

# Les systèmes informatiques fondés sur la confiance : un état de l'art.

Emmanuelle Gaillard

► **To cite this version:**

Emmanuelle Gaillard. Les systèmes informatiques fondés sur la confiance : un état de l'art.. [Rapport de recherche] Loria

Inria Grand Est. 2011, pp.23. <hal-00662479v2>

**HAL Id: hal-00662479**

**<https://hal.inria.fr/hal-00662479v2>**

Submitted on 4 Oct 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Les systèmes informatiques fondés sur la confiance : un état de l'art.

Emmanuelle GAILLARD

30 novembre 2011

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Définitions en sciences humaines . . . . .	2
1.1.1	Deutsch, le psychologue . . . . .	2
1.1.2	Luhman, le sociologue . . . . .	2
1.1.3	Gambetta, sciences sociales et politiques . . . . .	3
1.2	Quelques notions sur la confiance . . . . .	3
1.2.1	Réciprocité . . . . .	4
1.2.2	Mémoire . . . . .	4
1.2.3	Coopération . . . . .	4
1.2.4	Etre quelqu'un de confiant . . . . .	5
<b>2</b>	<b>Systèmes à base de confiance</b>	<b>6</b>
2.1	Systèmes de confiance à base de politique . . . . .	6
2.2	Système de confiance à base de réputation . . . . .	6
2.2.1	Modélisation de la confiance . . . . .	7
2.3	Système centralisé, système décentralisés . . . . .	7
2.3.1	Systèmes centralisés . . . . .	7
2.3.2	Systèmes distribués . . . . .	7
<b>3</b>	<b>Modèles computationnels</b>	<b>9</b>
3.1	Trust transitivity . . . . .	9
3.2	Réseaux bayésiens . . . . .	9
3.2.1	Modèle de Jøsang : bêta réputation . . . . .	10
3.2.2	D'autres systèmes bayésiens . . . . .	11
3.3	Valeurs discrètes . . . . .	11
3.3.1	Abdul-Rahman . . . . .	11
3.3.2	TidalTrust . . . . .	12
3.4	Modèles basés sur la croyance . . . . .	14
3.4.1	Jøsang : belief model . . . . .	14
3.4.2	Modèle de Barber . . . . .	14
3.5	Modèles flous . . . . .	14
3.5.1	Système REGRET . . . . .	14
3.6	Modèles de flux . . . . .	15
3.6.1	Google PageRank . . . . .	15
3.6.2	EigenTrust . . . . .	15
3.7	Un modèle plus général : Marsh . . . . .	16
<b>4</b>	<b>Les systèmes commerciaux</b>	<b>18</b>
4.1	Ebay . . . . .	18
4.2	Amazon . . . . .	18
4.3	Epinions . . . . .	18
<b>5</b>	<b>Conclusion</b>	<b>20</b>

# Chapitre 1

## Introduction

Bien que Luhman ait présenté la confiance comme un fait basique de la vie humaine [Luhmann, 1979], la confiance est un terme complexe qui ne connaît pas de définition consensuelle. Cette définition dépend du domaine et du courant de pensée des auteurs. Certains auteurs place le risque au coeur de la confiance [Deutsch, 1962] tandis que pour d'autres la situation est très importante [Barber, 1983, Marsh, 1994]. D'autres définissent la confiance comme la prévision de comportements, en effet, tout comme [Gambetta, 2000], Rempel voit la confiance comme une probabilité subjective sur de futurs événements, tandis que Stztomka voit la confiance comme un pari sur les futures actions des autres [Rempel and Holmes, 1986].

La littérature est donc très riche sur le concept de la confiance, cependant, certains auteurs se sont distingués dans leurs recherches en publiant des travaux signifiant sur la confiance dans leur domaine. [Deutsch, 1962], se base sur une perception individuelle de la confiance en parallèle à [Luhmann, 1979] qui place ce concept dans une dimension sociale tandis que [Gambetta, 2000] voit la confiance comme la prévision d'un comportement. Ces travaux introduisent la discussion sur la confiance, pour ensuite éclaircir les propriétés générales communes à de nombreuses définitions. Ces propriétés générales sont utilisées dans les systèmes à bases de confiance afin de mesurer le degré de confiance entre entités.

### 1.1 Définitions en sciences humaines

#### 1.1.1 Deutsch, le psychologue

[Deutsch, 1962] définit la confiance comme une décision par rapport à une perception individuelle des coûts et bénéfices dont dépend cette décision. Lors d'une décision de confiance, l'individu est confronté à un chemin ambigu dont les issues dépendantes d'une tierce personne peuvent être perçues positives ou négatives. L'individu perçoit les issues négatives plus importantes que les issues positives. En choisissant de faire confiance l'individu suppose que l'issue positive se produira plutôt que l'issue négative. L'individu est donc confiant aux capacités et intentions de la tierce personne dont dépend l'occurrence de l'issue positive. Par exemple, David vient de se casser le bras, il est incapable de conduire et il doit aller à son travail. Il fait le choix de faire confiance à Sylvie pour l'amener sur son lieu de travail. Cet exemple est bien selon Deutsch une situation de confiance, en effet, David se retrouve face à un choix dont l'issue dépend d'une tierce personne Sylvie, et l'issue positive (arriver sur son lieu de travail) est moins importante que l'issue négative (Sylvie peut avoir un accident). La définition de Deutsch rejoint celle de [Boon and Holmes, 1991] affirmant que celui qui fait confiance espère un aboutissement positif, tout en gardant à l'esprit le risque encouru d'une issue négative possible.

#### 1.1.2 Luhman, le sociologue

Tandis que [Deutsch, 1962] place le terme de confiance dans une dimension individuelle [Luhmann, 1979], tout comme [Barber, 1983], inscrit la confiance dans une réalité sociale multidimensionnelle. Selon [Luhmann, 1979], la confiance permettrait de faire diminuer la complexité de l'environnement, c'est pourquoi il la voit comme un « fait basique de la vie humaine ». Certaines situations peuvent connaître des possibilités de dénouements très élevées. Par exemple, lorsque je sors de chez moi le matin, je peux me faire enlever, je peux me faire tuer, me faire agresser, tomber malade, etc. Cependant, si je décide de sortir de chez moi et de faire confiance à ce que rien de tout cela n'arrive, je diminue la complexité de ce système. Luhman voit donc

la confiance comme un fait basique [Luhmann, 1979], car un individu effectue les choix de faire confiance quotidiennement afin de s'adapter dans son environnement.

Tout comme d'autres auteurs [Marsh, 1994, Deutsch, 1962], Luhman affirme que la confiance se base sur une notion de risque [Luhmann, 1979], en effet, c'est l'acceptation du risque de me faire tuer, qui permet de diminuer la complexité sociale. Selon lui, c'est seulement de par la présence de la possibilité d'un aboutissement négatif que la confiance est présente dans une situation. Luhman oppose l'acte de faire confiance (*trust*) et le sentiment de faire confiance (*confidence*), dans une situation de *confidence* l'alternative n'est pas possible donc il y a absence de risque.

Luhman décrit la confiance comme un phénomène sociale où des individus interagissent dans la société [Luhmann, 1979]. Ce phénomène sociale se traduit par le fait que celui qui fait confiance doit faire face à la liberté de celui à qui il fait confiance. En effet, celui à qui la confiance a été donnée choisit d'effectuer correctement ou non ce qu'on lui demande. Cependant, lorsque quelqu'un choisit de faire confiance en une autre personne, il possède certaines attentes envers cette personne malgré les libertés qu'elle peut avoir. Selon Barber qui se place aussi dans une dimension sociale de la croyance, l'individu qui fait confiance espère que l'autre remplira des compétences techniques ainsi que des responsabilités et des obligations morales [Barber, 1983].

### 1.1.3 Gambetta, sciences sociales et politiques

Afin de définir le terme de confiance, Gambetta introduit la notion de probabilité subjective et de circonstances.

"La confiance (ou symétriquement la défiance) est un niveau particulier de la probabilité subjective avec laquelle un agent accomplira une action spécifique, à la fois avant que nous ne puissions suivre chaque action (ou indépendamment de sa capacité de même pouvoir la tracer) et aussi dans un contexte dans lequel cela affecte notre propre action." [Gambetta, 2000]

Dans cette définition, Gambetta fait seulement référence à la confiance entre deux agents et non par exemple à la confiance entre un agent et son environnement. De plus, les actions des agents n'ont pas de conséquence sur la confiance car elle est évaluée à l'instant présent sans avoir de retour sur les actions des agents. Pour Gambetta, ce niveau particulier, en plus d'être dépendant de l'agent qui l'évalue est aussi dépendant des circonstances. En effet, dans deux circonstances différentes un agent pourra évaluer différemment la confiance qu'il a envers un autre agent. La confiance n'est donc pas généralisable, elle est spécifique du contexte.

