Politiques de bandits manchots et crédibilité dans les systèmes de réputation

Thibaut Vallée, Grégory Bonnet, François Bourdon

Normandie Université
GREYC, CNRS UMR 6072
F-14032 Caen, France
prénom.nom@unicaen.fr

RÉSUMÉ. *La robustesse des systèmes de réputation face aux manipulations a été largement étudiée. Cependant, les études portant sur l'utilisation des valeurs de réputation sont, elles, beaucoup plus rares. Dans cet article, nous proposons un modèle générique de système d'échange de services combinant système de réputation et bandits manchots. Nous étudions dans un premier temps l'effet des politiques de sélection des bandits manchots lorsque ces dernières sont utilisées sur les valeurs de réputation. Dans un second temps, nous définissons une mesure de crédibilité fondée sur la divergence de Kullback-Leibler afin de détecter et filtrer les faux témoignages. Cette mesure de crédibilité permet alors de définir des fonctions de filtrage pouvant être ajoutées à une fonction de réputation quelconque pour en accroître la robustesse.*

ABSTRACT. *Reputation systems robustness against manipulations has been widely studied. However, the studies of how to use the reputation values computed by those systems are rare. We propose an abstract model of service sharing system using a bandit-based reputation system. Firstly, we study how multi-armed bandit selection policies can be used in reputation system and their impact on robustness. Secondly, we propose a credibility measure based on Kullback-Leibler divergence to filter false feedbacks. This measure is used to define filtering function that can be added to any reputation function in order to increase its robustness.*

MOTS-CLÉS : *crédibilité, systèmes de réputation, manipulations, bandits manchots.*

KEYWORDS: *credibility, reputation System, manipulation, multi-armed bandits.*

DOI:10.3166/RIA.29.369-398

1. Introduction

Dans un système multi-agent lorsqu'un agent n'est pas capable de réaliser une tâche seul, il peut la déléguer à un autre agent. Pour cela, les agents doivent se partager leurs compétences et leurs connaissances. Dans un tel système, les agents sont à la fois fournisseurs et utilisateurs de services. Toutefois, si ces agents sont hétérogènes,

ils peuvent fournir des services de qualité variable, que cela soit dû à des erreurs de calcul, des défaillances ou encore à un comportement malveillant.

Afin d'aider les agents à décider avec qui ils doivent interagir [1], il est possible de mettre en place un système de réputation. Ce système permet à un agent de demander des services à un autre agent qui lui a été recommandé par des agents tiers. Pour cela, les agents évaluent leurs interactions passées et calculent une valeur représentant un niveau de confiance qu'ils ont envers les autres. Ces valeurs de confiance sont ensuite partagées entre les agents par le biais de témoignages et agrégées entre elles. Cette agrégation des témoignages produit une valeur de réputation pour chaque agent, qui est supposée refléter leur capacité respective à fournir des services de qualité. Malheureusement, dans un tel système ouvert et hétérogène, un agent malveillant peut manipuler sa valeur de réputation en fournissant de faux témoignages, en formant des coalitions avec d'autres agents malveillants, en introduisant de fausses identités (appelées Sybil) ou en changeant subitement de comportement.

De nombreux travaux proposent des systèmes de réputation robustes à des manipulations spécifiques. Cependant, ces études portent sur le calcul des valeurs de confiance et de réputation et non sur la manière dont les agents les utilisent. Or, la politique d'utilisation de ces valeurs a une influence sur le fonctionnement du système. En effet, si l'ensemble des agents décide de ne consommer des services qu'auprès des agents ayant la plus haute valeur de réputation, il sera difficile pour un agent malveillant seul de fournir des mauvais services sur le long terme. Cependant, une telle politique risque de surcharger de requêtes ces agents réputés et d'empêcher l'ouverture du système. Il est alors intéressant de remarquer que le problème qui consiste à décider avec qui interagir en utilisant des observations passées a été largement étudié dans le contexte des bandits manchots (*multi-armed bandits* ou MAB). Pour ce faire, les agents disposent de politiques qui leur permettent de décider avec qui interagir dans un compromis entre l'exploration pour obtenir de nouvelles observations et l'exploitation pour se servir des informations dont ils disposent. Or, cette problématique est la même dans un système de réputation où les agents cherchent à interagir majoritairement avec ceux ayant une bonne valeur de réputation tout en interagissant parfois avec d'autres agents afin d'obtenir des connaissances sur eux. Par ailleurs, les systèmes de réputation sont sensibles aux manipulations, en particulier les faux témoignages. La plupart des systèmes de réputation pondèrent alors les témoignages en fonction de la réputation des témoins. Cependant, cette approche confond fiabilité dans l'exécution d'un service et fiabilité dans la capacité à fournir des témoignages. C'est pourquoi des travaux récents proposent l'utilisation d'une valeur de crédibilité séparée de la réputation pour pondérer les témoignages (Malik, Bouguettaya, 2009 ; Koutrouli, Tsalgatidou, 2011).

Dans cet article, nous proposons de modéliser un système de réputation par des bandits manchots. Nous pouvons ainsi étudier comment l'utilisation des politiques MAB peut influencer la robustesse du système et ainsi permettre à un concepteur de

1. Deux agents interagissent lorsque l'un fournit un service à l'autre.

choisir la politique de sélection la plus adaptée. Par ailleurs, une modélisation d'un système de réputation inspirée de celle des bandits manchots nous permet de considérer les observations comme des variables aléatoires issues de fonctions de distribution de probabilité. Les propriétés de ces distributions nous permettent alors de définir une notion de crédibilité. Pour ce faire, nous proposons de calculer des différences entre témoignages et observations à l'aide d'une mesure de similarité fondée sur la divergence de Kullback-Leibler. Nous nous servons ensuite de cette mesure de crédibilité pour filtrer les témoignages.

Ainsi, nous proposons d'améliorer la robutesse des systèmes de réputation par (1) une politique pertinente de sélection et (2) un filtrage fondé sur la crédibilité. Nous présentons en section 2 l'état de l'art dans le domaine des systèmes de réputation et introduisons le problème des bandits manchots. En section 3, nous proposons un modèle général de système d'échange de services et faisons l'analogie entre ce modèle et celui des bandits manchots. Nous présentons en section 4 différentes politiques de sélection et stratégies de manipulation et en section 5 une étude empirique des performances de ces politiques lorsqu'une coalition d'agents malveillants manipule le système. Nous proposons en section 6 une notion de crédibilité fondée sur une divergence entre témoignages et observations ainsi que des fonctions de réputation utilisant cette crédibilité pour filtrer les faux témoignages et en section 7 comment ce filtrage influe sur la robustesse du système.

2. État de l'art

La notion de confiance a été introduite dans le contexte des systèmes multi-agents par (Marsh, 1994). Elle formalise une estimation du comportement futur d'un agent lorsqu'il existe un risque que celui-ci ait un comportement inattendu. (Resnick *et al.*, 2000) proposent trois axiomes fondamentaux pour définir un système de réputation : (1) les agents doivent interagir ensemble dans le futur ; (2) les agents doivent partager leurs valeurs de confiance par le biais de témoignages ; (3) ces témoignages doivent être utilisés par les agents pour déterminer à qui demander un service souhaité. Ainsi, la valeur de confiance d'un agent envers un autre est une évaluation des services que ce dernier a fourni au premier. La réputation d'un agent est, elle, une agrégation des valeurs de confiance des autres agents vis-à-vis de celui-ci.

De nombreux systèmes de réputation ont été proposés (Resnick *et al.*, 2000 ; Josang, Ismail, 2002 ; Kamvar *et al.*, 2003 ; Srivatsa *et al.*, 2005 ; Sabater *et al.*, 2006 ; Dini, Spagnolo, 2009 ; Hoffman *et al.*, 2009 ; Altman, Tennenholtz, 2010) et peuvent être classés en trois familles : symétrique (eBay (Resnick *et al.*, 2000)), assymétrique global (EigenTrust (Kamvar *et al.*, 2003)) et assymétrique personnalisé (BetaReputation (Josang, Ismail, 2002)). Une fonction de réputation est dite symétrique lorsque l'ordre d'aggrégation des témoignages n'influe pas sur la valeur de réputation des agents. De plus parmi les fonctions assymétriques, une fonction de réputation est dite globale (resp. personnalisée) lorsqu'une valeur de réputation ne dépend pas (resp. dépend) de celui qui la calcule. Parmi ces systèmes, trois sont particulièrement reconnus :

BetaReputation (Josang, Ismail, 2002), EigenTrust (Kamvar *et al*., 2003) et FlowTrust (Cheng, Friedman, 2005). BetaReputation utilise une loi bêta pour estimer la probabilité qu'un agent ait un bon comportement dans le futur. EigenTrust, qui a été démontré facilement manipulable par une coalition d'agents malveillants (Cheng, Friedman, 2006), se fonde sur le même principe que le Google PageRank : étant donné un graphe pondéré représentant les liens de confiance entre agents, la réputation d'un agent est la probabilité qu'un algorithme de marche aléatoire passe par le nœud de cet agent. Quant à FlowTrust, il définit la réputation des agents comme le flot maximum sur le graphe pondéré représentant les liens de confiance entre agents.

La robustesse des systèmes de réputation aux manipulations a été aussi fortement étudiée (Hoffman *et al*., 2009 ; Altman, Tennenholtz, 2010 ; Feldman *et al*., 2004 ; Cheng, Friedman, 2005 ; Altman, Tennenholtz, 2005). (Cheng, Friedman, 2005) ont prouvé qu'il n'existe pas de système de réputation symétrique robuste aux attaques Sybil [2] et que les systèmes assymétriques ne peuvent l'être que s'ils satisfont des conditions restrictives. (Altman, Tennenholtz, 2010 ; 2005) ont prouvé que la plupart des systèmes de réputation ne peuvent pas simultanément satisfaire un axiome de robustesse et d'autres axiomes fondamentaux. Les manipulations les plus couramment étudiées sont les faux témoignages et les blanchiments (Hoffman *et al*., 2009 ; Srivatsa *et al*., 2005). Pour lutter contre, la plupart des systèmes de réputation pondèrent les témoignages reçus par la réputation du fournisseur de ce témoignage. Par exemple, dans Eigentrust (Kamvar *et al*., 2003), un sous-ensemble d'agents qui sont supposés fiables a priori ont un poids plus important lors du calcul de la réputation. D'autres approches consistent à pondérer les témoignages des agents par un score correspondant à une affinité sociale entre eux (Sabater, Sierra, 2001). Cependant, certaines manipulations telles que l'attaque oscillante (Srivatsa *et al*., 2005) exploitent le fait que les agents ayant les meilleures réputations sont les plus à même de recommander d'autres agents. Des approches plus récentes consistent à différencier la capacité d'un agent à fournir un service et sa fiabilité dans les témoignages. Cette notion de fiabilité est appelée *crédibilité* (Selcuk *et al*., 2004 ; Zhao, Li, 2008 ; Malik, Bouguettaya, 2009 ; Koutrouli, Tsalgatidou, 2011 ; Noor *et al*., 2013). Par exemple, (Malik, Bouguettaya, 2009 ; Noor *et al*., 2013) pondèrent la réputation par, respectivement, la distance euclidienne entre les observations de l'agent et le témoignage moyen, ou une mesure de densité des témoignages. Cependant, cette pondération maintient un couplage entre fiabilité et crédibilité. De manière différente, (Zhao, Li, 2008) définissent un score de crédibilité par agent et filtrent les témoignages en éliminant ceux provenant d'agents ayant un score de crédibilité inférieur à un seuil. Cependant, ce seuil est un paramètre fixe donné a priori qui ne tient pas compte des connaissances de l'agent, ni de leur précision.

