ANR

8 projets depuis 2005, principalement "programme blanc"

# Enseignement

## A Télécom ParisTech

Forte participation aux enseignements en probabilité, statistique et traitement du signal.

## A l'extérieur

- M2 *Modélisation aléatoire* de Paris 7
- M2 *Ingénierie Mathématique* de Paris 11 Orsay
- ParisTech : Ponts, Polytechnique, ENSAE
- M2 *Mathématiques, Vision Apprentissage* de l'ENS Cachan
- M2 *Systèmes Avancés de Radiocommunication* (Supélec)
- M1 *Mathématiques de la Modélisation et de la Décision* de Paris-Dauphine

# Forces et faiblesses : facteurs internes

## Forces

- Ambition d'excellence scientifique
- Forte reconnaissance
- Variété des partenariats
- Renouvellement thématique
- Capacité de financement (ANR, CIFRE)

## Faiblesses

- Difficultés de recrutement doctoral (vivier restreint, thèmes un peu marginaux dans l'école doctorale)
- Positionnement singulier au sein de l'Institut Télécom, manque de reconnaissance interne

# Forces et faiblesses : facteurs externes

## Opportunités

- Bon réseau francilien et forte attractivité susceptibles de favoriser le développement de l'équipe

## Incertitudes

- Environnement en très forte évolution qui remet en cause les partenariats existants dans une logique "top-down"
- Devenir de l'équipe dans l'hypothèse d'un déménagement sur Palaiseau.

# Projet (1/2)

## Thématiques historiques

- Apprentissage statistique: dynamique **en progression**.
- Méthodes statistiques pour l'astrophysique: axe **très actif** à renforcer par un recrutement CNRS ?
- Traitement statistique du signal: axe majeur pour les **partenariats industriels**.
- Méthodes de Monte Carlo: axe **transversal** essentiel.
- Séries temporelles: **compétences fortes** à maintenir.

# Projet (2/2)

## Thématiques émergentes

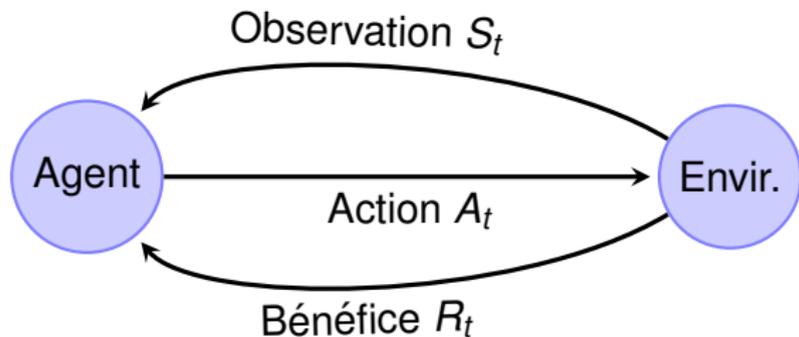
1. Processus de décision markoviens, apprentissage par renforcement (A. Garivier) : insertion naturelle en **apprentissage statistique** avec des applications et des outils **transversaux**: traitement statistique du signal, séries temporelles.
2. Traitement statistique collaboratif et distribué en grande dimension (P. Bianchi, J. Jakubowicz): en **synergie forte** avec la théorie des **matrices aléatoires**.

## Présentation de l'équipe

### Eclairage sur un nouveau thème

Qu'est ce que l'apprentissage par renforcement ?  
MDP et POMDP

# Apprentissage par Renforcement



dilemme  
exploration  
|  
exploitation

- L'agent est acteur et pas spectateur
- A chaque instant  $t$ , il choisit une action  $A_t \in A$  en fonction des observations et bénéfices passés  $(S_s, R_s)_{s < t}$  pour maximiser son bénéfice cumulé  $\sum_{t=1}^n R_t$
- Exemples: essais médicaux, robotique, proposition de contenu, publicité, internet mobile, ...

# Problèmes de bandits

- Environnement constant
- Conditionnellement aux actions  $(A_t)_{1 \leq t \leq n}$ , les bénéfices  $(R_t)_{1 \leq t \leq n}$  sont i.i.d. de moyenne  $\mu_{A_t}$
- But : jouer l'action  $a^*$  qui a le plus grand bénéfice moyen :

$$\mu_{a^*} = \max_{a \in A} \mu_a$$

- Mesure de performance : *regret cumulé*

$$R_n = \sum_{t=1}^n \mu_{a^*} - \mu_{A_t}$$

- Exemples : essais médicaux, channel sensing...