[McKnight and Chervany, 1996] reprennent les mêmes paramètres que Gambetta dans leur définition en ajoutant le fait que la confiance est présente malgré un aboutissement négatif possible.

Dans le domaine informatique, [Golbeck, 2005] se rapproche de la définition de Gambetta où la confiance est un engagement à croire au bon déroulement des futures actions d'une autre entité. Selon [Grandison, 2003], l'acte de faire confiance se réalise dans un contexte spécifique et se définit par une croyance quantifiée quant aux habilités de l'entité qui est crue. Cette quantification peut être une échelle de valeurs ou une simple classification.

Ces auteurs ont tous des définitions différentes sur la confiance cependant trois éléments sont présents dans chaque définition :

- Une personne ou entité faisant confiance (*trustor*)
- Une cible, une personne ou une entité à qui on fait confiance (*trustee*)
- Une situation

De plus, ces auteurs et d'autres étudiés dans des états de l'art sur la confiance [Jøsang et al., 2007], [Sabater and Sierra, 2002, Artz and Gil, 2007, Viljanen, 2005], ont permis de faire ressortir certaines généralités sur la confiance, présentées dans la section suivante.

## 1.2 Quelques notions sur la confiance

### Risque, coûts et bénéfices

Comme nous l'avons vu plus haut, Luhman introduit la notion de risque dans sa définition de la confiance [Luhmann, 1979]. Dans la relation entre deux personnes, cette notion de risque vient renforcer le lien de

confiance. Plusieurs auteurs ont proposés des solutions pour mesurer le risque encouru dans une situation de confiance.

- Marsh a représenté le risque en un rapport entre les coûts et les bénéfices [Marsh, 1994] :

$$risque = \frac{coûts}{bénéfices}$$

- [Boyle and Bonacich, 1970] ont utilisés la théorie des jeux avec le dilemme du prisonnier pour une estimation du risque. Cependant leur proposition n'est possible que lorsque les gains et les pertes sont connus, elle est difficilement applicable lorsque la situation est incertaine.

## Les caractéristiques de Trust

- La confiance n'est pas symétrique : « Alice a confiance en Bob » ne veut pas dire que « Bob a confiance en Alice ».
- La confiance n'est pas distributive : « Alice a confiance en (Bob et David) » ne veut pas dire que « Alice a confiance en Bob et Alice a confiance en David ».
- La confiance n'est pas associative : « (Alice a confiance en Bob) a confiance en Alice » n'est pas possible tandis que « (Alice a confiance en (Bob qui a confiance en Alice)) » est possible.
- La confiance n'est pas généralisable : « Alice a confiance en Bob dans le domaine de la mécanique » ne signifie pas que « Alice a confiance en Bob dans le domaine culinaire ».
- La confiance n'est pas intrinsèquement transitive : si « Alice a confiance en Bob » et « Bob a confiance en David » alors « Alice a confiance en David » n'est pas forcément vrai. Une discussion sur la transitivité est détaillée plus bas en 3.1.

### 1.2.1 Réciprocité

[Mui et al., 2002] définissent la réciprocité en « Soit gentil avec ceux qui sont gentils avec toi ». Dans [Boyd and Richerson, 2009], la réciprocité permet de faire diminuer le caractère aléatoire des interactions sociales. C'est cette idée que Marsh utilise dans sa représentation de la confiance [Marsh, 1994], en effet, si un agent a fait une faveur à autre agent il espérera une autre faveur en retour. Donc si un agent A a fait une faveur à un agent B, et que cette agent B ne lui a pas encore retourné la pareil, l'agent A peut augmenter la valeur de la confiance qu'il a en B. D'autres cas sont discutés dans [Marsh, 1994].

### 1.2.2 Mémoire

La mémoire est un facteur important dans le domaine de la confiance, différentes question peuvent se poser : la confiance doit-elle être accordée en prenant en compte toutes les précédentes interactions ? Ou doit-elle seulement prendre en compte les  $n$  dernières interactions ? Certains modèles ont mis en place un seuil qui prend en compte seulement les  $n$  dernières interactions tandis que d'autres, comme [Jøsang and Ismail, 2002], ont introduit un facteur d'oubli.

### 1.2.3 Coopération

Le concept de coopération est lié au concept de confiance, pour coopérer avec quelqu'un j'ai besoin de lui faire confiance, de savoir qu'il ne va pas me trahir ou qu'il ait les compétences nécessaires. Selon [Williams, 1988], la confiance est une notion indispensable pour assurer une coopération réussie. Marsh a établi un modèle pour calculer un seuil de coopération [Marsh, 1994] ; pour coopérer avec un agent, la confiance envers cet agent dans une situation particulière doit atteindre ce seuil de coopération. Le système REGRET [Sabater and Sierra, 2002] base son évaluation du degré de confiance sur le fait que deux agents peuvent coopérer, être en compétition ou être en transaction. [Muy et al.] dans leur système, avancent que les concepts de coopération, réciprocité et réputation sont liés entre eux.

#### **1.2.4 Etre quelqu'un de confiant**

Selon Deutsch, les personnes fiables sont plus aptes à faire confiance que les personnes les moins fiables [Deutsch, 1962]. Ceci s'explique par le fait que si vous êtes quelqu'un de fiable vous aurez tendance à penser naïvement que toutes les personnes sont pareils, par contre, si vous n'êtes pas fiables, vous savez de quoi vous êtes capables donc de quoi sont capables les autres. Cette idée se rapproche du concept d'identité projective en psychologie. Marsh a utilisé cette notion en avançant le fait qu'on peut estimer la valeur de confiance d'un agent en connaissant à quelle valeur de confiance cet agent nous estime [Marsh, 1994].

## Chapitre 2

# Systemes à base de confiance

Dans la littérature, il existe principalement deux types de systèmes de confiance :

- Systèmes de confiance à base de politique
- Systèmes de confiance à base de réputation

Le premier est fondé sur le principe de compétences et d'autorisations requises pour effectuer une action tandis que le deuxième que nous allons développer plus en détail dans cette revue, utilise la réputation d'une entité permettant d'accorder ou non la confiance envers cette entité. Dans cette section, nous présenterons les systèmes centralisés et les systèmes décentralisés qui sont utilisées pour les systèmes de confiance à base de réputation.

### 2.1 Systèmes de confiance à base de politique

Ces systèmes utilisent des politiques pour définir si une entité à l'autorisation ou non à un accès. Ces systèmes sont basés sur la définition de gestion de la confiance (*trust management*).

Selon [Blaze et al., 2003], « la gestion de la confiance est une approche unifiée pour spécifier et interpréter des politiques, des qualifications et des relations permettant d'autoriser ou non des actions. » Les systèmes se basant sur cette définition de gestion de la confiance gèrent les accès et les actions dans le système grâce à des autorisations. Ces autorisations peuvent se faire sur les qualifications. Dans les systèmes de gestion de la confiance, la confiance a une valeur binaire autorisant ou non une entité d'effectuer une action. Les systèmes KeyNote [Blaze et al., 1999] et PolicyMaker [Blaze et al., 1996] reposent sur la gestion de la confiance.

### 2.2 Système de confiance à base de réputation

Les systèmes de confiance à base de réputation utilisent les interactions, ou expériences directes, entre entités ou /et l'expérience des autres pour le choix de faire confiance à une autre entité. La réputation est vue comme un facteur guidant (il peut être unique ou associé à d'autres) le processus de décision de confiance. Le dictionnaire anglais d'Oxford définit la réputation comme « l'estimation commune ou générale d'une personne avec le respect de la personnalité ou d'autres qualités ».

La confiance et la réputation sont des termes très proches et extrêmement liés. [Mui et al., 2002], voient la réputation comme la perception des intentions et des normes d'un autre agent, tandis qu'ils définissent la confiance comme un prévision subjective d'un agent sur le comportement d'un autre agent. La réputation est la perception d'une entité en fonction des expériences passée mais ne suppose rien sur les comportements futurs de cette entité tandis que la confiance fait une prévision sur ces comportements futurs en prenant en compte par exemple les expériences passées. La réputation sert à évaluer la confiance, les définitions des deux termes sont très proches car dans de nombreux systèmes, seule la réputation est le support de l'évaluation de la valeur de la confiance. La confiance liée ici à la réputation concerne la catégorie « trust Personnel » de [McKnight and Chervany, 1996] où une personne croit une autre personne dans une situation spécifique.