De plus, les études sur la robustesse des systèmes de réputation ne s'intéressent pas à la manière dont les valeurs de réputation sont utilisées. Il convient donc de

2. Cette manipulation consiste à s'introduire dans le système sous de multiples fausses identités pour constituer une coalition d'agents malveillants (Douceur, 2002).

se poser une question : comment utiliser les valeurs de réputation pour déterminer à qui demander un service souhaité ? Usuellement, dans les systèmes de réputation, un agent demande le service désiré à l'agent ayant la meilleure réputation, même si certains auteurs proposent des heuristiques probabilistes (Kamvar *et al.*, 2003). Ce problème de décision a été étudié dans un autre contexte, celui des bandits manchots (MAB) (Robbins, 1952). La définition canonique d'un problème MAB est la suivante : considérons une machine à sous avec plusieurs bras, chacun ayant une fonction de gain suivant une loi de distribution a priori inconnue, quelle séquence de bras un agent doit-il tirer afin de maximiser son gain ? De nombreux modèles de MAB existent : à plusieurs joueurs (Liu, Zhao, 2010), à fonction de gain stationnaire ou non (Koulouriotis, Xanthopoulos, 2008), à possibilité de tirer plusieurs bras simultanément (Anantharam *et al.*, 1987), ou même avec adversaire (Auer *et al.*, 1995). Dans tous les cas, l'agent dispose d'une politique de sélection lui permettant de minimiser son regret, c'est-à-dire la différence entre le gain obtenu et le gain qu'aurait eu l'agent si, à chaque pas de temps, il avait choisi le meilleur bras. Toutes ces politiques, telles que UCB, Poker, ε-glouton (Vermorel, Mohri, 2005 ; Auer, Ortner, 2010), sont des compromis entre l'exploitation et l'exploration.

Dans cet article, nous faisons l'analogie entre les deux problèmes de décisions : la sélection d'agents évalués par un système de réputation et la sélection de bras évalués par une estimation de leur fonction de gain. Ce lien a déjà été mis en évidence par (Awerbuch, Kleinberg, 2005) mais uniquement en se concentrant sur la question de la mise à jour des valeurs de confiance. Pour notre part, cette approche nous permet d'étudier deux manières d'accroître la robustesse des systèmes de réputation. La première consiste à se servir d'une politique de sélection de bandits manchots afin de décider avec qui interagir. La seconde consiste à se servir des estimations de gain des agents pour définir une mesure de crédibilité des témoignages. Comme (Malik, Bouguettaya, 2009), nous considérons qu'un témoignages est crédible s'il est similaire aux observations des autres agents. Cependant, nous proposons une mesure de dissimilarité fondée sur la divergence de Kullback-Leibler (Kullback, 1997) pour définir la crédibilité, et contrairement aux approches qui pondèrent les témoignages, notre mesure de crédibilité est associée aux témoignages. De plus, contrairement aux approches qui se fondent sur un seuil donné a priori, nous proposons une notion de seuil dépendante des connaissances de l'agent qui évalue la crédibilité. Enfin, nous proposons des fonctions de filtrage pour écarter les témoignages non crédibles.

3. Modèle formel

Afin d'aider la lecture de cet article, le tableau 1 résume nos principales notations.

3.1. Analogie avec un bandit manchot

Considérons un système multi-agent où chaque agent peut fournir des services et demander à d'autres de lui en fournir. Afin de ne pas perdre en généralité, nous considérons ces services comme abstraits. Un tel système est appelé un *système d'échange*

Tableau 1. Notations

Système d'échange de services	
N :	Ensemble des agents
S :	Ensemble des services
N_x :	Ensemble des agents pouvant fournir le service s_x
$\varepsilon_{i,x}$:	Expertise de l'agent a_i pour le service s_x
v_i :	Fonction d'évaluation de l'agent a_i
π_i :	Politique de sélection de l'agent a_i
Système de réputation	
f_i :	Fonction de réputation de l'agent a_i,
$O_{i,k,x}$:	Ensemble des observations de l'agent a_i vis-à-vis de $\varepsilon_{k,x}$
$F_{i,j,k,x}$:	Témoignages de l'agent a_j fournis à a_j vis-à-vis de $\varepsilon_{k,x}$
$\mathcal{F}_i$:	Ensemble des témoignages et des observations de l'agent a_i
Fonction de filtrage	
ϕ_i :	Fonction de filtrage de l'agent a_i
$D_{i,j,k,x}$:	Divergence de Kullback-Leibler entre $O_{i,k,x}$ et $F_{i,j,k,x}$
$\mu_{i,k,x}$:	Moyenne des valeurs de $O_{i,k,x}$
$\sigma_{i,k,x}$:	Écart-type des valeurs de $O_{i,k,x}$
$\mu_{i,j,k,x}$:	Moyenne des valeurs de $F_{i,j,k,x}$
$\sigma_{i,j,k,x}$:	Écart-type des valeurs de $F_{i,j,k,x}$

de services : lorsqu'un agent a besoin d'un service qu'il ne peut pas réaliser lui-même, il demande à un autre agent de le lui fournir. Son objectif est alors de recevoir le service désiré avec la meilleure qualité possible.

DÉFINITION 1. — *Un* système d'échange de services *est un tuple $\langle N, S\rangle$ où N est l'ensemble des agents et S l'ensemble des services qui peuvent être fournis. Notons par $N_x \subseteq N$ l'ensemble des agents capables de réaliser le service $s_x \in S$.*

Dans la suite de cet article, nous écrirons *système* pour système d'échange de services. Ce système est supposé ouvert, c'est-à-dire qu'à tout moment un agent peut le quitter (cesser de participer à l'échange de service en étant retiré de N) et de nouveaux agents peuvent le rejoindre (en étant ajouté dynamiquement à N) afin de fournir ou consommer des services. À chaque fois qu'un agent a besoin d'un service, il doit décider à quel autre agent le demander. Ce problème est similaire à celui des bandits manchots. Le tableau 2 résume cette analogie entre système d'échange de services et MAB. Pour cela, considérons un joueur et une machine à sous ayant plusieurs bras. Chacun de ces bras est associé à une fonction de gain dont la loi de probabilité est a priori inconnue. Le problème est alors de décider quelle séquence de bras tirer afin de maximiser le gain cumulé. Considérons un agent $a_i \in N$ et un service $s_x \in S$. L'agent a_i peut modéliser son problème de sélection de fournisseur par un bandit manchot m_x où il associe un bras $m_{x,k}$ à chaque agent $a_k \in N_x$. L'espérance de gain du bras $m_{x,k}$ (inconnue par a_i) représente la capacité de l'agent a_k à fournir le service s_x. Demander le service s_x à l'agent a_k correspond alors à tirer le bras $m_{x,k}$. Dans un bandit manchot tout comme dans le système d'échange de service, les expériences

passées sont utilisées pour estimer la qualité d'un service futur (le gain) d'un agent (d'un bras) s'il est sélectionné. Dans le cadre des systèmes d'échange de services, les agents peuvent aussi échanger des informations et approximer l'estimation du gain espéré par une fonction de réputation. Sous hypothèse de corrélation entre réputation d'un agent et estimation de l'espérance de gain d'un bras, les témoignages du système de réputation sont les observations du bandit manchot.

SUPPOSITION 2. — $\forall\, a_i, a_j, a_k \in N$ et $\forall\, s_x \in S$ si la réputation de a_j selon a_i pour le service s_x (notée $f_i(a_j, s_x) \in \mathbb{R}$) est supérieure à la réputation de a_k selon a_i pour le service s_x (notée $f_i(a_k, s_x) \in \mathbb{R}$) alors l'estimation du gain espéré du bras $m_{x,j}$ est supérieure que celle du bras $m_{x,k}$. □

Cependant, certains agents, appelés agents malveillants, peuvent volontairement fournir des services de mauvaise qualité (par exemple fournir un virus). C'est pourquoi il nous faut considérer dans le MAB la présence d'adversaires qui choisissent le gain fourni par leurs bras (Auer *et al.*, 1995).

Tableau 2. Analogie entre système d'échange de services et MAB

	Système d'échange de services	**MAB**
Objectif	Maximiser la qualité des services reçus	Maximiser le gain
Acteurs	Consommateurs et fournisseurs	Joueurs et bandits
Interactions	Demander un service	Tirer un bras
Capacité	Expertise	Fonction de distribution des gains
Gain	Qualité d'un service	Gain d'un bras
Observations	Évaluation de la qualité	Observations passées
Communication	Témoignage sur un autre agent	Témoignage sur un bras
Réputation	Comportement futur espéré	Gain espéré
Politique de selection	Déterminer le futur fournisseur de service	Déterminer le futur bras à utiliser
Manipulations	Agents malveillants	Adversaire

3.2. Modèle formel

Nous définissons donc un agent dans un système d'échange de services ainsi :

DÉFINITION 3. — *Un* agent $a_i = \langle \overrightarrow{\varepsilon_i}, v_i, \mathcal{F}_i, f_i, \pi_i \rangle$ *est une entité autonome qui peut fournir et recevoir des services où* $\overrightarrow{\varepsilon_i} \in \mathbb{R}^{|S|}$ *désigne un vecteur d'expertise ;* v_i *une fonction d'évaluation ;* $\mathcal{F}_i$ *un ensemble de témoignages;* f_i *une fonction de réputation ;* π_i *une politique de sélection.*

Pour un service $s_x \in S$, l'*expertise* de l'agent $a_k \in N_x$, notée $\varepsilon_{k,x} \in \mathbb{R}$, est sa capacité à fournir le service s_x avec une bonne qualité lorsqu'un autre agent le lui demande. Même si la qualité d'un service dépend de l'expertise de son fournisseur, elle est sujette à l'évalutation du demandeur. Cette évalutation peut être fondée sur de nombreux facteurs et est donc subjective. Par exemple, supposons que l'agent a_i demande à l'agent a_k de convertir un fichier *.doc* en fichier *.pdf* car il ne dispose pas d'une fonction *doc2pdf*. La qualité de ce service peut être fondée sur le fait de recevoir le fichier sans erreur d'encodage, mais aussi sur le temps mis par a_k pour le lui fournir. Afin de rester général, nous considérons qu'un agent $a_i \in N$ ayant demandé le ser-

vice $s_x \in S$ à l'agent $a_k \in N_x$ à l'instant t reçoit une *observation* $v_i(a_k, s_x, t) \in V_x$ où v_i désigne la fonction d'évaluation de l'agent a_i et V_x une échelle d'évaluation commune à l'ensemble des agents pour le service s_x. Pour simplifier la lecture, nous notons $v^t_{i,k,x}$ pour désigner $v_i(a_k, s_x, t)$. L'objectif d'un agent $a_i \in N$ qui désire le service $s_x \in S$ est de le recevoir avec la meilleure qualité possible. Pour cela, a_i peut utiliser ses observations sur le système afin de demander le service à un agent $a_k \in N_x$ en lequel il aurait le plus confiance. Dans toute la suite, nous notons $O_{i,k,x}$ l'ensemble des observations de l'agent a_i pour le service s_x fourni par a_k. Nous supposons a priori que les observations des agents sont sans erreur. Cependant, le cas où les observations sont incertaines est en partie pris en compte lors de l'utilisation de la mesure de crédibilité présentée en section 6.1.

Comme il est possible que les agents aient individuellement peu d'observations sur les autres agents, ils peuvent partager ces dernières par le biais de témoignages. Un agent a_j peut fournir à un agent a_i à propos du service s_x rendu par a_k un *témoignage* noté $F_{i,j,k,x}$. Ce témoignage est équivalant aux observations de a_j ($F_{i,j,k,x} = O_{j,k,x}$) sauf en cas de manipulation comme présenté en section 4.2. Nous notons $\mathcal{F}_i$ l'union des observations de a_i et de l'ensemble des témoignages qu'il a reçu pour l'ensemble des services. L'agent a_i peut alors utiliser $\mathcal{F}_i$ pour estimer l'expertise d'un agent pour un service donné. Cette estimation est la *réputation* de l'agent pour ce service. Nous supposons que chaque agent dispose d'une *fonction de réputation* $f_i : N \times S \times 2^{\mathcal{F}} \rightarrow \mathbb{R}$ (par abus de notation $\mathcal{F}$ est l'ensemble des témoignages possibles) qui calcule la réputation des agents. Cette fonction abstraite doit être instanciée et nous présentons en section 3.3 les différentes fonctions que nous considérons dans cet article.