# Upper Confidence Bound

- Algorithmes optimistes :

*Fais comme si tu te trouvais dans l'environnement qui t'est le plus favorable parmi tous ceux qui rendent les observations assez vraisemblables*

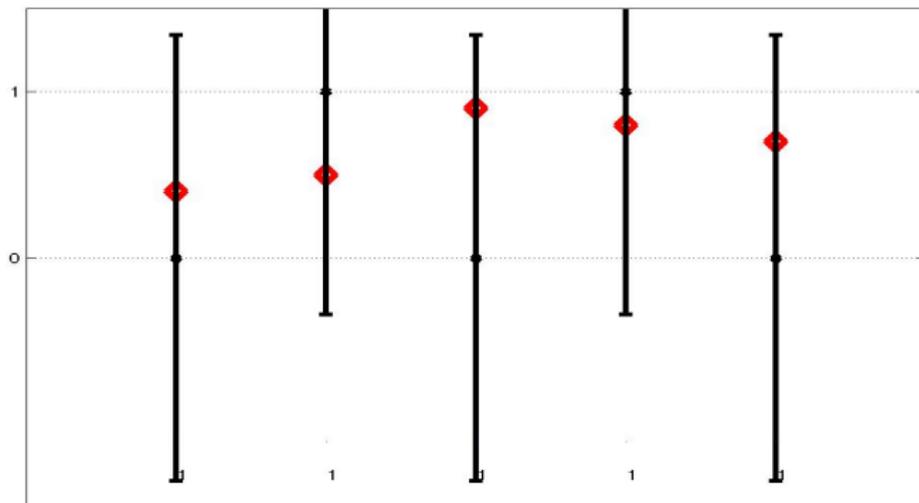
- Ici : UCB (Upper Confidence Bound) = établir une borne supérieure de l'intérêt de chaque action, et choisir celle qui est la plus prometteuse

⇒ le regret grandit comme  $C|A| \log n$ , où  $C$  dépend de

$$\Delta = \min_{\mu_a < \mu_{a^*}} \mu_{a^*} - \mu_a$$

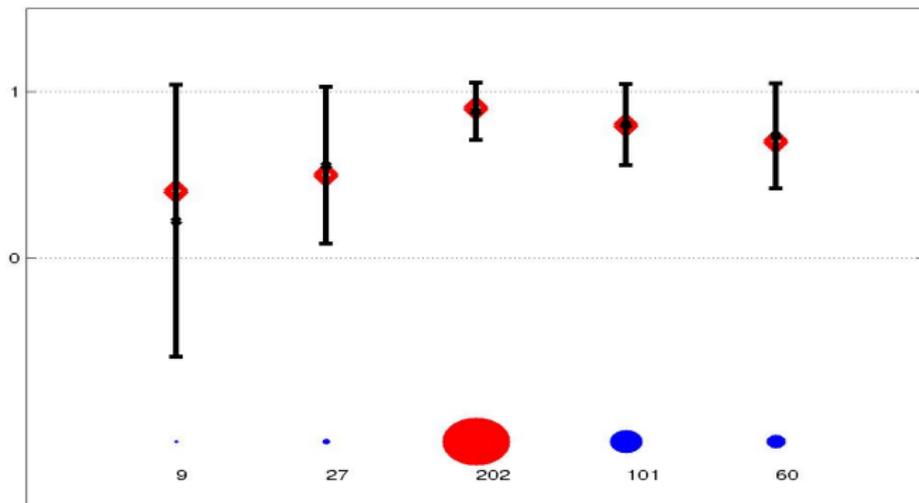
et c'est (presque) optimal

# Upper Confidence Bound



Début

# Upper Confidence Bound



Début

## “Discounted UCB”

- Environnement non stationnaire : l'intérêt moyen  $\mu_a$  de chaque action évolue brutalement dans le temps
- Motivation : channel sensing
- Idée simple D-UCB : introduire un facteur d'escompte dans l'estimation des performances
- Contribution : analyse de cet algorithme et proposition d'une variante

⇒ Equilibre biais-variance, étude de martingales auto-normalisées

# Contrôle des martingales auto-normalisées

- $(R_t)_{t \geq 1}$  = i.i.d. tq  $X_t \in [0, B]$   $\mathbb{P}$ -p.s. et  $\mu_t = \mathbb{E}[R_t]$ .
- $\mathcal{F}_t$  = filtration croissante tq  $\sigma(R_1 \dots, R_t) \subset \mathcal{F}_t$  et  $\forall s > t, R_s \perp \mathcal{F}_t$ .
- $(\epsilon_t)_{t \geq 1}$  = suite  $\mathcal{F}_t$ -prévisible de variables de Bernoulli,  $\gamma \in ]0, 1[$
- $S_t(\gamma) = \sum_{s=1}^t \gamma^{t-s} R_s \epsilon_s, \quad M_t(\gamma) = \sum_{s=1}^t \gamma^{t-s} \mu_s \epsilon_s$
- $N_t(\gamma) = \sum_{s=1}^t \gamma^{t-s} \epsilon_s, \quad n_t(\gamma) = \sum_{s=1}^t \gamma^{t-s}$