Dans les systèmes de confiance à base de réputation, il existe des processus pour collecter, réunir ou répartir la réputation dans le système. Selon [Resnick et al., 2000], ces systèmes possèdent trois propriétés : i. après chaque interaction, il y a espérance d'une prochaine interaction, ii. les interactions individuelles sont capturée et distribuées (à une entité centrale ou aux autres entités) et iii. les interactions sont le guide des décisions sur



les futures interactions entre entités. Ces systèmes reposent aussi sur des mécanismes computationnels afin de calculer cette réputation.

### 2.2.1 Modélisation de la confiance

Cette approche correspond à la façon de représenter et d'évaluer la confiance. De nombreuses propositions afin de représenter la confiance ont été faites dans la littérature. Selon, [Abdul-Rahman and Hailes, 1996], « les valeurs de confiance sont utilisées pour représenter les différents niveaux de confiance qu'un agent peut avoir envers un autre agent ». Cette définition est la base de nombreux systèmes, mais les différents niveaux peuvent être représentés de différentes façons. Ils peuvent correspondre à une valeur unique qui peut être une valeur continue dans un ensemble donné [Marsh, 1994, Jøsang and Pope, 2005], une valeur discrète [Golbeck, 2005], des valeurs floues [Sabater and Sierra, 2002] ou encore des étiquettes discrètes [Abdul-Rahman and Hailes, 2000]. Selon Marsh [Marsh, 1994], cette valeur n'a pas d'unité, il existe seulement un seuil déterminant si l'agent peut être digne de confiance ou non. [Abdul-Rahman and Hailes, 1996] proposent aussi de ne pas représenter la confiance comme une valeur absolue mais comme une relation entre agents.

## 2.3 Système centralisé, système décentralisés

Les systèmes de confiance, notamment ceux basés sur la réputation, peuvent fonctionner de deux manières, la première avec une unité centralisée qui gère les informations permettant d'estimer la confiance entre entités, ou la seconde de manière distribuée où chaque entité stocke ses propres informations à propos des autres entités.

Pour décrire ces deux types de systèmes, nous nous plaçons dans les systèmes de confiance à base de réputation, afin d'homogénéiser le vocabulaire.

### 2.3.1 Systèmes centralisés

Dans un système centralisé, la réputation de chaque agent est disponible à tout moment grâce à la collecte au fur et à mesure des scores donnés par tous les agents qui ont interagi avec l'agent donné. A chaque fois qu'un agent  $A$  décide d'interagir avec un agent  $B$ , l'agent  $A$  évalue cette interaction et envoie le score de l'évaluation à une unité centrale. Cette dernière stocke alors le score et l'utilise pour mettre à jour l'ancien score de réputation de  $B$  dans le système. Ainsi, dans un système centralisé, à tout moment un agent peut interroger l'unité centrale afin de connaître le score de réputation d'un agent en particulier.

Un système centralisé doit donc être doté d'un protocole de communication centralisé afin que l'autorité centrale puisse collecter les scores, et d'un moteur de calcul de réputation pour calculer un score global pour un agent donné.

### 2.3.2 Systèmes distribués

Un système distribué est un système sans aucune composante centrale. Il n'y a donc aucun lieu de stockage pour enregistrer les scores de réputation des agents. Chaque agent enregistre lui-même le résultat de chaque interaction directe qu'il a eu avec un agent donné. Ce score peut alors être fourni à un autre agent grâce à une requête. Ceci implique que si un agent  $A$  veut interagir avec un agent  $B$ , pour connaître la fiabilité de  $B$ , il doit trouver les agents qui ont interagis directement avec  $B$ .  $A$  calcule alors le score de réputation de  $B$  grâce à tous les scores collectés.

Un système distribué doit être doté d'un protocole de communication entre agents et d'une méthode de calcul du score de réputation basée sur les scores des autres agents et éventuellement d'autres facteurs.

Les systèmes distribués pose le problème d'inférer la confiance soit par la transitivité, concept discuté en 3.1, soit, par exemple par voisins les plus proches, car il est trop coûteux de récupérer toutes les interactions directes qu'a pu avoir l'agent cible. Les systèmes pair-à-pair (*Peer-to-peer* (P2P)) présentés ci-dessous sont des systèmes de cette catégorie.

Les systèmes P2P sont des systèmes qui peuvent être vus comme des systèmes distribués où chaque pair est un client ainsi qu'un serveur. Dans ces systèmes, il existe deux phases ; la première consiste à choisir le serveur qui fournit le service recherché par le client tandis que la deuxième phase correspond à la phase de

téléchargement de la ressource voulue. Cependant, dans ce genre de système il existe des serveurs non fiables qui ne fournissent pas exactement ce à quoi le client s'attend et détruisent ainsi l'utilité de tout le système. C'est pourquoi, un système de confiance basé sur la réputation est devenu souvent indispensable. Ainsi, un pair peut choisir un autre pair en fonction de sa fiabilité. De nombreux systèmes de confiance ont été mis en place pour le P2P. Selon EigenTrust [Kamvar et al., 2003], les systèmes à base de réputation dans les P2P sont utiles pour déterminer les serveurs offrant des ressources non fiables et les serveurs offrant les ressources de meilleures qualités.

## Chapitre 3

# Modèles computationnels

Les systèmes reposant sur la confiance doivent, comme nous l'avons remarqué ci-dessous, posséder une méthode de calcul du degré de confiance. En effet, le calcul de la confiance doit permettre de prendre une décision, par exemple un degré élevé de confiance envers une entité permet de juger cette entité fiable et de prendre la décision de lui faire confiance. De nombreux auteurs ont proposé des modèles pour représenter et calculer la confiance dans les systèmes. Ces modèles sont classés dans des catégories : modèles bayésiens, modèles de valeurs discrètes, modèles basés sur la croyance, modèles flous et modèles de flux. Après avoir défini un concept clé de la confiance, la transitivité présente dans de nombreux modèles, nous présentons dans ce papier les modèles les plus connus et représentatifs de chaque catégorie

### 3.1 Trust transitivity

[Marsh, 1994] et [Gerck, 1997], nous prouvent que la transitivité n'est pas une propriété de la confiance. Or, beaucoup de systèmes basés sur la confiance reposent sur ce concept de transitivité. Dans ces systèmes, lorsqu'une entité ne possède pas de liens directs avec une entité cible (avec qui elle n'a pas eu d'expérience directe), elle calcule son degré de confiance en combinant la confiance qu'elle a envers des entités avec qui elle a eu des expériences directes et la confiance que ces entités ont avec l'entité cible. Pour dévier ce problème, les auteurs ont définis deux types de confiance [Jøsang and Pope, 2005] représentés par la figure 3.1 :

- La confiance de référence : fait référence à une tierce personne. Le calcul du degré de confiance d'une entité  $A$  envers une entité  $B$  dépend de la recommandation d'une autre entité  $C$ . Selon [Abdul-Rahman and Hailes, 1996], la recommandation est une information sur la réputation d'une entité, échangé entre deux agents.
- La confiance fonctionnelle : correspond à la fin du chemin de transitivité, c'est le degré de confiance qu'à une entité envers l'entité cible.

[Abdul-Rahman and Hailes, 1996], définissent 4 conditions pour que l'affirmation qui suit soit possible : « Si  $A$  a confiance en  $B$  et que  $B$  a confiance en  $C$ , alors  $A$  a confiance en  $C$  » :

- $B$  transmet sa croyance en  $C$  à  $A$  comme une recommandation ;
- $A$  considère  $B$  comme quelqu'un qui est une entité faisant une recommandation ;
- $A$  peut juger de la qualité du degré de confiance que  $B$  a envers  $C$  ;
- La confiance n'est pas absolue,  $A$  peut avoir plus ou moins confiance envers  $C$  que  $B$ .

De plus, Jøsang ajoute que cette transitivité de la confiance n'est effective que si la confiance de référence et la confiance fonctionnelle se fait pour le même contexte [Jøsang and Pope, 2005]. Par exemple, si Marc demande un bon mécanicien à Alice, elle lui proposera quelqu'un en qui elle a confiance dans le domaine de la mécanique et non dans le domaine culinaire, sinon la transitivité de la confiance n'est pas valable.