Un agent $a_i \in N$ qui désire le service $s_x \in S$ doit ensuite décider à quel autre agent le demander. La *politique de sélection* de l'agent a_i $\pi_i : S \rightarrow N$ permet à a_i de choisir un fournisseur pour le service s_x à partir des valeurs de réputation de tous les agents. Cette politique doit être instanciée et nous présentons en section 4.1 les différentes politiques que nous considérons dans cet article. Ainsi, l'architecture générale du système est résumée sur la figure 1. Débutant par un besoin de service (centre de la figure), chaque agent utilise sa politique de sélection π_i pour déterminer à quel agent le demander (flèche 1). Une fois ce service reçu (flèche 2), l'agent l'évalue et met à jour ses observations. Il fournit ensuite ces observations comme témoignages aux autres agents (flèche 5) et obtient à son tour des témoignages (flèche 6). La fonction de réputation f_i agrège les observations de l'agent avec les témoignages reçus afin d'obtenir une estimation de l'expertise des agents. Parallèlement, lorsqu'un agent reçoit une demande de service (flèche 3), il le fournit s'il en est capable (flèche 4).

Afin d'étudier comment des politiques de sélection et une notion de crédibilité influent sur la robustesse du système, nous considérons différentes instanciations de la fonction de réputation.

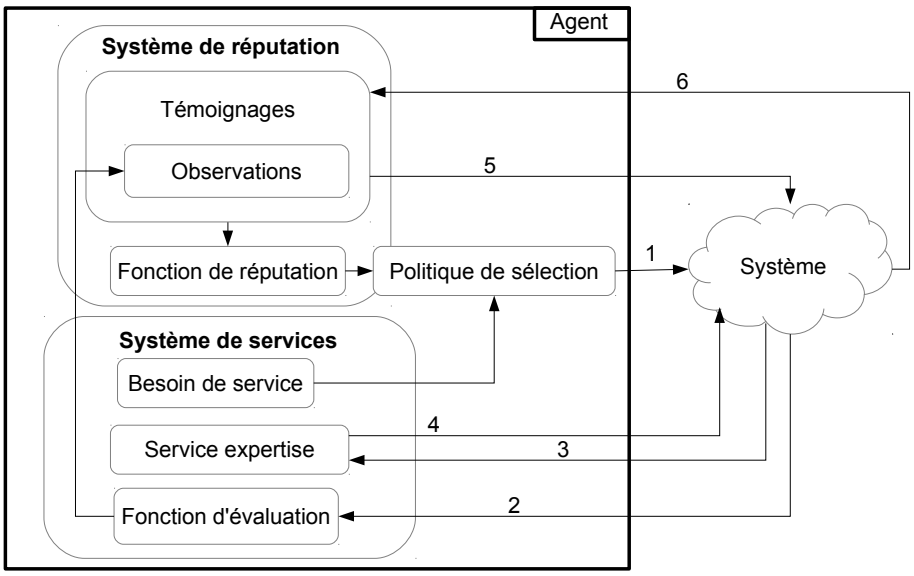

Figure 1. Système d'échange de services

3.3. Fonctions de réputation

Il convient de définir comment est calculée la réputation d'un agent. Il existe de nombreuses fonctions de réputation. Une fonction classique dans les problèmes de bandits manchots, que nous appellons estimation personnelle, consiste à ne pas prendre en compte les témoignages et de considérer la réputation d'un agent comme le gain moyen des services qu'il a fournis.

DÉFINITION 4. — *L'*estimation personnelle *de a_i est la fonction de réputation $f_i(a_k, s_x, \mathcal{F}_i) = \mu_{i,k,x}$ où $\mu_{i,k,x}$ est la moyenne de $O_{i,k,x}$.*

Une autre fonction de réputation triviale, que nous appellons estimation collective, consiste à prendre en compte tous les témoignages et considérer la réputation comme le gain moyen calculé à partir des observations et des témoignages.

DÉFINITION 5. — *L'*estimation collective *de a_i est la fonction de réputation $f_i(a_k, s_x, \mathcal{F}_i) = \mu_{N,k,x}$ où $\mu_{N,k,x}$ désigne la moyenne des observations présentes dans l'ensemble $\bigcup_{a_j \in N} F_{i,j,k,x}$.*

Dans cet article, nous considérons aussi trois fonctions de réputation classiquement étudiées dans la littérature : EigenTrust (Kamvar *et al.*, 2003), BetaReputation (Josang, Ismail, 2002) et FlowTrust (Feldman *et al.*, 2004). Ces trois fonctions sont fondées sur un *graphe de confiance* calculé à partir des témoignages : un graphe orienté dont les arcs (a_i, a_j) représentent le fait qu'il y ait eu au moins un service fourni à a_i par a_j et sont étiquettés par les observations de a_i envers a_j. La définition formelle de ces systèmes telles que nous les avons implantés est donnée ci-dessous.

BetaReputation considère la réputation d'un agent comme une espérance de gain modélisée par une loi de bêta densité. C'est une approche assymétrique où chaque agent communique à ses voisins dans le graphe de confiance la réputation qu'il a calculée et pondère les témoignages reçus par la réputation des agents qui les fournissent.

DÉFINITION 6. — *Soit $\mathbb{P}_i$ l'ensemble des chemins entre a_i et a_k sur un graphe de confiance, $r_{i,k,x}$ (resp. $s_{i,k,x}$) le nombre de bon (resp. mauvais) services*[3] *s_x fournis à a_i par a_k. La fonction* BetaReputation *de a_i est $f_i(a_k, s_x, \mathcal{F}_i) = \frac{r_{i,k,x} - s_{i,k,x}}{r_{i,k,x} + s_{i,k,x} + 2}$ où :*

$$r_{i,k,x} = \sum_{P \in \mathcal{P}_{i,k}} \prod_{(a_j, a_{j'}) \in P} \frac{2 r_{j,j',x} r_{j',k,x}}{(s_{j,j',x} + 2)(r_{j',k,x} + s_{j',k,x} + 2) + 2 r_{j,j',x}}$$

$$s_{i,k,x} = \sum_{P \in \mathcal{P}_{i,k}} \prod_{(a_j, a_{j'}) \in P} \frac{2 r_{j,j',x} s_{j',k,x}}{(s_{j,j',x} + 2)(r_{j',k,x} + s_{j',k,x} + 2) + 2 r_{j,j',x}}$$

Dans EigenTrust, la réputation d'un agent est la probabilité qu'un marcheur aléatoire sur le graphe de confiance atteigne le nœud qui lui est associé où la probabilité de transition d'un nœud à l'autre est proportionnelle à la différence entre les bons services et les mauvais services fournis.

DÉFINITION 7. — *Soit $\mathbb{C}$ la matrice d'adjacence d'un graphe de confiance. La fonction* EigenTrust *de a_i est : $f_i(a_k, s_x, \mathcal{F}_i) = (1 - a)(\mathbb{C}^T)^n \vec{c_i} + a p_k$, où $a \in [0, 1]$ est un facteur d'exploration, $\vec{c_i}$ le vecteur normalisé de confiance de a_i dans lequel $c_{i,k,x} = \max(0, r_{i,k,x} - s_{i,k,x})$ et p_k une valeur de confiance a priori envers a_k.*

Dans FlowTrust, la réputation d'un agent a_k selon un agent a_i est le flot maximum de a_i vers a_k sur le graphe de confiance où la capacité d'un arc (a_i, a_j) est la moyenne des gains $\mu_{i,k,x}$ de a_i.

DÉFINITION 8. — *Soit $\mathcal{P}_{i,k}$ l'ensemble des* chemins disjoints *entre a_i et a_k sur un graphe de confiance et $\mu_{i,k,x}$ la moyenne des $O_{i,k,x}$. La fonction* FlowTrust *de a_i est : $f_i(a_k, s_x, \mathcal{F}_i) = \sum_{P \in \mathcal{P}_{i,k}} \min\{\mu_{j,j',x} | (a_j, a_{j'}) \in P\}$.*

4. Stratégies honnêtes et stratégies de manipulation

4.1. Politiques de sélection

Si la réputation d'un agent est une estimation du gain espéré lors des futures interactions, il convient de définir comment un agent l'utilise pour maximiser son gain.

DÉFINITION 9. — *Soit $a_i \in N$ un agent désirant recevoir le service $s_x \in S$. La* politique de sélection *Π_i définit à quel agent $a_k \in N_x$ demander ce service.*

3. Un bon service est un service ayant un gain positif et un mauvais service un service ayant un gain négatif.

Nous proposons ici d'utiliser les politiques canoniques des bandits manchots comme politiques de sélection d'un système d'échange de services. Ces politiques permettent aux agents d'interagir majoritairement avec les fournisseurs de services ayant une bonne réputation (puisqu'ils sont supposés être ceux maximisant l'espérance de gain) et d'interagir occasionnellement avec les autres agents afin de vérifier si leur mauvaise réputation n'est pas due à un manque d'observations. Ce sont donc des compromis entre l'exploitation des connaissances des agents et l'exploration du système permettant d'affiner ces connaissances.

Nous adaptons ici deux d'entre elles, UCB (*Upper Confidence Bound*) et la politique ε-gloutonne, et en proposons une troisième : l'ε-élitisme. Il existe bien entendu d'autres politiques mais nous considérons celles-ci car UCB est une des politiques les plus performantes, ε-glouton est une politique naïve qui sert classiquement de point de comparaison et l'ε-élitisme est la politique implicitement utilisée dans les systèmes de réputation.

La première est l'une des plus utilisées dans le cadre des bandits manchots. En effet, UCB permet de borner le regret des agents (différence entre le gain total obtenu et le gain maximum si le meilleur bras avait toujours été sélectionné) (Robbins, 1952 ; Auer *et al.*, 1995). Dans le cadre des systèmes de réputation, UCB consiste à demander un service à l'agent maximisant la réputation à laquelle s'ajoute un facteur d'exploration. Ce facteur a pour objectif d'inciter les agents à interagir avec ceux sur lesquels il dispose de peu d'information. UCB garantit ainsi un minimum d'interactions avec chaque fournisseur afin que la fonction de réputation retourne une estimation suffisamment précise de l'expertise. Il est important de noter que ceci permet de maintenir la propriété d'ouverture du système en incitant à interagir avec tout nouvel arrivant.

DÉFINITION 10. — *L'agent $a_i \in N$ désirant le service $s_x \in S$ suit la politique* UCB *s'il sélectionne l'agent $a_k \in N_x$ qui maximise $f_i(a_k, s_x, \mathcal{F}_i) + \sqrt{\frac{2\ln(1+n_x)}{1+n_{k,x}}}$ où $n_{k,x}$ est le nombre de fois que l'agent a_k a fourni le service s_x à a_i et n_x le nombre de fois que a_i a reçu le service s_x.*

La politique ε-gloutonne consiste à demander le service désiré à l'agent capable de le fournir avec la meilleure valeur de réputation, tout ayant une certaine probabilité d'explorer uniformément le système (Auer *et al.*, 2002). Ce facteur d'exploration est la probabilité de sélectionner aléatoirement uniformément un autre agent que celui maximisant la valeur de réputation. Comme UCB, la politique ε-gloutonne garantit une propriété d'ouverture du système mais, contrairement à UCB, sans pour autant avantager les nouveaux entrants.

DÉFINITION 11. — *L'agent $a_i \in N$ suit une politique* ε-gloutonne *s'il demande le service $s_x \in S$ à l'agent $a_k \in N_x$ qui maximise $f_i(a_k, s_x, \mathcal{F}_i)$ avec une probabilité $1-\varepsilon$ ou, avec une probabilité ε, sélectionne aléatoirement uniformément a_k dans N_x.*

Nous proposons une troisième politique appelée ε-élitisme. Un agent suivant cette politique sélectionne le futur fournisseur de services uniformément parmi les $\lceil \varepsilon \times |N_x| \rceil$ agents de N_x ayant les meilleures valeurs de réputation. Contrairement aux

politiques précédentes, la politique ε-élitiste ne respecte pas la propriété d'ouverture du système car il n'y a pas de facteur d'exploration permettant d'éventuellement sélectionner des nouveaux entrants mais elle permet de ne pas surcharger de demandes l'agent ayant la meilleure réputation. Nous pouvons remarquer que les systèmes de réputation appliquent classiquement une politique purement élitiste (soit $\frac{1}{|N_x|}$-élitiste).