**Théorème:** pour tout  $t$  et tout  $\delta, \eta > 0$ ,

$$\mathbb{P} \left( \frac{S_t(\gamma) - M_t(\gamma)}{\sqrt{N_t(\gamma^2)}} > \delta \right) \leq \left[ \frac{\log n_t(\gamma)}{\log(1 + \eta)} \right] \exp \left( -\frac{2\delta^2}{B^2} \left( 1 - \frac{\eta^2}{16} \right) \right).$$

## “GLM - UCB”

- Problème de bandit avec information contextuelle  $m_a$  sur chaque action  $a$
- Modèle linéaire généralisé :

$$p_{\theta}(r|a) = \exp \left( r \langle \theta, m_a \rangle - b(\langle \beta, m_a \rangle) + c(r) \right)$$

- Permet de traiter les bénéfices binaires et entiers (régression logistique, Poissonnienne)
- Exemple : publicité sur internet
  - ⇒ le regret dépend de la *dimension du paramètre*, pas du nombre de bras

# Processus de Décision Markoviens

- Le système est dans un état  $S_t$  qui évolue de façon markovienne :

$$S_{t+1} = f(S_t, A_t, U_t) \text{ et } R_t = g(S_t, A_t)$$

- Exemples : channel sensing, robotique, etc.
- Même en connaissant les paramètres, la politique optimale n'est pas évidente (*planification*)

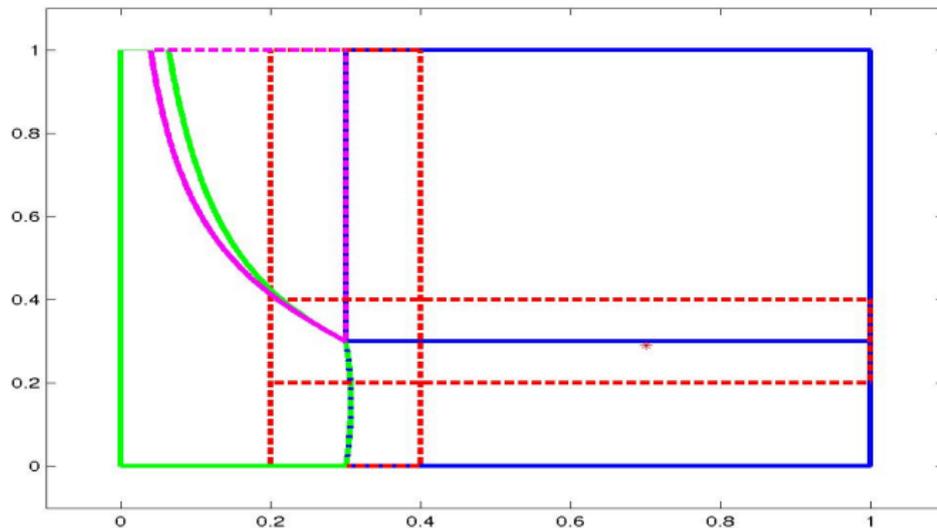
# “Küllback-Leibler UCB”

- Stratégie optimiste : à chaque instant
  - Considère l'ensemble de tous les MDP (transitions + lois des bénéfiques) qui rendent les observations assez vraisemblables
  - Trouve celui dont la valeur est la plus grande
  - Joue *pendant un certain temps* selon la politique optimale de ce MDP
- Voisinages du maximum de vraisemblance : utilisation de l'*information de Küllback-Leibler*
- Même type d'inégalités sur les sommes auto-normalisées

# POMDP : “Tiling Algorithm”

- L'état courant n'est pas directement observé
- Exemple : “optimal channel sensing” (on n'observe que le canal que l'on utilise)
- Cadre paramétrique, politique optimale connue selon la valeur des paramètres
- Pavage de l'espace des paramètres en *zones* et *frontières*
- Première phase : exploration jusqu'à la convergence dans une zone ou dans une frontière
- Deuxième phase : exploitation de la politique obtenue

# POMDP : “Tiling Algorithm”



# POMDP : “Tiling Algorithm”