### 3.2 Réseaux bayésiens

Selon [Pearl, 1988], un réseau bayésien est un graphe direct acyclique  $B = (X, A)$ , où  $X$  est un ensemble de variables et  $A$  est un ensemble d'arcs entre ces nœuds. Chaque variable de  $X$  est un état logique possible

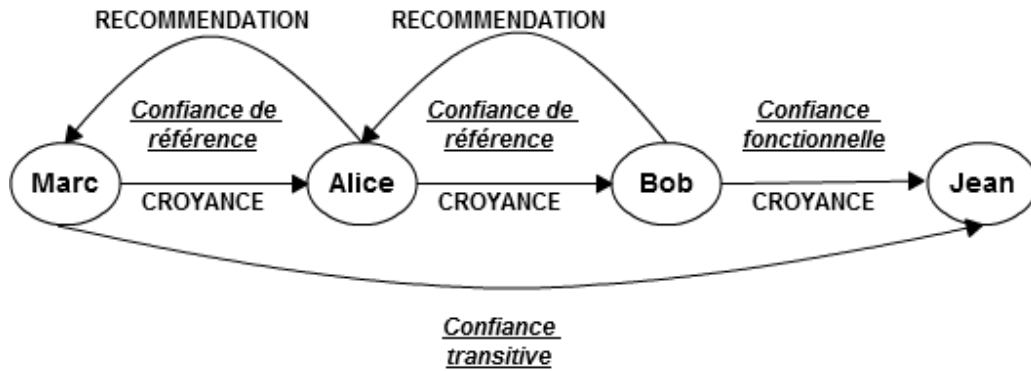


FIGURE 3.1 – Les différents types de confiance.

du monde. Par exemple, VRAI ou FAUX, ou encore FIABLE ou NON FIABLE. S'il y a un lien entre deux variables appartenant à  $X$ , par exemple si  $x, y \in X$  et  $x \rightarrow y \in A$  alors la valeur de  $y$  dépend de  $x$ . Certains modèles de confiance se basent sur cette théorie en percevant la confiance d'une entité envers une autre entité comme une probabilité qu'un événement se déroule.

### 3.2.1 Modèle de Jøsang : bêta réputation

Le modèle de confiance de Jøsang nommée *Subjective logic* [Jøsang and Ismail, 2002] repose sur la théorie de probabilité bayésienne ainsi que la théorie de la croyance qui sera présentée dans une autre partie. L'approche bayésienne repose sur la bêta distribution, permettant d'exprimer la probabilité incertaine qu'une future interaction ou expérience sera positive. En d'autres termes, la bêta distribution représente le fait qu'un résultat futur positif  $x$  est quelque chose d'incertain et qu'il peut être estimé. La bêta distribution est ici utilisée car les valeurs qu'elle prend sont comprises entre 0 et 1.

La bêta distribution, est représentée par une fonction de densité de probabilité (PDF)  $f$  de la probabilité de la variable  $p$ . Cette fonction prend deux paramètres  $\alpha$  et  $\beta$  en entrée et est décrit comme suit :

$$f(p|\alpha\beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1} \quad (3.1)$$

Avec  $0 \leq p \leq 1$ ,  $\alpha > 0$ ,  $\beta > 0$ ,  $\Gamma$  la fonction Gamma et :

$$\alpha = r + 1$$

$$\beta = s + 1$$

où  $r$  est le nombre d'interactions positives et  $s$  le nombre d'interactions négatives.

La probabilité est donc calculée en fonction d'événements précédemment collectés, et la représentation de  $f$  dans un graphe permet de visualiser la probabilité incertaine que l'événement  $x$  se produise dans la prochaine interaction future. Jøsang modélise la confiance par la probabilité subjective qu'une future expérience sera positive. Cette modélisation correspond à la valeur de probabilité d'espérance de la bêta distribution :

$$E(p) = \frac{\alpha}{(\alpha + \beta)} \quad (3.2)$$

Si par exemple on étudie l'événement positif  $x$  et que  $p = 0.8$ , on peut dire que la probabilité que l'événement  $x$  se produise est incertain, mais que la probabilité la plus plausible que  $x$  se produise est de 0.8.

Jøsang se base sur ces définitions pour définir une fonction de réputation correspondant à la bêta distribution 3.1 :

$$\vartheta(p|r_T^X s_T^X) = \frac{\Gamma(r_T^X + s_T^X + 2)}{\Gamma(r_T^X + 1)\Gamma(s_T^X + 1)} p^{r_T^X - 1} (1 - p)^{s_T^X}$$

Avec  $0 \leq p \leq 1$ ,  $r_T^X > 0$ ,  $s_T^X > 0$  et  $\Gamma$  la fonction Gamma.

$\vartheta$  est appelé la réputation de  $T$  par  $X$ .  $r_T^X$  représente les expériences positives de l'agent  $X$  avec la cible  $T$ , tandis que  $s_T^X$  représente les expériences négatives de  $X$  avec  $T$ . Le couple  $(r_T^X, s_T^X)$  représente les paramètres de réputation de  $T$  par  $X$ .

D'après 3.2 Jøsang modélise la valeur de probabilité d'espérance de la fonction de réputation :

$$E(\vartheta(p|r_T^X s_T^X)) = \frac{r_T^X + 1}{r_T^X + s_T^X + 2}$$

La fonction de réputation est bien subjective puisqu'elle se base sur les expériences passées de  $X$  envers  $T$ .

Enfin, Jøsang modélise un score de réputation :

$$Rep(r_T^X s_T^X) = E(\vartheta(p|r_T^X s_T^X)) - 0.5 \times 2 = \frac{r_T^X - s_T^X}{r_T^X + s_T^X + 2}$$

Ce score de réputation est une mesure de la réputation, donc comment un agent est évalué pour ses futures interactions.

De plus, Jøsang introduit deux opérateurs :

- Consensus : qui combine les réputations de deux agents à propos d'une même cible
- Actualisation : qui utilise la transitivité de la croyance. La réputation est calculée en prenant en compte la réputation que l'agent  $A$  a envers l'agent  $B$  et la réputation que l'agent  $B$  a envers la cible  $T$ .

Cette approche est complétée par la théorie de la croyance que Jøsang a introduit. Elle est expliquée dans la partie 3.4.1

### 3.2.2 D'autres systèmes bayésiens

[Mui et al., 2002] ont aussi réalisé un modèle reposant sur la bêta distribution, où la décision de coopération entre un agent et un autre dépend de la réputation basée sur le résultat des rencontres précédentes entre ces deux agents.

[Zeng et al., 2006] ont proposés un système de confiance pour l'encyclopédie en ligne de Wikipédia<sup>1</sup>. Ce système repose sur les réseaux dynamiques bayésiens pour représenter la fiabilité d'un article. Un nœud du réseau correspond à une version  $i$  d'un article écrit par l'auteur  $j$ , au contenu ajouté par rapport à la version  $i - 1$  et au contenu supprimé par rapport à la version  $i - 1$ . Ce système satisfait la propriété de Markov où l'état de l'article  $i$  dépend de l'état de l'article  $i - 1$ . La fiabilité d'une version  $\in [0, 1]$  d'un article dépend alors de la version précédente, de l'auteur de la dernière version, de la quantité de texte ajouté et de texte supprimé. Si un article obtient la note de 0.7 grâce à ces facteurs, cette note sera plutôt entre 0.65 et 0.75 car cette note peut dépendre d'autres facteurs, par exemple l'intérêt de l'auteur pour le domaine. Afin de représenter l'incertain, la fiabilité est représentée grâce à la bêta distribution.

## 3.3 Valeurs discrètes

### 3.3.1 Abdul-Rahman

Dans leur modèle, [Abdul-Rahman and Hailes, 2000] rejettent le fait de définir le niveau de confiance comme une probabilité, car pour eux une probabilité n'a de sens que pour des événements similaires répétés. De même, ils rejettent la transitivité de la confiance. Dans ce modèle, la confiance est définie comme une mesure

1. [www.wikipedia.org](http://www.wikipedia.org)

subjective (croyance) de l'expérience personnelle dans un contexte particulier, et cette mesure subjective est propagée dans le système sous le terme de réputation.

La croyance d'un agent envers un autre agent dans un contexte spécifique est représenté par 4 valeurs : très fiable, fiable, douteux, très douteux. L'expérience directe d'un agent avec un autre agent admet 4 valeurs : très bonne, bonne, mauvaise, très mauvaise.

Le but dans ce modèle est d'obtenir une mesure de la distance sémantique entre la recommandation d'un agent  $b$  pour un agent  $c$  faite à l'agent  $a$  et l'expérience personnelle directe que  $a$  a réellement perçue. Par exemple, si Alice demande à Bob à quel degré elle peut avoir confiance en Jean et que Bob lui répond que Jean est très fiable et qu'Alice s'aperçoit finalement que Jean n'est pas très fiable mais juste fiable, alors Alice adaptera les prochaines recommandations de Bob en les abaissant d'un niveau. Cette distance sémantique se nomme valeur de recommandation de confiance.