DÉFINITION 12. — *Soit $a_i \in N$ un agent désirant recevoir le service $s_x \in S$. Soit $N_{x,\varepsilon} \subseteq N_x$ tel que $|N_{x,\varepsilon}| = \lceil \varepsilon \times |N_x| \rceil$ et que $\forall a_j \in N_{x,\varepsilon}, \nexists a_k \in N_x \setminus N_{x,\varepsilon}$: $f_i(a_j, s_x, \mathcal{F}_i) < f_i(a_k, s_x, \mathcal{F}_i)$. L'agent a_i suit une politique* ε-élitiste *s'il sélectionne aléatoirement uniformément a_k dans $N_{x,\varepsilon}$.*

4.2. Agents malveillants et manipulations

L'utilisation de systèmes de réputation dans un système d'échange de services a pour objectif de garantir aux agents de recevoir les services avec la meilleure qualité. Cependant, dans un système ouvert, certains agents (appelés agents malveillants) peuvent avoir comme objectif de fournir de mauvais services. Par exemple, dans un système pair-à-pair d'échange de fichiers tel que Gnutella (Ripeanu, 2001), un agent malveillant peut chercher à propager des virus.

DÉFINITION 13. — *Soit un agent malveillant $a_j \in N$ ayant une expertise $\varepsilon_{j,x}$. a_j fournit un bon service s_x s'il le fournit avec une qualité correspondante à son expertise. a_j fournit un mauvais service s_x s'il le fournit avec une qualité* $\min(V_x)$.

Si un tel agent malveillant peut être détecté par des systèmes de réputation, plusieurs agents malveillants peuvent former une coalition afin de manipuler le système. Par ailleurs, un agent malveillant seul peut s'introduire dans le système sous de multiples fausses identités (appelées agents Sybil (Douceur, 2002)) et ainsi former une coalition malveillante. Nous considérerons ici les trois types de manipulations classiquement utilisés dans la littérature : les faux témoignages, le blanchiment et l'attaque oscillante (Hoffman *et al.*, 2009).

Comme la réputation des agents est fondée sur l'utilisation de témoignages, une manipulation consiste à fournir de faux témoignages. Nous considérons deux types de faux témoignages : la promotion et la diffamation. La première consiste à fournir des témoignages afin d'augmenter artificiellement la valeur de réputation d'un agent. À l'inverse, la diffamation a pour objectif de diminuer la réputation de l'agent. Dans les deux cas, si la manipulation est efficace, les agents malveillants apparaîtront comme les meilleurs fournisseurs, leur permettant ainsi de fournir de mauvais services.

DÉFINITION 14. — *Soit un agent malveillant $a_j \in N$. Soit un service $s_x \in S$ et deux agents $a_i \in N$ et $a_k \in N_x$. L'agent a_j fournit un* faux témoignage *à l'agent a_i vis-à-vis de l'expertise de a_k pour le service s_x s'il lui communique des témoignages $F_{i,j,k,x}$ tels que $F_{i,j,k,x} \neq O_{j,k,x}$. Soit $\mu_{j,k,x}$ la moyenne des véritables observations de a_j (fondée sur $O_{j,k,x}$) et $\mu_{i,j,k,x}$ la moyenne des observations fournies en témoignage (fondée sur $F_{i,j,k,x}$).*

– *si $\mu_{j,k,x} < \mu_{i,j,k,x}$ alors l'agent a_j* promeut *a_k,*

– *si* $\mu_{j,k,x} > \mu_{i,j,k,x}$ *alors* a_j diffame a_k.

Comme nous considérons un système multi-agent ouvert, si un agent malveillant a une valeur de réputation trop faible, il lui est possible de quitter le système pour revenir sous une autre identité. Cette manipulation, appelée blanchiment, a pour objectif de réinitialiser la réputation de l'agent en obtenant la même réputation qu'un nouvel agent qui vient de joindre le système pour la première fois.

DÉFINITION 15. — *Soit un agent malveillant* $a_j \in N$. a_j *effectue un* blanchiment *s'il quitte le système pour le rejoindre sous une autre identité* a'_j.

Nous considérons ici que les agents malveillants peuvent changer d'identité à volonté. Notons qu'il est difficile d'empêcher une telle manipulation tout en satisfaisant la propriété d'ouverture du système. Cependant, certaines approches telles que l'utilisation de *captchas*, de frais d'inscription ou de puzzles cryptographiques permettent de rendre le blanchiment coûteux (Borisov, 2006).

Si la promotion, la diffamation et le blanchiment peuvent être effectués par un agent seul en un court instant, une coalition d'agents malveillants peut également manipuler le système d'échange de services sur le long terme. L'une de ces manipulations est l'attaque oscillante. Dans une telle manipulation, la coalition d'agents malveillants est partitionée en deux sous-ensembles M_1 et M_2. Ces sous-groupes ont alors un comportement coordonné. Les agents du premier groupe fournissent des services de bonne qualité afin de bénéficier d'une bonne valeur de réputation. Dans le même temps, ils promeuvent les agents du second groupe afin d'accroître la réputation de ces derniers. Les agents du second groupe fournissent quant à eux volontairement de mauvais services et diffament les agents honnêtes.

Lorsque la réputation d'un agent du second groupe tombe en dessous de celle d'un des agents du premier groupe, ils échangent leurs rôles : l'agent de M_1 fournit désormais les mauvais services et diffame tandis que celui du groupe M_2 fournit de bons services et promeut. Notons que l'attaque oscillante peut être combinée avec du blanchiment au moment où les agents de M_1 et de M_2 changent de rôle. Dans cet article, nous considérons l'attaque oscillante suivante :

DÉFINITION 16. — *Soit* $M \subset N$ *une coalition d'agents malveillants. Soit* M_1 *et* M_2 *un partitionnement de* M. *La coalition* M *effectue une* attaque oscillante *en appliquant la stratégie suivante :*

– *les agents de* M_1 *promeuvent ceux de* M_2,

– *les agents de* M_2 *diffament les agents de* $N \setminus M$,

– *les agents de* M_1 *fournissent les services en fonction de leur expertise,*

– *les agents de* M_2 *fournissent volontairement de mauvais services,*

– *si la réputation d'un agent de* M_2 *est inférieure à celle d'un agent de* M_1, *l'agent de* M_2 *se blanchit et intègre* M_1 *tandis que l'agent de* M_1 *intègre* M_2.

5. Impact des politiques sur la robustesse des fonctions de réputation

5.1. Mesures expérimentales

Afin d'évaluer les politiques de sélection face aux manipulations, nous considérons deux mesures : le regret des agents honnêtes et le coût de la manipulation. Le regret d'un agent est une mesure classique dans les problèmes de bandits manchots. Intuitivement, le regret d'un agent désigne la différence entre le gain qu'il aurait pu gagner s'il avait toujours demandé les services aux meilleurs fournisseurs et le gain qu'il a réellement obtenu en suivant sa politique de sélection. Le regret des agents peut donc être vu comme une mesure d'efficacité du système puisque minimiser le regret et maximiser le gain des agents est équivalent.

DÉFINITION 17. — *Soit un agent $a_i \in N$ ayant un ensemble d'observations $O_i = \{v^1_{i,k_1,x_1}, \ldots, v^n_{i,k_n,x_n}\}$. Soit $\varepsilon^t_{\star,x}$ l'expertise du meilleur fournisseur du service s_x à l'instant t. Le* regret *de a_i est donné par* $r_i = \sum_{t=1}^{n} \varepsilon^t_{\star,x_t} - v^t_{i,k_t,x_t}$.

Comme certaines manipulations telles que l'attaque oscillante impliquent que les agents malveillants fournissent parfois de bons services (ce qui est contraire à leur objectif) afin de maintenir une haute valeur de réputation, nous considérons le coût de la manipulation comme le ratio de bons services fournis par les agents malveillants sur l'ensemble des interactions passées.

DÉFINITION 18. — *Soit $M \subset N$ une coalition d'agents malveillants. Soit $n_{i,k,x}$ le nombre de fois que l'agent a_k a fourni le service s_x à l'agent a_i et $n^+_{i,k,x}$ le nombre de fois où l'agent a_k a fourni le service s_x avec une bonne qualité. Le* coût *de la manipulation est donné par :*

$$\textit{coût} = \frac{\sum_{a_k \in M} \sum_{s_x \in S} n^+_{i,k,x}}{\sum_{a_k \in N} \sum_{s_x \in S} n_{i,k,x}}$$

5.2. Protocole expérimental

Nous considérons dans nos expérimentation un système d'échange de services $\langle N, S\rangle$ où initialement $|N| = 100$ et $|S| = 10$. Parmis les 100 agents, 10 sont considérés comme appartenant à une même coalition malveillante et ils appliquent une attaque oscillante. L'expertise des agents pour un service est tirée aléatoirement uniformément entre 0 et 1, chaque agent pouvant fournir entre 0 et 5 services. La qualité des services est évalué sur l'intervalle $[-1, 1]$. Nous considérons le temps comme discret. À chaque pas de temps, les agents demandent un service qu'ils ne peuvent pas fournir eux-même. Afin de simplifier notre étude, nous supposons que chaque service est fourni en un pas de temps et qu'un agent peut fournir simultanément autant de ser-

vices que demandé. À chaque pas de temps, un nouvel agent peut rejoindre le système ou le quitter avec une probabilité de $0,01$.

Dans nos simulations, nous considérons que nos agents n'ont a priori aucune connaissance initiale et vont interagir durant 200 pas de temps. Nous réitérons ces simulations 50 fois et calculons la moyenne des métriques présentées précédemment. Dans ces simulations, nous considérerons trois politiques de sélection : UCB, $0,1$-gloutonne et $0,1$-élitiste que nous appliquons sur l'estimation collective, BetaReputation, EigenTrust et FlowTrust. Nous comparons ces résultats avec l'estimation personnelle utilisant UCB, ce qui correspond à un problème classique de bandits manchots. Intuitivement, l'estimation collective doit être très sensible aux manipulations alors que l'estimation personnelle (n'utilisant aucun témoignage) n'est sensible qu'aux changements de comportements.

5.3. Résultats expérimentaux

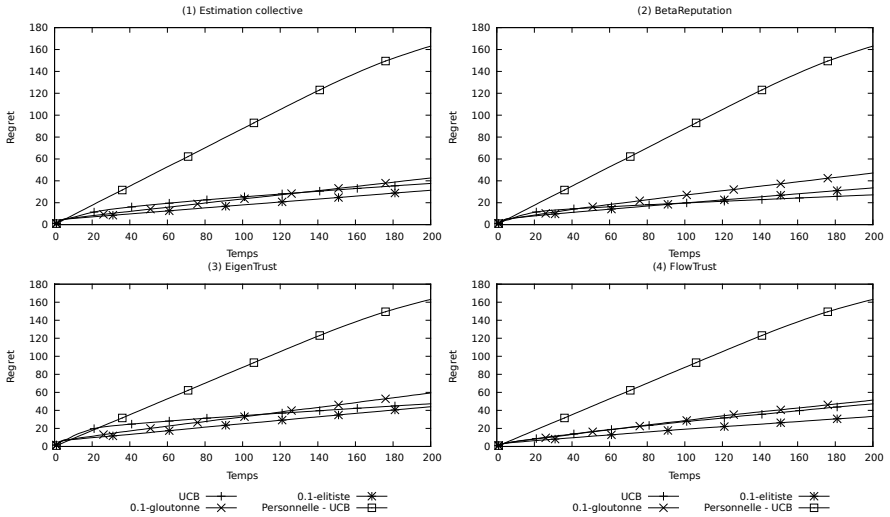

Figure 2. Regret en l'absence de manipulation selon (gauche à droite et haut en bas) (1) estimation collective ; (2) BetaReputation ; (3) EigenTrust ; (4) FlowTrust

La figure 2 nous montre l'intérêt des agents à coopérer en l'absence de manipulation dans le système. Nous considérons ici que les agents malveillant fournissent uniquement des mauvais services mais n'appliquent pas d'attaque oscillante. Indépendament de la politique de sélection utilisée, l'échange d'information permet aux agents d'avoir un regret très bas contrairement à l'estimation personnelle qui nécessite que les agents explorent chaque fournisseur, ce qui leur confère un regret important. La politique $0,1$-élitiste est celle qui minimise le regret sur l'estimation collective (figure 2.1) et FlowTrust (figure 2.4). UCB devient rapidement la politique qui minimise le regret sur BetaReputation (figure 2.2). Enfin, si UCB est initialement la politique

de sélection la moins performante sur EigenTrust (figure 2.3), la vitesse de croissance de son regret devient rapidement quasi-nulle, tendant ainsi à minimiser le regret. Dans les quatre cas, la politique $0,1$-gloutonne est celle qui fournit le plus haut regret. En effet, le facteur d'exploration de cette politique amène souvent les agents à interagir avec des agents pourtant identifiés comme ayant une faible expertise.

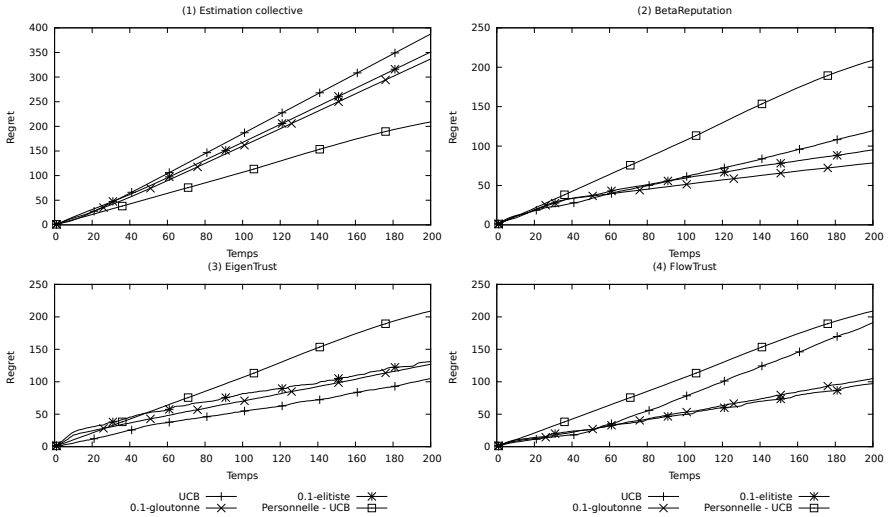

Figure 3. Regret en présence de manipulations selon (gauche à droite et haut en bas) (1) estimation collective ; (2) BetaReputation ; (3) EigenTrust ; (4) FlowTrust

La figure 3 présente le même système en présence de manipulations. Ainsi, le regret des agents est globalement plus important (y compris avec l'estimation personnelle puisque les agents malveillants affectuent une attaque oscillante). Ici, l'estimation collective (figure 3.1) est si peu robuste que toutes les politiques de sélection tendent à interagir avec les agents malveillants. Dans les systèmes de réputation non triviaux, si UCB était la politique la plus efficace précédemment, elle dégrade largement les performances de BetaReputation et FlowTrust (figures 3.2 et 3.4) en raison de son facteur d'exploration qui incite à interagir avec les agents qui viennent d'effectuer un blanchiment. À l'inverse, EigenTrust (figure 3.3) qui est le système de réputation le plus sensible aux manipulations est plus performant avec UCB : le facteur d'exploration va permettre aux agents honnêtes diffamés d'être tout de même sélectionnés. Remarquons que la politique $0,1$-gloutonne qui est la moins performante en l'absence de manipulation est ici celle qui minimise le regret sur BetaReputation.

Si manipuler les systèmes de réputation permet aux agents malveillants de fournir de mauvais services, ceux-ci fournissent également des services de bonne qualité afin de maintenir leurs valeurs de réputation. La figure 4 nous montre ce coût (définition 18). Nous pouvons constater qu'avec UCB, le coût initial est très élevé puis chute

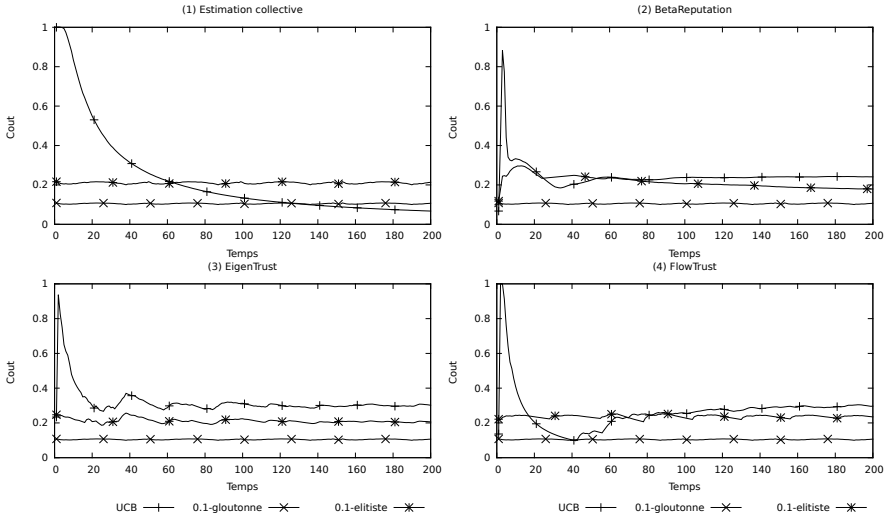

Figure 4. Coût de la manipulation selon (gauche à droite et haut en bas) (1) estimation collective ; (2) BetaReputation ; (3) EigenTrust ; (4) FlowTrust

rapidement. En effet, les agents honnêtes commencent par explorer et interagissent avec les agents malveillants qui ne sont pas promus et qui fournissent de bons services puis, très rapidement, le facteur d'exploration s'annule et les agents vont interagir avec les agents malveillants qui fournissent de mauvais services. BetaReputation, EigenTrust et FlowTrust (figures 4.2, 4.3 et 4.4) présentent des hausses occasionnelles du coût de la manipulation dues aux blanchiments. Avec les politiques de sélection $0,1$-gloutonne et $0,1$-élitiste, le coût de la manipulation est globalement constant pour toutes les politiques : comme les agents malveillants fournissant de bons services ne sont pas diffamés, ils appartiennent aux 10 % des meilleurs fournisseurs des services et peuvent donc être sélectionnés par le facteur d'exploration.

Nous pouvons donc conclure que les politiques de sélections ont une forte influence sur l'efficacité des fonctions de réputation. En l'absence de manipulation, UCB est la politique qui minimise le regret mais elle rend les fonctions de réputation beaucoup plus sensibles au blanchiment et au changement de comportement. Cependant, de telles manipulations ont un fort coût puisque les agents malveillants doivent fournir des bons services après blanchiment. À l'inverse, la politique $0,1$-gloutonne qui est la moins efficace en l'absence de manipulation devient la plus performante en présence de manipulations. En effet, son facteur d'exploration lui permet d'interagir avec des agents honnêtes même s'ils ont une faible réputation due aux diffamations. Malgré les manipulations, les trois politiques appliquées sur BetaReputation, EigenTrust et FlowTrust donnent un regret inférieur à celui de l'estimation personnelle. L'estimation collective, elle, reste toujours sensible aux faux témoignages. Cependant, il convient de remarquer que l'influence des politiques diffèrent selon la fonction de réputation utilisée. Ainsi, l'étude de l'influence des paramètres de ces fonctions (comme

l'ajout d'un facteur d'oubli pour BetaReputation ou une variation des confiances a priori d'EigenTrust) serait pertinente pour généraliser complètement nos résultats.

6. Mesure de credibilité et fonctions de filtrage

Si l'utilisation de politiques de bandits manchots influe sur la robustesse des systèmes de réputation, les faux témoignages restent un problème. Cependant, notre modèle nous permet de définir une notion de crédibilité pour filtrer les faux témoignages.

En effet, nous considérons les observations comme des variables aléatoires issues de fonctions de distribution de probabilité de paramètres inconnus, mais indépendantes de l'agent qui a reçu le service. Ainsi, avec suffisament d'observations, les estimations faites par chaque agent doivent converger. Nous proposons ici d'utiliser la divergence de Kullback-Leibler pour comparer ces estimations des fonctions de distribution de probabilité, et ce afin de définir si un témoignage est crédible ou non. Cette mesure de crédibilité permet aux agents de déterminer si un témoignage est faux et ainsi filtrer les témoignages provenant d'agents malveillants.

6.1. La divergence de Kullback-Leibler, une mesure de crédibilité

Dans notre modèle de bandit manchot, l'expertise de a_k pour le service s_x correspond à l'espérance de gain moyen lorsqu'un agent a_i lui demande ce service. À partir de ses observations directes $O_{i,k,x}$, l'agent a_i peut calculer le gain moyen qu'il a reçu lors de ses interactions avec l'agent a_k. Notons par $\mu_{i,k,x}$ ce gain moyen et par $\sigma_{i,k,x}$ l'écart-type. Bien que la qualité des services fournis par a_k ne suit pas nécessairement une loi normale, l'agent a_i peut l'approximer par la loi $\mathcal{N}(\mu_{i,k,x}, \sigma^2_{i,k,x})$. De même, graçe aux témoignages qu'il a reçus, a_i peut calculer $\mu_{i,j,k,x}$ et $\sigma_{i,j,k,x}$ le gain moyen et l'écart-type fondé sur les témoignages de a_j et ainsi obtenir l'approximation $\mathcal{N}(\mu_{i,j,k,x}, \sigma^2_{i,j,k,x})$. Ainsi, sous l'hypothèse que la qualité des services fournis est indépendante de l'agent recevant le service, deux agents doivent obtenir les mêmes estimations pour un grand nombre d'observations.

SUPPOSITION 19. — Si $O_{i,k,x}$ et $F_{i,j,k,x}$ sont des observations du service s_x fourni par a_k alors ces observations proviennent de la même fonction de distribution de probabilité. Ainsi, pour $n = |O_{i,k,x}|$ et $m = |F_{i,j,k,x}|$:

$$\lim_{n,m\to\infty} \mathcal{N}(\mu_{i,k,x}, \sigma^2_{i,k,x}) = \mathcal{N}(\mu_{i,j,k,x}, \sigma^2_{i,j,k,x})$$

Si $F_{i,j,k,x}$ est un faux témoignage alors :

$$\lim_{n,m\to\infty} \mathcal{N}(\mu_{i,k,x}, \sigma^2_{i,k,x}) \neq \mathcal{N}(\mu_{i,j,k,x}, \sigma^2_{i,j,k,x})$$

□

Notons que cette supposition n'a de sens que si les observations des agents sont sans erreurs. Dans le cas contraire, l'erreur d'observation peut être considérée comme

du bruit et si le bruit d'un témoignage est trop important, ce dernier fausse l'estimation de l'expertise. Par ailleurs, comme les agents ne disposent que d'un nombre fini d'observations, leurs estimations différent nécessairement. Ainsi, une mesure de crédibilité des témoignages nécessite de prendre en compte ces deux points.

Pour cela, nous proposons d'utiliser la divergence de Kullback-Leibler pour mesurer la différence entre deux témoignages. La divergence de Kullback-Leibler entre deux fonctions de probabilité $f(x)$ et $g(x)$ est :

$$D_{KL}(f||g) = \int f(x) \log \frac{f(x)}{g(x)} d(x)$$

DÉFINITION 20. — *La* divergence de Kullback-Leibler *entre les observations de a_i et les témoignages de a_j vis-à-vis de $\varepsilon_{k,x}$ est :*

$$D_{i,j,k,x} = D_{KL}(\mathcal{N}(\mu_{i,k,x}, \sigma^2_{i,k,x})||\mathcal{N}(\mu_{i,j,k,x}, \sigma^2_{i,j,k,x}))$$

Si les témoignages fournis par l'agent a_j sont similaires aux observations de l'agent a_i, $D_{i,j,k,x} \simeq 0$. Inversement, si $D_{i,j,k,x}$ est supérieur à un seuil δ, cela signifie que l'agent a_i et l'agent a_j n'ont pas la même estimation de $\varepsilon_{k,x}$. Cela peut être dû à plusieurs facteurs. Soit les agents n'ont pas suffisament d'observations pour avoir une bonne estimation de $\varepsilon_{k,x}$, soit ils évaluent la qualité des services sur des critères différents ($v_i \neq v_j$), soit les témoignages de a_j sont faux. Dans le premier cas, après quelques interactions supplémentaires, $D_{i,j,k,x}$ tendra vers 0. Dans les deux autres cas, cela signifie que l'agent a_i ne peut pas considérer comme crédibles les témoignages de l'agent a_j car ils sont soit faux, soit inutiles.