Si un agent  $x$  veut connaître la fiabilité d'un auteur  $a_1$  de livre de sciences fictions, voici la méthodologie à suivre dans ce modèle :

1.  $x$  demande des recommandations aux agents qui ont eu une expérience directe avec  $a_1$ .
2. La deuxième étape consiste à obtenir la valeur de recommandation de confiance de chaque agent faisant une recommandation en fonction des recommandations passées.
3. Ensuite, chaque recommandation actuelle de chaque agent reçue par  $x$  sur l'auteur  $a_1$  est adaptée.
4. En fonction de sa recommandation adaptée et de sa valeur de recommandation de confiance, chaque agent se voit attribuer un poids contribuant à la valeur de réputation finale de  $a_1$ .
5. La dernière étape consiste à mettre à jour les expériences entre  $x$  et chaque agent qui a fait une recommandation ainsi que l'expérience directe entre  $x$  et  $a_1$ .

Dans ce modèle, seulement les agents connus participent à la recommandation, les recommandations des agents non-connus serviront à la prochaine recommandation. Les expériences directes servent seulement à calculer la distance sémantique mais ne sont pas réellement combinées aux recommandations. Comme les auteurs rejettent la transitivité de la confiance, les recommandations de recommandations ne sont pas traitées. Étant donné que les valeurs ne sont pas des valeurs numériques, faire des calculs n'est pas facile mais la plus grosse critique pouvant être faite sur ce modèle est la manière d'attribuer les poids, celle-ci est faite ultérieurement sans aucune justification.

### 3.3.2 TidalTrust

[Golbeck, 2005] propose l'algorithme TidalTrust qui est testé dans le réseau social FilmTrust<sup>2</sup> qui est un système de recommandations de films comprenant 400 utilisateurs. Un utilisateur évalue des films entre une demi-étoile et 4 étoiles, et chaque utilisateur peut évaluer sa confiance envers un autre utilisateur avec 10 valeurs discrètes comprises dans l'intervalle  $[1, 10]$ . Pour Golbeck, représenter la confiance par des valeurs discrètes est plus instinctif que de la représenter dans un intervalle continu. TidalTrust est un algorithme qui permet à une source d'inférer le score d'un film  $m$  à travers les scores de recommandations d'évaluateurs. Cet algorithme calcule un score de recommandations qui est une moyenne pondérée. La moyenne pondérée du score d'un film reflète l'opinion, ou encore la recommandation des utilisateurs de la communauté. Pour calculer ce score, TidalTrust procède par plusieurs étapes :

1. Le système cherche les évaluateurs du film  $m$  que la source connaît directement ;
2. S'il n'y a pas de connexions directes entre les évaluateurs de  $m$  et la source, le système cherche les utilisateurs connectés avec la source par un chemin de longueur 2, c'est à dire les utilisateurs connus par les utilisateurs que la source connaît directement. Le processus continue à chercher en augmentant de 1 la longueur du chemin à chaque itération, jusqu'à ce qu'un chemin soit trouvé.
3. Le score de confiance est calculé à partir de tous les utilisateurs compris dans le chemin trouvé.

---

2. <http://trust.mindswap.org/FilmTrust/>

Pour un nœud source donné  $s$  dans l'ensemble des nœuds  $S$ , le score de recommandation  $r_{sm}$  inféré par  $s$  pour un film  $m$  est défini par :

$$r_{sm} = \frac{\sum_{s \in adj(s)} t_{si} \times r_{im}}{\sum_{s \in adj(s)} t_{si}}$$

où  $adj(s) \in S$  sont les nœuds directement connectés à  $s$  dans le réseau.

Voici un exemple où nous considérons que :

- Marc a un score de confiance envers Alice égal à 7 ;
- Alice a un score de confiance envers Bob égal à 4 ;
- Alice a un score de confiance envers David égal à 8 ;
- Bob a un score de confiance envers Jean égal à 8 ;
- David a un score de confiance envers Marie égal à 3 ;
- Jean a évalué le film  $m$  avec 3 étoiles ;
- Marie a évalué le film  $m$  avec 2 étoiles.

La figure 3.2 représente ces affirmations.

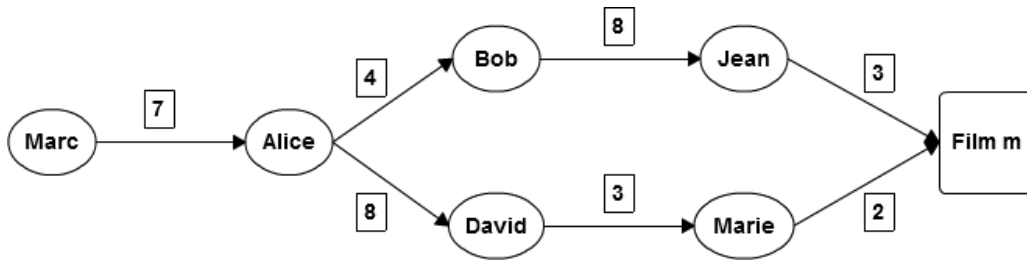


FIGURE 3.2 – Liens de confiance de l'exemple.

Le score de recommandation de Marc à propos du film  $m$  est calculé de la façon suivante :

$$r_{Marc \rightarrow m} = \frac{t_{Marc \rightarrow Alice} \times r_{Alice \rightarrow m}}{t_{Marc \rightarrow Alice}} \quad (3.3)$$

$$= \frac{t_{Marc \rightarrow Alice} \times \left( \frac{t_{Alice \rightarrow Bob} \times r_{Bob \rightarrow m} + t_{Alice \rightarrow David} \times r_{David \rightarrow m}}{t_{Alice \rightarrow Bob} + t_{Alice \rightarrow David}} \right)}{t_{Marc \rightarrow Alice}} \quad (3.4)$$

$$= \frac{t_{Marc \rightarrow Alice} \times \left( \frac{t_{Alice \rightarrow Bob} \times \frac{t_{Bob \rightarrow Jean} \times r_{Jean \rightarrow m}}{t_{Bob \rightarrow Jean}} + t_{Alice \rightarrow David} \times \frac{t_{David \rightarrow Marie} \times r_{Marie \rightarrow m}}{t_{David \rightarrow Marie}} \right)}{t_{Marc \rightarrow Alice}} \quad (3.5)$$

$$= \frac{7 \times \left( \frac{4 \times \frac{8 \times 3}{8} + 8 \times \frac{3 \times 2}{3}}{8 + 4} \right)}{7} \quad (3.6)$$

$$= 2,33 \in [0, 4] \quad (3.7)$$

Pour éviter des chaînes de trop longues tailles des limites peuvent être fixées. Cependant, la longueur du chemin n'est pas pris en compte. En effet, un score obtenu grâce à un utilisateur en lien direct avec la source dans le réseau aura la même valeur qu'un score obtenu avec un chemin de longueur 3 par exemple. De plus, cette approche ne prend pas en compte l'incertain, c'est-à-dire qu'elle ne tient pas compte des différents chemins possibles, mais calcule le score de recommandation avec le premier chemin trouvé. Le risque n'est pas non plus étudié ici, tout comme la décision de faire confiance ; il n'y a pas de seuil permettant à l'utilisateur de choisir ou non le film.

## 3.4 Modèles basés sur la croyance

### 3.4.1 Jøsang : belief model

Jøsang complète son modèle *subjective logic* [Jøsang and Ismail, 2002] par une théorie de la croyance. Une mesure spécifique de la croyance, appelée l'opinion, est calculée pour représenter une croyance spécifique. La mesure de la croyance est utilisée dans des situations d'ignorance et d'incertain. L'opinion est un triplet défini comme suit :

$$\omega_x^A = (b, d, u)$$

où  $b + d + u = 1$  avec  $b \in [0, 1]$  représente la probabilité de vérité,  $d \in [0, 1]$  la probabilité de fausseté et  $u \in [0, 1]$  la probabilité d'incertitude. Grâce à  $u$  une entité peut exprimer son incapacité à évaluer la valeur de probabilité de  $x$ .

Jøsang a réalisé une combinaison du modèle de bêta probabilité et du modèle de croyance :

$$b = \frac{r}{r + s + 2}$$

$$d = \frac{s}{r + s + 2}$$

$$u = \frac{2}{r + s + 2}$$

Le modèle de confiance de Jøsang *subjective logic* qui vient d'être présenté à été utilisé dans des systèmes de commerces électronique [Jøsang, 1999].