Pour fixer le seuil δ à partir duquel a_i considère comme non crédibles des témoignages, nous proposons d'utiliser l'erreur type de l'estimateur. Nous considérons ici l'erreur type de la moyenne (SEM $= \sigma_{i,k,x}/\sqrt{n}$) qui correspond à la confiance de a_i dans son estimation de $\mu_{i,k,x}$. L'approximation de la fonction de distribution de probabilité par une loi normale permet à l'agent a_i de déterminer avec une confiance de 95 % que la moyenne réelle des gains espérés se trouve dans l'intervalle :

$$[\mu_{i,k,x} - \frac{1.96 \times \sigma_{i,k,x}}{\sqrt{n}}, \mu_{i,k,x} + \frac{1.96 \times \sigma_{i,k,x}}{\sqrt{n}}]$$

Ainsi, l'agent a_i peut utiliser sa propre SEM pour fixer δ et calculer si les témoignages de a_j sont crédibles.

DÉFINITION 21. — *Soit $F_{i,j,k,x}$ le témoignage que a_j a fourni à a_i vis-à-vis de $\varepsilon_{k,x}$. $F_{i,j,k,x}$ est* KL-crédible *si $D_{i,j,k,x} \leq \delta$, où :*

$$\delta = D_{KL}(\mathcal{N}(\mu_{i,k,x}, \sigma^2_{i,k,x})||\mathcal{N}(\mu_{i,k,x} + \frac{1.96 \times \sigma_{i,k,x}}{\sqrt{n}}, \sigma^2_{i,k,x}))$$

Utiliser la divergence de Kullback-Leibler comme mesure de crédibilité et l'erreur type de la moyenne pour fixer dynamiquement le seuil présentent plusieurs avantages.

Comme la divergence de Kullback-Leibler est fortement liée à l'entropie, un témoignage divergent apporte de nouvelles informations utiles lorsque l'agent ne dispose que de peu d'observations. À l'inverse, plus l'agent dispose d'informations, moins un nouveau témoignage est supposé apporter une information utile. Comme l'erreur type de la moyenne dépend du nombre d'observations, plus l'agent en dispose, moins un témoignage divergent est supposé crédible, et inversement. Ainsi, cette notion de crédibilité est dynamique car elle peut être remise en cause au cours du temps au fur et à mesure que l'agent obtient de nouvelles informations.

L'assymétrie de la divergence de Kullback-Leibler nous permet de représenter le fait que si l'agent a_i considère comme non crédible le témoignage d'un agent a_j, l'agent a_j peut quant à lui considérer le témoignage de l'agent a_i comme crédible car lui-même ne dispose pas du même nombre d'observations. Enfin, la prise en compte de l'erreur type de la moyenne dans la définition du seuil de crédibilité permet à un agent de considérer que ses observations sont en partie imparfaites.

Dans toute la suite, nous notons par $KL_i(F_{i,j,k,x})$ (resp. $\neg KL_i(F_{i,j,k,x})$) si le témoignage $F_{i,j,k,x}$ est KL-crédible (resp. non KL-crédible) du point de vue de a_i.

Supposons que l'échelle d'évaluation de la qualité du service s_x soit définie sur $[-1, 1]$. Fixons les observations de a_i telles que $\mu_{i,k,x} = 0$ et $\sigma_{i,k,x} = 0,4$. La figure 5 représente $D_{i,j,k,x}$ en fonction de $\mu_{i,j,k,x}$ et de $\sigma_{i,j,k,x}$. Le blanc correspond ici à une configuration où $D_{i,j,k,x} > 1$. Notons que si $n = 5$, l'agent a_i a peu d'observations directes. Son erreur type est donc importante ($SEM \simeq 0,28$) et l'agent est susceptible d'accepter des témoignages qui divergent de ces observations ($\delta \simeq 0,41$). En revanche, si a_i a suffisament d'observations (par exemple $n = 50$, $SEM \simeq 0,08$ et $\delta \simeq 0,079$), il est susceptible de rejeter des témoignages en les considérant soit faux, soit n'apportant pas suffisament de nouvelles informations pour être utiles.

Bien qu'une connaissance de δ permette à un agent malveillant de construire un témoignage qui pourrait être accepté comme crédible, un tel faux témoignage sera alors peu divergent des observations de l'agent manipulé. Ainsi, il sera nécessaire d'avoir un grand nombre de faux témoignages pour affecter la décision de l'agent. Remarquons que ceci s'applique aussi aux témoignages d'agents honnêtes : ils doivent aussi être nombreux pour affecter la décision de l'agent.

6.2. Filtrer les témoignages non KL-crédibles

Afin de diminuer l'influence des faux témoignages dans le système de réputation, nous proposons d'introduire dans les fonctions de réputation un mécanisme de filtrage des témoignages jugés non crédibles.

DÉFINITION 22. — *La* fonction de filtrage *de l'agent a_i est la fonction $\phi_i(\mathcal{F}_i)$ qui retourne l'ensemble des témoignage que a_i considère comme crédible.*

Nous proposons ici d'utiliser dans une fonction de réputation uniquement les témoignages retournés par la fonction de filtrage.

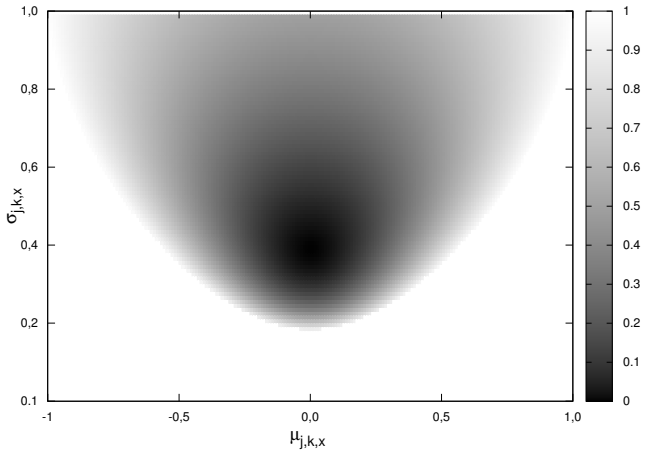

Figure 5. $D_{i,j,k,x}$ *en fonction de* $\mu_{i,j,k,x}$ *et de* $\sigma_{i,j,k,x}$

DÉFINITION 23. — *Soit un service* $s_x \in S$ *et deux agents* $a_i \in N$ *et* $a_k \in N_x$. *La* réputation crédible *de* a_k *pour le service* s_x *fondée sur la fonction de réputation* f_i *et la fonction de filtrage* ϕ_i *est définie par* $f_i(a_k, s_x, \phi_i(\mathcal{F}_i))$.

Comme la fonction de filtrage dépend de a_i, la réputation crédible est assymétrique au sens de (Cheng, Friedman, 2005) : elle peut donc être robuste aux attaques Sybil.

Bien qu'il existe de nombreuses fonctions de filtrage, nous proposons dans cet article trois fonctions de filtrage différentes : la première utilise trivialement la mesure de crédibilité que nous avons définie précédemment et les deux autres ont pour objectif de tenir compte de l'incertitude de l'agent, respectivement en généralisant ses observations ou en faisant appel aux observations d'autres agents.

Ainsi, une méthode intuitive pour filtrer les témoignages est de ne considérer que les témoignages KL-crédibles (définition 21) du point de vue de l'agent a_i.

DÉFINITION 24. — *Soit un service* $s_x \in S$ *et deux agents* $a_i \in N$ *et* $a_k \in N_x$. *La fonction de* KL-filtrage *est la fonction* ϕ_i *définie par :*

$$\phi_i(\mathcal{F}_i) = \{F_{i,j,k,x} \in \mathcal{F}_i | KL_i(F_{i,j,k,x})\}$$

Avec cette approche, un témoignage de a_j est considéré comme crédible ou non indépendamment de la crédibilité de ses autres témoignages. Or, si un agent n'a pas toujours les observations lui permettant de juger correctement un témoignage, il peut en avoir pour juger un autre témoignage provenant du même agent. Si nous faisons l'hypothèse qu'un agent mentant sur un témoignage a une forte probabilité de mentir sur un autre, nous pouvons considérer que si a_j n'est pas crédible sur un sous-ensemble de ses témoignages alors aucun de ses témoignages n'est crédible.

DÉFINITION 25. — *Soient deux agents* $a_i, a_j \in N$. *L'agent* a_j *est* k-crédible *si :*

$$\forall a_{k'} \in N, s_x \in S : |\{F_{i,j,k',x} \in \mathcal{F}_i | \neg KL_i(F_{i,j,k',x})\}| \leq k$$

Dans toute la suite, nous notons par $KL_i(N) \subseteq N$ l'ensemble des agents considérés comme k-crédibles par a_i. Nous pouvons ainsi définir une fonction de filtrage plus drastique qui rejette tous les témoignages provenant des agents non k-crédibles.

DÉFINITION 26. — *La* fonction de filtrage par k fautes *est la fonction* ϕ_i *telle que :*

$$\phi_i(\mathcal{F}_i) = \{F_{i,j,k',x} \in \mathcal{F}_i | a_j \in KL_i(N) \wedge KL_i(F_{i,j,k',x})\}$$

Remarquons que même si l'agent a_j est k-crédible, le sous-ensemble de ses témoignages qui ne sont pas KL-crédibles sont tout de même filtrés. Ainsi, le filtrage par k fautes est une généralisation de KL-filtrage. En effet, plus k est proche de 0, moins un agent accepte de témoignages crédibles car l'agent qui les fournit ne l'est pas. Inversement, plus k est grand, plus le filtrage par k fautes est proche du KL-filtrage.

Les deux fonctions de filtrage précédentes sont fondées sur les observations de l'agent a_i. La troisième fonction de filtrage que nous proposons permet d'utiliser les témoignages des autres agents pour déterminer si un témoignage est crédible, en s'inspirant d'une procédure de stochocratie. En politique, la stochocratie désigne un État dont le gouvernement est sélectionné aléatoirement. Les membres d'un tel gouvernement sont ainsi considérés comme moins sensibles à des manipulations (Delannoi, Dowlen, 2010). Dans notre contexte, nous proposons d'utiliser la stochocratie afin de juger si un témoignage est crédible : la fonction de filtrage par k-stochocratie accepte un témoignage si, parmi un sous-ensemble de k agents tirés aléatoirement uniformément dans N, une majorité d'entre eux le juge comme KL-crédible.

DÉFINITION 27. — *Le témoignage* $F_{i,j,k',x}$ *est dit* crédible par k-stochocratie *si, pour un sous-ensemble* $N' \subseteq N \setminus \{a_j, a_{k'}\}$ *de* k *agents tirés aléatoirement uniformément au moins* $\lceil k/2 \rceil$ *agents de* N' *jugent* $F_{j,k',x}$ *comme KL-crédible.*

Dans la suite, nous notons L_i l'ensemble des témoignages considérés comme crédibles par k-stochocratie par l'agent a_i. Nous pouvons ainsi définir la fonction de filtrage qui rejette tous les témoignages qui ne sont pas crédibles par k-stochocratie.

DÉFINITION 28. — *La* fonction de filtrage par k-stochocratie *est la fonction* ϕ_i *où :*

$$\phi_i(\mathcal{F}_i) = \{F_{i,j,k',x} \in \mathcal{F}_i | L_i(F_{i,j,k',x})\}$$

Notons que, dans la fonction de filtrage par k-stochocratie, les observations de l'agent sont elles aussi soumises au processus de filtrage. Ainsi, si a_i a un SEM important, ses observations peuvent ne pas être prises en compte lors du calcul de la réputation. Ceci permet de tenir compte d'une possible incertitude sur les observations.

Nous pouvons aussi nous interroger sur l'hypothèse implicite qui rend cette fonction efficace : seule une minorité d'agents juges peuvent être malveillants. En effet, comme certains des agents malveillants peuvent être sélectionnés parmi l'ensemble des juges, il est possible que ceux-ci amènent un faux témoignage à être considéré comme crédible. Cependant, le processus de k-stochocratie ne rend crédible un faux

témoignage que si et seulement si une majorité des juges le considère comme KL-crédible. Un tel cas n'arrive que si les agents honnêtes sélectionnés ont fourni peu de témoignages ou si la majorité des k agents sélectionnés appartiennent à la même coalition malveillante. La probabilité qu'au moins $\lceil k/2 \rceil$ de ces mauvais juges soient sélectionnés suit une loi hypergéométrique. Ainsi, un faux témoignage $F_{i,j,k',x}$ sera jugé par k-stochocratie comme crédible avec une probabilité p s'il existe l autres agents $a_z \in N \setminus \{a_j, a_{k'}\}$ tels que $KL_z(F_{i,j,k',x})$ et que :

$$\sum_{K=\lceil k/2 \rceil}^{k} \binom{l}{K} \binom{|N|-2-l}{k-K} \geq p \binom{|N|-2}{k}$$

Par exemple, considérons un système d'échange de services avec $|N| = 100$, $l = 20$ et un agent honnête utilisant la 10-stochocratie. La probabilité qu'un faux témoignage soit jugé comme crédible est de $0,0278$. Notons que plus k est petit, plus la probabilité qu'un faux témoignage soit jugé crédible est importante. En revanche, un k plus grand implique un temps de calcul plus important.

7. Impact de la crédibilité sur la robustesse des fonctions de réputation

7.1. Protocole expérimental

Pour évaluer l'efficacité des fonctions de filtrage, nous considérons deux protocoles expérimentaux. Le premier est le même que celui décrit en section 5.2 : 100 agents interagissant durant 200 pas de temps dont 10 agents malveillants effectuant une attaque oscillante. Dans le contexte de notre application, il s'agit d'une phase d'initialisation puisque aucun agent n'a de connaissance sur les autres. Notre second protocole correspond à une phase de fonctionnement nominal : 100 agents dont 10 malveillants ont déjà interagit durant 100 pas de temps lorsque 20 nouveaux agents honnêtes rejoignent le système. Ces nouveaux agents n'ont donc aucune connaissance a priori sur les autres agents et utilisent les témoignages qu'ils reçoivent pour calculer les valeurs de réputation. Nous mesurons alors le regret moyen de ces 20 agents durant les 200 pas de temps suivant.

Nous comparons trois fonctions de réputation (estimation collective, BetaReputation et FlowTrust) sans et avec nos trois fonctions de filtrage (KL-filtrage, filtrage par 10 fautes et 10-stochocratie [4]). Des expérimentations ont été menées avec EigenTrust. Cependant, comme il a été prouvé par (Cheng, Friedman, 2006) qu'il suffit d'un unique faux témoignage pour manipuler le système, le regret des agents est identique avec ou sans filtrage. Afin de mettre en difficulté notre approche, nous ne présentons ici que les résultats avec la politique de sélection UCB qui est la plus sensible aux attaques oscillantes. Enfin, nos résultats sont comparés à la fonction d'estimation personnelle. Nous considérons trois métriques : le regret (définition 17), le rappel et la

4. Le paramètre $k = 10$ a été fixé par des expérimentations non présentées dans cet article.

précision. Le rappel et la précision sont des mesures courantes dans le domaine de la classification (Bramer *et al.*, 2007). Ces mesures nous permettent de déterminer si l'utilisation de la divergence de Kullback-Leibler permet d'évaluer correctement la crédibilité des témoignages. Le rappel est le taux de faux témoignages filtrés. La précision est le ratio de faux témoignages filtrés parmi l'ensemble de tous les témoignages filtrés.

DÉFINITION 29. — *Soit un agent $a_i \in N$. Soit TP l'ensemble des faux témoignages filtrés et FN l'ensemble des faux témoignages considérés comme crédibles par a_i. Le* rappel *de la fonction de filtrage de l'agent a_i est :*

$$\textit{rappel}(\phi_i) = \frac{|TP|}{|TP| + |FN|}$$

DÉFINITION 30. — *Soit un agent $a_i \in N$. Soit TP l'ensemble des faux témoignages filtrés et TN l'ensemble des vrais témoignages considérés comme non crédibles par a_i. La* précision *de la fonction de filtrage de l'agent a_i est :*

$$\textit{precision}(\phi_i) = \frac{|TP|}{|TP| + |TN|}$$

7.2. Influence sur le regret

La figure 6 montre le regret des agents selon les fonctions de filtrage utilisées sur les différentes fonctions de réputation. Les gains obtenus sont donnés sur le tableau 3 où les valeurs correspondent à la réduction de regret apporté par les fonctions de filtrage. Comme nous l'avons vu précédement (figure 3.1), l'estimation collective est très manipulable. Cependant, un filtrage permet de réduire fortement l'influence des faux témoignages et ainsi diminuer le regret. Remarquons sur la figure 6.2 que la 10-stochocratie est beaucoup plus performante en fonctionnement nominal qu'en phase d'initialisation. En effet, cette fonction utilise les témoignages des autres agents pour détecter les faux témoignages. Or, en phase nominale, les nouveaux agents se reposent sur les observations précises des agents ayant déjà interagit. Dans les autres cas, les agents doivent obtenir plusieurs observations avant d'avoir une estimation correcte de l'expertise des fournisseurs.

Les figures 6.3 et 6.4 montrent que les fonctions de filtrage permettent de détecter les promotions et ainsi obtenir un regret plus faible sur FlowTrust. En effet, seules les promotions sont efficaces sur FlowTrust. Une fois celles-ci détectées, les agents ayant les meilleures réputations sont nécessairement ceux ayant une bonne expertise. Comme précédement, la 10-stochocratie a de meilleurs résultats en phase nominale. BetaReputation étant peu sensible aux manipulations, les fonctions de filtrage apporte naturellement un gain plus faible, voir même une augmentation du regret en phase nominale. Cependant ce résultat provient d'une moyenne sur l'ensemble de l'expérimentation. Dans les 40 premiers pas de temps, il y a effectivement une augmentation

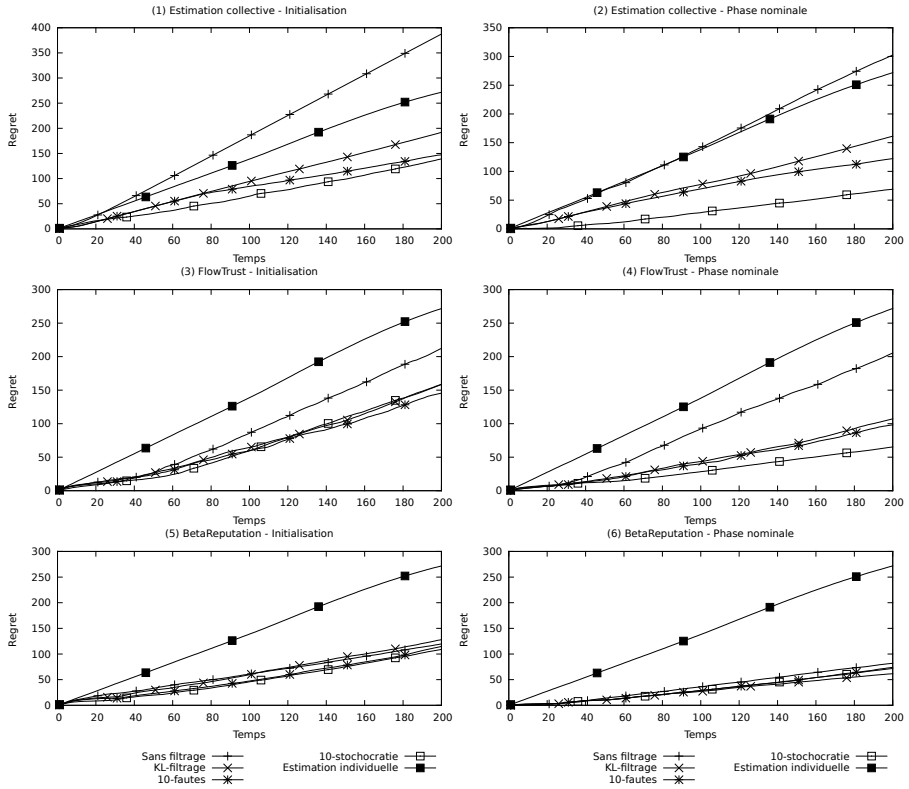

Figure 6. Regret : (1) estimation collective - initialisation ; (2) estimation collective - phase nominale ; (3) FlowTrust - initialisation ; (4) FlowTrust - phase nominale ; (5) BetaReputation - initialisation ; (6) BetaReputation - phase nominale

du regret. Après cela, les fonctions de filtrage obtiennent un regret inférieur. Il serait alors intéressant d'étudier si un agent malveillant ne pourrait pas profiter de cette sensibilité pour construire une manipulation efficace.

Tableau 3. Gains apportés par les fonctions de filtrage

Phase	Système de réputation	KL-filtrage	10-fautes	10-stochocratie
Initialisation	Estimation collective	$0,49$	$0,55$	$0,59$
	FlowTrust	$0,18$	$0,25$	$0,29$
	BetaReputation	$0,04$	$0,2$	$0,27$
Nominale	Estimation collective	$0,43$	$0,49$	$0,83$
	FlowTrust	$0,37$	$0,43$	$0,56$
	BetaReputation	$-0,05$	$-0,13$	$-0,11$
	(entre $t = 1$ et $t = 40$)	$-1,24$	$-1,44$	$-1,32$
	(entre $t = 41$ et $t = 200$)	$0,24$	$0,18$	$0,17$

7.3. Rappel et précision des fonctions de filtrage

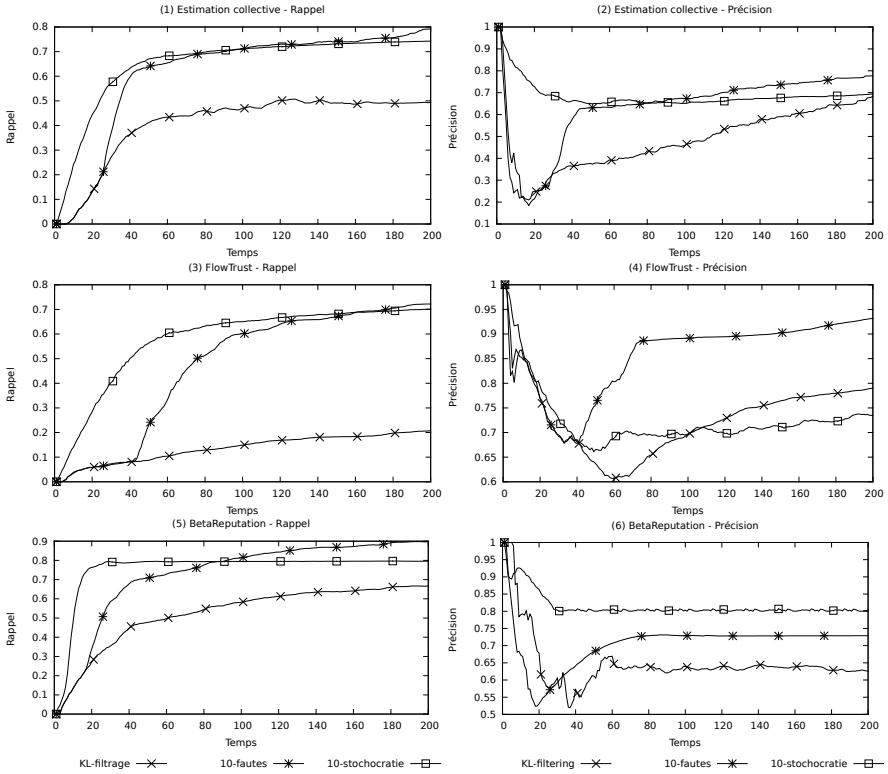

Figure 7. Rappel et précision des fonctions des filtrages sur (1 et 2) estimation collective ; (3 et 4) FlowTrust ; (5 et 6) BetaReputation