### 3.4.2 Modèle de Barber

[Barber and Kim, 2001] est un modèle multi-agents fondé sur la révision de la croyance, il utilise la connaissance à propos de la fiabilité d'une source d'information, la réputation de cette source. La confiance est donc définie ici comme la croyance sur les habilités et l'intention de la source d'information à fournir des informations correctes. Une information est considérée comme fiable si l'agent apportant l'information est considéré comme fiable. Le modèle est basé sur une approche bayésienne, en effet, la croyance de la réputation d'un agent  $a$  à propos d'une source d'information se calcule en fonction de la croyance en la réputation d'agents qui sont en lien avec l'agent  $a$  et avec cette source d'information. Par exemple, l'agent  $a$  demande à l'agent  $b$  sa croyance en la réputation de l'agent  $c$ . A ce moment-là, l'agent  $a$  met en place un processus de révision de la croyance, pour réviser sa croyance en la réputation de  $c$ .

Les auteurs, ont construit un algorithme de révision de la croyance distribuée pour les systèmes multi-agents fondés sur la réputation des sources d'information. Par exemple, si je veux réviser ma croyance envers la source d'information  $q$ , je peux demander aux sources d'informations  $s_1$  et  $s_2$  leur croyance en la réputation de  $q$ . Cette croyance en la réputation de  $q$  est modulée en fonction de ma propre croyance en la réputation de  $s_1$  et  $s_2$ . Grâce à ces critères, le modèle calcule le degré de croyance que  $q$  soit vraie et que  $q$  soit faux. Le résultat final consiste à chercher la valeur de certitude révisée.

## 3.5 Modèles flous

Les modèles flous représentent la confiance ou encore la réputation comme des concepts linguistiques flous. Ils décrivent à quel degré un agent peut être fiable en formalisant des règles reposant sur des mesures floues.

### 3.5.1 Système REGRET

Dans le système de commerce électronique REGRET [Sabater and Sierra, 2002], la réputation est un concept multi dimensionnel. Elle accepte une dimension individuelle, sociale et ontologique. Dans l'environnement, les agents peuvent coopérer, être en compétition ou commercer.



- Dimension individuelle : il s’agit de l’expérience directe entre l’agent et l’agent cible. Un contrat entre deux agents est défini par le prix, la qualité et la date de livraison du produit. La confiance est modélisée en fonction de la différence entre le contrat initial et les résultats de la vente. Cette différence va permettre de classer l’agent cible dans une catégorie, par exemple, la catégorie « fait payer trop chère ».
- Dimension sociale : l’agent fait appel à des témoins qui ont déjà un score de réputation pour l’agent cible ou à des voisins de l’agent cible. Cette dimension fait appel à des règles floues. En fonction de la relation du témoin avec la cible et du degré de cette relation, l’agent assignera un degré de croyance aux informations apportées par le témoin ou le voisin. Par exemple,
 

SI le témoin coopère fortement avec l’agent cible  
ALORS les informations du témoin sont très mauvaises.
- Dimension ontologique : combine la réputation de différents aspects afin de calculer une réputation complexe. Par exemple, mettre l’agent cible dans une nouvelle catégorie qui est une combinaison de deux catégories.

## 3.6 Modèles de flux

Les modèles de flux calculent la confiance grâce à des itérations transitives à travers des boucles ou encore de longues chaînes. Dans certains modèles la confiance totale est constante à travers toute la communauté, donc l’augmentation du score de confiance d’une des entités se fait au détriment des autres.

### 3.6.1 Google PageRank

L’algorithme PageRank de google [Page et al., 1998], permet de classer les résultats d’une recherche de pages web. Cette classification est basée sur la confiance, puisqu’elle est basée sur les liens entre pages. Dans une page, le lien vers une autre page est vu comme une preuve de confiance. Le PageRank d’une page mesure en fait la probabilité qu’à un surfeur, qui est sur une page donnée, d’arriver sur cette page. C’est une représentation de chaîne de Markov avec le suivi d’un très grand nombre de liens. Cette méthode repose sur l’hypothèse qu’un surfeur qui se retrouve sur une page, choisit aléatoirement de suivre un lien de cette page. Voici la méthode de calcul du PageRank :

$$R(u) = c \sum_{v \in B_u} \frac{R(v)}{N_v}$$

Avec  $R(u)$  le pageRank de la page  $u$ ,  $c$  un facteur de normalisation, et  $N_v$  le nombre de liens présents dans la page  $v$ .

Le PageRank d’une page est d’autant plus important que les PageRank des pages qui pointent vers elles sont importants. En effet, chaque page qui pointe vers cette page contribue à une fraction du PageRank de cette page. Dans ce calcul, l’attribution du PageRank d’une page est récursif. On peut résumer ce calcul par trois affirmations :

- Plus une page a un PageRank important, plus elle contribue au PageRank des pages vers lesquelles elle pointe ;
- Plus une page contient de lien, moins elle contribue au PageRank des pages vers lesquelles elle pointe ;
- Plus une page reçoit de liens, plus son PageRank est important.

### 3.6.2 EigenTrust

Le modèle [Kamvar et al., 2003] a été réalisé pour les systèmes Pair-à-Pair, il permet d’assigner à chaque pair une valeur globale unique dérivée grâce à l’histoire de ce dernier. Pour être constant, ce score ne requiert pas la somme de tous les scores de confiance. Les scores de réputation dans ce modèle sont calculés à travers une longue chaîne transitive de calculs répétés et itératifs jusqu’à ce que les scores de confiance des agents convergent vers une valeur stable.

L’interaction d’un pair  $i$  avec un pair  $j$  peut être positive ou négative. Le nombre d’interactions positives qu’a eu  $i$  avec  $j$  est noté  $sat(i, j)$  et le nombre d’interactions négatives qu’a eu  $i$  avec  $j$  est noté  $unsat(i, j)$ .

Le score local de croyance normalisé est noté :

$$c_{ij} = \frac{\max(s_{ij}, 0)}{\sum_{l \in L} \max(s_{il}, 0)} \in [0, 1]$$

avec  $L$  l'ensemble des pairs qui ont eu une expérience direct avec  $i$  et  $s_{ij} = \text{sat}(i, j) - \text{unsat}(i, j)$ .

Un score de confiance hors du groupe local de  $i$  peut ensuite être calculé :

$$t_{ik} = \sum_{j \in L} c_{ij} c_{jk}$$

Ce score permet de réunir dans un score unique la confiance fonctionnelle et la confiance de référence.

Dans le modèle,  $C = [c_{ij}]$  représente la matrice de tous les scores de confiance normalisée dans la communauté. Tandis que  $\vec{c}_i$  représente le vecteur des valeurs de confiances locales de  $i$ .

$\vec{t}_i$  quant à lui représente le vecteur contenant les valeurs de confiance  $t_{ik}$  où  $i$  et  $k$  sont séparés par  $n$  nœuds intermédiaires :

$$\vec{t}_i = C^n \vec{c}_i$$

Lorsque  $n$  est important, le vecteur  $\vec{t}_i$  converge vers la même valeur quel que soit le pair  $i$ . Le vecteur  $\vec{t}_i$  correspond à un vecteur de confiance globale et correspond à la valeur de confiance globale qu'a la communauté envers  $k$ . Les limites de ce modèle correspondent au fait que le score normalisé supprime les scores de satisfaction négative. De plus, un nouvel arrivant aura la même valeur qu'un mauvais pair.

### 3.7 Un modèle plus général : Marsh

Dans son modèle [Marsh, 1994], Marsh se place dans le courant des systèmes multi-agents, mais ne traite de la confiance qu'entre deux agents. Pour se faire, il définit trois types de confiance :

- La confiance basique constitue la disposition générale d'un agent  $x$  à faire confiance

$$T_x \in [-1, 1[$$

- La confiance générale est la confiance d'un agent  $x$  envers un autre agent  $y$  sans prendre en compte une situation spécifique. Cette connaissance implique la connaissance de l'autre agent notée  $K_x(y)$ . Marsh accepte la méfiance complète correspondant à  $-1$  mais rejette la confiance complète  $1$  :

$$T_x(y) \in [-1, 1[$$

- La confiance situationnelle est la confiance d'un agent  $x$  envers un autre agent  $y$  dans une situation donnée  $\alpha$  :

$$T_x(y, \alpha) \in [-1, 1[$$

Pour la confiance situationnelle, Marsh utilise la définition de [Rempel and Holmes, 1986], où il affirme que la confiance se trouve dans les gens et dans des situations spécifiques.