La figure 7 présente les rappels et précisions des différentes fonctions de filtrage sur les trois fonctions de réputation que sont l'estimation collective, FlowTrust et BetaReputation en phase d'initialisation. Le KL-filtrage permet de détecter 50 % des faux témoignages pour l'estimation collective (figure 7.1), 65 % pour BetaReputation (figure 7.5) et seulement 20 % sur FlowTrust (figure 7.3). Dans les deux premiers cas, les agents sont sensibles aux promotions et interagissent avec les agents malveillants dans les premiers pas de temps. Après quelques interactions, la divergence entre leurs observations et les faux témoignages est suffisante pour détecter les promotions. Les agents peuvent alors interagir avec les agents honnêtes même si ceux-ci sont diffamés, ce qui permet alors de détecter les diffamations [5]. Comme FlowTrust est insensible aux diffamations, le KL-filtrage ne détecte que les promotions et les

5. Notons que les diffamations qui ne sont pas détectées par le KL-filtrage correspondent en fait à des diffamation envers des agents honnêtes ayant une faible expertise.

diffamations vis-à-vis des meilleurs fournisseurs de services : les agents ne disposent alors pas de suffisament d'observations pour détecter les autres diffamations. Le filtrage par 10 fautes obtient un rappel identique à celui du KL-filtrage dans les premiers pas de temps. Après cela, la majorité des faux témoignages est soudainement détectée car, les agents malveillants n'étant plus crédibles sur certains témoignages, aucun de leurs témoignages ne l'est. La 10-stochocratie permet de détecter sur les trois fonctions de réputation environ 70 % des faux témoignages. En effet, les agents n'ont pas besoin d'avoir des observations pour déterminer si un témoignage est crédible. Les figures 7.2, 7.4 et 7.6 montrent que certains vrais témoignages sont considérés comme non crédibles par les fonctions de filtrage. Il s'avère que ces témoignages considérés comme non crédibles portent sur les agents ayant une mauvaise expertise. Comme les agents ont peu d'intérêt à interagir avec eux, ils n'ont que peu d'observations et une SEM importante. Ainsi, les témoignages sont rejetés car trop divergents des rares d'observations.

De manière générale, la précision décroit rapidement lors des premiers pas de temps avant de remonter progressivement. En effet, les agents débutent avec une SEM importante puis explorent petit à petit à l'aide d'UCB. La précision du filtrage par 10 fautes suit initialement la même décroissance avant de soudainement remonter lors de la détection massive de faux témoignages. Ces résultats nous permettent de confirmer que nos fonctions de filtrage sont efficaces pour détecter les faux témoignages sans faire un trop grand nombre d'erreurs. Le KL-filtrage et le filtrage par 10 fautes nécessitent quelques observations initiales pour être efficaces. À l'inverse, la 10-stochocratie utilise les témoignages des autres agents afin de décider de la crédibilité de chaque témoignage, réduisant ainsi le besoin en observations directes. Dans les trois cas, l'utilisation de la divergence de Kullback-Leibler entre les observations d'un agent et les témoignages reçus est une mesure efficace de crédibilité. Les agents peuvent donc l'utiliser pour filtrer les faux témoignages et ainsi augmenter la robustesse de leurs fonctions de réputation. Notons que l'efficacité des fonctions de filtrage dépendent de la fonction de réputation. Il est donc intéressant d'étudier quels paramètres du système influent sur l'efficacité de chaque fonction de filtrage, afin de généraliser ces résultats et déterminer quelle fonction de filtrage est la plus adaptée.

8. Conclusion

Dans cet article, nous étudions la robustesse d'un système de réputation modélisé comme un problème de bandits manchots. Ce modèle considère un agent comme un joueur et la réputation des autres agents comme l'espérance de gain du joueur. Les agents coopèrent en échangeant des témoignages de leurs observations. Cependant, des agents malveillants peuvent alors manipuler le système sous forme de faux témoignages, de blanchiments et d'attaques oscillantes.

Nous avons tout d'abord étudié l'utilisation des valeurs de réputation dans le processus de sélection d'un fournisseur de service. Ce problème de sélection dans les systèmes de réputation étant similaire à celui des bandits manchots, nous avons étudié

l'influence des politiques UCB, ε-gloutonne et ε-élitiste sur la robustesse de ces systèmes. Si en l'absence de manipulation UCB est la politique qui minimise le regret, elle rend le système de réputation beaucoup plus sensible à des manipulations telles que le blanchiment ou l'attaque oscillante (bien que ces manipulations aient un coût élevé). À l'opposé, une politique gloutonne moins efficace en l'absence de manipulation est plus robuste bien que diminuant le coût des manipulations. Nous avons ensuite étudié l'utilisation d'une mesure de crédibilité pour détecter les faux témoignages et ainsi accroître la robustesse des fonctions de réputation. Les gains des agents étant modélisés comme des variables aléatoires, nous proposons d'utiliser la divergence de Kullback-Leibler comme mesure de crédibilité des témoignages. Nous proposons trois fonctions de filtrage (KL-filtrage, filtrage par k-fautes et k-stochocratie) fondées sur cette mesure de crédibilité, permettant ainsi de n'utiliser que des témoignages crédibles lors du calcul de la réputation d'un agent. Nous avons montré empiriquement que ces trois fonctions de filtrage permettent de détecter efficacement les faux témoignages et ainsi fortement réduire le regret des agents.

Les deux approches, politiques de sélection et fonctions de filtrage, que nous proposons sont génériques et peuvent être utilisées sur de multiples fonctions de réputation. Cependant, il serait intéressant de relacher certaines de nos hypothèses et d'aller plus loin dans l'étude formelle de ces propositions. Par exemple, nous avons considéré ici que l'expertise des agents honnêtes est stationnaire. Il serait intéressant d'étudier si les politiques de sélection et de filtrage restent efficaces sur des systèmes non stationnaires. Dans la même optique, il est intéressant de considérer la présence d'agents aux observations imparfaites et d'étudier si les fonctions de filtrage restent aussi performantes en termes de précision et de rappel. D'un point de vue formel, il s'agirait de prouver (ou non) qu'une politique en domine une autre et sous quelles conditions. Enfin, puisque les fonctions de filtrage sont des heuristiques, il serait particulièrement intéressant de les généraliser au sein d'une fonction de filtrage unique. Enfin, une approche que nous souhaitons étudier est le paramétrage dynamique du tuple $\langle f_i, \Pi_i, \phi_i \rangle$ en fonction de l'évolution du regret des agents. Dans une telle approche, un agent disposerait de plusieurs politiques de sélection, fonctions de réputation et fonctions de filtrage et déterminerait à chaque pas de temps lesquelles utiliser en fonction de ses connaissances. Pour ce faire, la première étape serait d'étudier d'un point de vue formel sous quelles conditions l'un de ces tuples en domine un autre.

Bibliographie

Altman A., Tennenholtz M. (2005). Ranking systems: the PageRank axioms. In *6th ACM Conference on Electronic Commerce*, p. 1–8.

Altman A., Tennenholtz M. (2010). An axiomatic approach to personalized ranking systems. *Journal of the ACM*, vol. 57, n° 4, p. 201–236.

Anantharam V., Varaiya P., Walrand J. (1987). Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays. *IEEE Automatic Control*, vol. 32, n° 11, p. 968–976.

Auer P., Cesa-Bianchi N., Fischer P. (2002). Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, vol. 47, n° 2-3, p. 235–256.

Auer P., Cesa-Bianchi N., Freund Y., Schapire R. (1995). Gambling in a rigged casino: the adversarial multi-armed bandit problem. In *36th Annual Symposium on Foundations of Computer Science*, p. 322–331.

Auer P., Ortner R. (2010). UCB revisited: Improved regret bounds for the stochastic multi-armed bandit problem. *Periodica Mathematica Hungarica*, vol. 61, n° 1-2, p. 55–65.

Awerbuch B., Kleinberg R. (2005). Competitive collaborative learning. In *Learning theory*, p. 233–248.

Borisov N. (2006). Computational puzzles as Sybil defenses. In *6th IEEE International Conference on Peer-to-Peer Computing*, p. 171–176.

Bramer M., Bramer M., Bramer M. (2007). *Principles of data mining* (vol. 131). Springer.

Cheng A., Friedman E. (2005). Sybilproof reputation mechanisms. In *3rd Workshop on the Economics of Peer-to-Peer Systems*, p. 128–132.

Cheng A., Friedman E. (2006). Manipulability of PageRank under Sybil strategies. In *1st Workshop on the Economics of Networks, Systems and Computation.*

Delannoi G., Dowlen O. (2010). *Sortition, theory and practice*. Academic UK and USA.

Dini F., Spagnolo G. (2009). Buying reputation on eBay: Do recent changes help? *International Journal of Electronic Business*, vol. 7, n° 6, p. 581–598.

Douceur J. (2002). The Sybil attack. In *1st International Workshop on Peer-to-Peer Systems*, p. 251–260.

Feldman M., Lai K., Stoica I., Chuang J. (2004). Robust incentive techniques for peer-to-peer networks. In *5th ACM Conference on Electronic Commerce*, p. 102–111.

Hoffman K., Zage D., Nita-Rotaru C. (2009). A survey of attack and defense techniques for reputation systems. *ACM Computing Surveys*, vol. 42, n° 1.

Josang A., Ismail R. (2002). The Beta reputation system. In *15th Bled Electronic Commerce Conference*, p. 41–55.

Kamvar S., Schlosser M., Garcia-Molina H. (2003). The EigenTrust algorithm for reputation management in P2P networks. In *12th International World Wide Web Conference*, p. 640–651.

Koulouriotis D., Xanthopoulos A. (2008). Reinforcement learning and evolutionary algorithms for non-stationary multi-armed bandit problems. *Applied Mathematics and Computation*, vol. 196, n° 2, p. 913–922.

Koutrouli E., Tsalgatidou A. (2011). Credibility enhanced reputation mechanism for distributed e-communities. In *19th Euromicro International Conference on Parallel, Distributed and Network-Based Processing*, p. 627–634.

Kullback S. (1997). *Information theory and statistics*. Courier Dover Publications.

Liu K., Zhao Q. (2010). Distributed learning in multi-armed bandit with multiple players. *IEEE Signal Processing*, vol. 58, n° 11, p. 5667–5681.

Malik Z., Bouguettaya A. (2009). Rateweb: Reputation assessment for trust establishment among webservices. *International Journal on Very Large Data Bases*, vol. 18, n° 4, p. 885–911.

Marsh S. P. (1994). *Formalising trust as a computational concept*. Thèse de doctorat non publiée, University of Stirling.

Noor T., Sheng Q., Ngu A., Alfazi A., Law J. (2013). Cloud armor: a platform for credibility-based trust management of cloud services. In *22nd ACM International Conference on Information and Knowledge Management*, p. 2509–2512.

Resnick P., Kuwabara K., Zeckhauser R., Friedman E. (2000). Reputation systems. *ACM Communications*, vol. 43, n° 12, p. 45–48.

Ripeanu M. (2001). Peer-to-peer architecture case study: Gnutella network. In *1st International Conference on Peer-to-Peer Computing*, p. 99–100.

Robbins H. (1952). Some aspects of the sequential design of experiments. *Journal of the AMS*, vol. 58, n° 5, p. 527–535.

Sabater J., Paolucci M., Conte R. (2006). Repage: Reputation and image among limited autonomous partners. *Journal of Artificial Societies and Social Simulation*, vol. 9, n° 2, p. 1–18.

Sabater J., Sierra C. (2001). Social ReGreT, a reputation model based on social relations. *ACM SIGecom Exchanges*, vol. 3, n° 1, p. 44–56.

Selcuk A. A., Uzun E., Pariente M. R. (2004). A reputation-based trust management system for P2P networks. In *14th IEEE International Cluster Symposium on Computing and the Grid*, p. 251–258.

Srivatsa M., Xiong L., Liu L. (2005). TrustGuard: countering vulnerabilities in reputation management for decentralized overlay networks. In *14th International World Wide Web Conference*, p. 422–431.

Vermorel J., Mohri M. (2005). Multi-armed bandit algorithms and empirical evaluation. In *16th European Conference on Machine Learning*, p. 437–448.

Zhao H., Li X. (2008). H-trust: A robust and lightweight group reputation system for peer-to-peerdesktop grid. In *28th International Conference on Distributed Computing Systems*, p. 235–240.