Marsh propose une méthode pour calculer la confiance situationnelle qui prend en compte la confiance générale qu'un agent a envers un autre agent, ainsi que l'importance et l'utilité de la situation :

$$T_x(y, \alpha) = U_x(\alpha) * I_x(\alpha) * \widehat{T_x(y)}$$

Avec :

- $U_x(\alpha)$  : l'utilité de la situation  $\alpha$  pour  $x$ , c'est-à-dire, ce que va lui rapporter cette situation ;
- $I_x(\alpha)$  : l'importance de la situation  $\alpha$  pour  $x$ , c'est-à-dire, est-ce qu'il y a beaucoup de chances que l'agent obtienne des bénéfices ? ;
- $\widehat{T_x(y)}$  : la confiance générale qui prend en compte toutes les situations possibles, correspondant à chaque  $T_x(y, \alpha)$ . Différentes manières de calculer  $\widehat{T_x(y)}$  sont proposées. En fonction de si l'agent est optimiste, il le calculera en fonction du plus élevé de tous les  $T_x(y, \alpha)$ , si l'agent est pessimiste, il le calculera en fonction du moins élevé de tous les  $T_x(y, \alpha)$  et si l'agent est rationnel, il calculera une moyenne des  $T_x(y, \alpha)$ .

Dans son modèle, un agent peut coopérer avec un autre si sa confiance situationnelle est supérieure à un seuil de coopération. Selon Marsh, « Le seuil de coopération est une mesure subjective, temperé par des croyances objectives. » Pour calculer le seuil de coopération, Marsh intègre les risque perçus de la situation, les compétences perçues de l'autre agent dans la situation, et la reciprocité :

$$CooperationThreshold_x(\alpha) = \left( \frac{PerceivedRisk(\alpha)}{PerceivedCompetence_x(y, \alpha) \times \widehat{T_x(y)}} \times I_x(\alpha) \right) \times ReductionPercentage(y, \alpha)$$

avec  $ReductionPercentage(y, \alpha)$  qui correspond à la retroprocité. Dans [Marsh, 1994], Marsh explique comment calculer  $PerceivedRisk(\alpha)$ ,  $PerceivedCompetence_x(y, \alpha)$  et  $ReductionPercentage(y, \alpha)$ .

Les travaux de Marsh sont encore aujourd'hui une référence dans les modèles sur la confiance. En effet, il a réuni dans son modèle de nombreux concepts introduit par différents auteurs : le risque, les compétences, l'utilité, l'importance, la situation, la retroprocité. De plus, il a étudié le facteur de la mémoire et comment calculer la confiance entre un autre agent en fonction de la confiance qu'il nous attribue grâce à un processus itératif.

## Chapitre 4

# Les systèmes commerciaux

### 4.1 Ebay

Ebay<sup>1</sup> est un célèbre site de vente aux enchères où des vendeurs particuliers vendent des articles divers à des acheteurs. Après chaque vente, l'acheteur et le vendeur se donne mutuellement un score pour évaluer la qualité de l'échange. Ebay est un système de confiance centralisée basée sur la réputation. Son fonctionnement est très simple puisque le score global de réputation est calculé en soustrayant les scores négatifs aux scores positifs. Cependant, ce système introduit de nombreux biais, il y a un effet de réciprocité entre l'acheteur et le vendeur ; de peur d'obtenir une note négative très peu ose mettre de tels scores. Une solution serait de ne noter que les vendeurs par exemple. De plus, un utilisateur avec un score très mauvais peut se réinscrire sous un nouveau nom et ainsi effacer son passé non fiable . Le dernier biais concerne les nouveaux arrivants dans le système ; un nouvel arrivant qui n'a encore participé à aucune vente obtient par défaut la note de 0. Par contre, un utilisateur qui aura un score de 20 en négatif et 30 en positif aura une meilleure réputation dans le système que le nouvel arrivant, ce qui n'est pas du tout représentatif de la réalité. En effet, si vous devez décider de vous faire opérer par un nouveau chirurgien qui n'a pas encore réalisé d'opérations ou par un chirurgien qui a tué 20 patients sur 50, qui choisiriez-vous ? La réponse paraît simple.

### 4.2 Amazon

Amazon<sup>2</sup> est un site de vente en ligne qui s'est d'abord spécialisé dans les livres pour s'élargir au matériel informatique, vêtement, accessoires beauté, etc. Un membre peut noter un produit avec un système d'étoile allant de 1 à 5. Des non membres peuvent voter si le commentaire a été utile ou non, et les commentaires sont classés en fonction de leur utilité. A côté du commentaire d'un membre, il est indiqué combien de personnes ont voté ce commentaire utile sur le nombre de vote au total.

Cependant, des problèmes apparaissent du fait que des non membres peuvent voter, en effet une personne peut voter plusieurs fois pour un même commentaire et le faire monter dans le classement. Une solution trouvée par les créateurs du site a été de limiter un vote par adresse IP pour un commentaire donné, mais cette solution reste partielle puisqu'un utilisateur peut la contourner en utilisant un autre ordinateur.

### 4.3 Epinions

Epinions<sup>3</sup> est un site recueillant des critiques sur tout type de produits allant de l'électroménager à des films. Tout membre inscrit peut faire une critique sur un produit avec des rubriques définies ultérieurement et spécifiquement pour chaque catégorie d'articles, comme la qualité du son ou de l'image pour une télévision. Le produit obtient ensuite une note grâce aux critiques. Chaque critique peut être évaluée par un autre membre inscrit, grâce aux quatre niveaux : pas utile, un peu utile, bien utile et très utile. Les critiques seront donc ensuite classées en fonction de leur qualité. De plus, un membre peut mettre un autre membre dans son réseau

---

1. <http://www.ebay.fr>

2. <http://amazon.com>

3. <http://www.epinions.com>

de confiance. Ainsi, chaque membre dispose d'un réseau de confiance, où les membres présents sont ceux qu'il considère les plus fiables. Ce réseau est composé de plusieurs niveaux :

- Membres de confiance élevée : j'ai directement mis ces membres dans mon réseau de confiance ;
- Membres de confiance moyenne : ces membres sont ceux qui ont été directement mis dans le réseau de confiance de mes membres de confiance élevée ;
- Membres de confiance faible : ces membres sont ceux qui ont été directement mis dans le réseau de confiance de mes membres de confiance moyenne ;
- Membres de confiance inexistantes : je n'ai aucun lien avec ces membres.

Ces catégories permettent un affichage personnalisé de la page de critiques d'un produit où les membres à qui j'ai directement confiance apparaissent en première position. Cependant, ce fonctionnement indirect des réseaux de confiance peut être discuté, car comme il a été vu plus haut, la confiance n'est pas transitive.

## Chapitre 5

# Conclusion

Ce papier est un état de l'art sur le terme complexe de la confiance. Nous avons tout d'abord défini la notion de confiance autour de la littérature, notamment grâce à [Deutsch, 1962] qui définit la confiance en une perception individuelle des coûts et bénéfices d'une situation, à [Luhmann, 1979] qui avance que la confiance permet de diminuer la complexité sociale et enfin à [Gambetta, 2000] qui voit la confiance comme une mesure subjective de l'occurrence d'un certain comportement. Des caractéristiques générales de la confiance ont ensuite émergées telles que la notion de risque, de coopération ou encore de réciprocité. Ces fondements nous ont permis par la suite de présenter les systèmes utilisant la confiance, ceux-ci sont classés en deux grandes catégories : les systèmes de confiance basés sur la sécurité et ceux basés sur la réputation. Nous nous sommes concentrés sur les systèmes de confiance basés sur la réputation qui peuvent être des systèmes centralisés ou distribués. Pour que ces systèmes puissent utiliser la confiance afin de prendre des décisions, évaluer la fiabilité d'une entité par exemple, ils doivent posséder un moteur de calcul. Nous avons donc naturellement présentés des modèles de calcul de la confiance. La confiance peut être représentée comme une valeur continue correspondant à une probabilité par exemple, une étiquette nominative, une valeur discrète ou encore une valeur floue. La question de la transitivité de la confiance a aussi été traitée, en effet, la confiance n'est mathématiquement pas transitive [Marsh, 1994], alors que des auteurs l'utilisent comme un concept transitif [Jøsang and Pope, 2005]. Ces auteurs contournent en réalité ce problème en utilisant la notion de recommandation.

La confiance représentée avec des valeurs discrètes est plus facile à interpréter et à attribuer pour des utilisateurs mais les calculs sur ces valeurs sont difficiles et ne prennent pas en compte l'incertain ou encore le risque. Les modèles utilisant des valeurs floues sont implémentables dans des systèmes multi-agents mais difficilement applicable à un système avec de réels utilisateurs. Les réseaux bayésiens utilisant des valeurs continues sont les plus adaptés pour de réels utilisateurs et pour appliquer des calculs, même si leur représentation est parfois complexe. Ces systèmes permettent par exemple de prendre en compte les expériences directes et les recommandations, ainsi que l'incertitude d'une prévision. Le modèle le plus complet, bien que limité à deux agents dans les systèmes multi-agents, correspond à [Marsh, 1994], qui prend en compte le risque, les compétences, la réciprocité, la coopération, l'importance, l'utilité et la fiabilité même de l'agent qui mesure la confiance de l'autre agent.

Finalement, des modèles commerciaux utilisant la confiance pour recommander des produits ou encore des vendeurs par le biais de la réputation, ont été présentés. Ces sites ont un fonctionnement plutôt satisfaisant mais connaissent de nombreux biais. Par exemple, quelqu'un qui a une réputation négative peut avoir une meilleure réputation qu'un nouvel arrivant ou encore quelqu'un peut redevenir un membre neutre en changeant d'identité.

[Cojan et al., 2011, Dufour-Lussier et al., 2011, Gaillard et al., 2011]

# Bibliographie

- [Abdul-Rahman and Hailes, 1996] Abdul-Rahman, A. and Hailes, S. (1996). A Distributed Trust Model. *Proceedings of the 1997 workshop on New security paradigms NSPW 97*, pages 48–60.
- [Abdul-Rahman and Hailes, 2000] Abdul-Rahman, A. and Hailes, S. (2000). Supporting Trust in Virtual Communities. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 00(c) :9.
- [Artz and Gil, 2007] Artz, D. and Gil, Y. (2007). A survey of trust in computer science and the Semantic Web. *Web Semantics : Science, Services and Agents on the World Wide Web*, 5(2) :58–71.
- [Barber, 1983] Barber, B. (1983). *The logic and limits of trust*. Rutgers University Press.
- [Barber and Kim, 2001] Barber, K. and Kim, J. (2001). Belief revision process based on trust : Agents evaluating reputation of information sources. *Trust in Cyber-societies*, pages 73–82.
- [Blaze et al., 1999] Blaze, M., Feigenbaum, J., and Keromytis, A. D. (1999). KeyNote : Trust Management for Public-Key Infrastructures. *Lecture Notes in Computer Science*, 1550(59-63) :59—63.
- [Blaze et al., 1996] Blaze, M., Feigenbaum, J., and Lacy, J. (1996). Decentralized trust management. *Proceedings 1996 IEEE Symposium on Security and Privacy*, (96-17) :164–173.
- [Blaze et al., 2003] Blaze, M., Ioannidis, J., and Keromytis, A. (2003). Experience with the keynote trust management system : Applications and future directions. *Trust Management*, pages 1071–1071.
- [Boon and Holmes, 1991] Boon, S. D. and Holmes, J. G. (1991). The dynamics of interpersonal trust : Resolving uncertainty in the face of risk. In Hinde, R. A. and Groebel, J., editors, *Cooperation and prosocial behaviour*, chapter 11, pages 190–211. Cambridge University Press.
- [Boyd and Richerson, 2009] Boyd, R. and Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society of London - Series B : Biological Sciences*, 364(1533) :3281–3288.
- [Boyle and Bonacich, 1970] Boyle, R. and Bonacich, P. (1970). The development of trust and mistrust in mixed-motive games. *Sociometry*, 33(2) :123–139.
- [Cojan et al., 2011] Cojan, J., Dufour-Lussier, V., Gaillard, E., Lieber, J., Nauer, E., and Toussaint, Y. (2011). Knowledge extraction for improving case retrieval and recipe adaptation. In *Computer Cooking Contest Workshop*, London, Royaume-Uni.
- [Deutsch, 1962] Deutsch, M. (1962). Cooperation and trust : Some theoretical notes. In Jones, M. R., editor, *Nebraska Symposium On Motivation*, volume 10, pages 275–320. University of Nebraska Press.
- [Dufour-Lussier et al., 2011] Dufour-Lussier, V., Lieber, J., Nauer, E., and Toussaint, Y. (2011). Improving case retrieval by enrichment of the domain ontology. In *19th International Conference on Case Based Reasoning - ICCBR'2011*, London, Royaume-Uni.
- [Gaillard et al., 2011] Gaillard, E., Lieber, J., and Nauer, E. (2011). Adaptation knowledge discovery for cooking using closed itemset extraction. In *The Eighth International Conference on Concept Lattices and their Applications - CLA 2011*, Nancy, France.
- [Gambetta, 2000] Gambetta, D. (2000). Can we trust trust. *Trust Making and Breaking Cooperative Relations*, chapter13p :213–237.
- [Gerck, 1997] Gerck, E. (1997). Trust Properties. <http://mcwg.org/mcg-mirror/trustdef.htm>.
- [Golbeck, 2005] Golbeck, J. A. (2005). *Computing and applying trust in web-based social networks*. PhD thesis, University of Maryland.

- [Grandison, 2003] Grandison, T. W. A. (2003). Trust management for internet applications. Technical report.
- [Jøsang, 1999] Jøsang, A. (1999). Trust-based decision making for electronic transactions. In *Proceedings of the 4th Nordic Workshop on Secure Computer Systems NORDSEC'99 Stockholm University Sweden*. Citeseer.
- [Jøsang and Ismail, 2002] Jøsang, A. and Ismail, R. (2002). The beta reputation system. *Proceedings of the 15th Bled Electronic Commerce Conference*, 160 :324–337.
- [Jøsang and Pope, 2005] Jøsang, A. and Pope, S. (2005). Semantic constraints for trust transitivity. In *Reproduction*, volume 43, pages 59–68. Australian Computer Society, Inc.
- [Jøsang et al., 2007] Jøsang, A., Ismail, R., and Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2) :618 – 644. <ce :title>Emerging Issues in Collaborative Commerce</ce :title>.
- [Kamvar et al., 2003] Kamvar, S. D., Schlosser, M. T., and Garcia-Molina, H. (2003). The Eigentrust algorithm for reputation management in P2P networks. *Proceedings of the twelfth international conference on World Wide Web WWW 03*, 03 :640.
- [Luhmann, 1979] Luhmann, N. (1979). *Trust and power : two works*. Number pts. 1-2 in UMI Books on Demand. Wiley.
- [Marsh, 1994] Marsh, S. P. (1994). *Formalising Trust as a Computational Concept*. PhD thesis, University of Stirling.
- [McKnight and Chervany, 1996] McKnight, D. H. and Chervany, N. L. (1996). The meanings of trust. *Measurement*, 55455(612) :86.
- [Mui et al., 2002] Mui, L., Mohtashemi, M., and Halberstadt, A. (2002). A computational model of trust and reputation. *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 00(c) :2431–2439.
- [Page et al., 1998] Page, L., Brin, S., Motwani, R., and Winograd, T. (1998). The PageRank Citation Ranking : Bringing Order to the Web. *World Wide Web Internet And Web Information Systems*, 54(2) :1–17.
- [Pearl, 1988] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference*, volume 88 of *Morgan Kaufmann Series in Representation and Reasoning*. Morgan Kaufmann.
- [Rempel and Holmes, 1986] Rempel, J. and Holmes, J. (1986). How do i trust thee ? *Psychology Today*, 28-34.
- [Resnick et al., 2000] Resnick, P., Kuwabara, K., Zeckhauser, R., and Friedman, E. (2000). Reputation systems. *Communications of the ACM*, 43(12) :45–48.
- [Sabater and Sierra, 2002] Sabater, J. and Sierra, C. (2002). Reputation and social network analysis in multi-agent systems. *Proceedings of the first international joint conference on Autonomous agents and multiagent systems part 1 AAMAS 02*, 15-19 :475.
- [Viljanen, 2005] Viljanen, L. (2005). Towards an ontology of trust. In *Proceedings of the 2nd International Conference on Trust, Privacy and Security in Digital Business (TrustBus'05)*.
- [Williams, 1988] Williams, B. (1988). Formal Structures and Social Reality. In Gambetta, D., editor, *Making Sense of Humanity*, pages 3–13. Cambridge University Press.
- [Zeng et al., 2006] Zeng, H., Alhossaini, M. A., Ding, L., Fikes, R., and McGuinness, D. L. (2006). Computing trust from revision history. In *Proceedings of the 2006 International Conference on Privacy Security and Trust Bridge the Gap Between PST Technologies and Business Services PST 06*, number January, page 1. Proceedings of the 2006 International Conference on Privacy, Security and Trust : Bridge the Gap Between PST Technologies and Business Services, ACM Press.