

S&F: Évaluation de la fiabilité des sources et des faits

Quentin Elsaesser¹ Patricia Everaere² Sébastien Konieczny¹

¹ CRIL, CNRS - Université d'Artois, France

² CRIStAL, Université de Lille, France

elsaesser@cril.fr

patricia.everaere-caillier@univ-lille.fr

konieczny@cril.fr

Résumé

Dans ce travail, nous proposons une famille de méthodes qui permettent de calculer conjointement la fiabilité d'un ensemble de sources d'information et la confiance des faits sur un ensemble d'objets, en confrontant les points de vue des sources. Nous utilisons une méthode de vote pour l'évaluation de la confiance des sources, en utilisant les arguments du théorème du jury de Condorcet afin d'identifier la vérité et les sources fiables. Nous discutons des propriétés théoriques générales que de tels opérateurs devraient satisfaire, et nous étudions quelles sont les propriétés satisfaites par nos méthodes. Nous proposons une étude expérimentale qui montre que nous sommes plus performants que les méthodes de l'état de l'art pour trouver la vérité parmi les faits possibles. Nous montrons que nous pouvons également évaluer de manière adéquate la fiabilité des sources d'information.

Abstract

In this work we propose a family of methods that allow to jointly compute the reliability of a set of information sources and the confidence of the facts on a set of objects, by confronting the sources points of view. We use a (scoring-based) voting method for the evaluation of the trust of the sources, using Condorcet's Jury Theorem arguments in order to identify the truth and the reliable sources. We discuss general theoretical properties that such operators should satisfy, and we study what are the properties satisfied by our methods. We provide an experimental study that shows that we perform better than state of the art methods on the task of finding the truth among the possible facts. We show that we can also adequately evaluate the reliability of the sources of information.

1 Introduction

Il existe de nombreuses applications où l'on reçoit des informations (généralement contradictoires) de différentes sources. À l'aide de ces informations, on doit se forger une opinion. Dans cette situation, une façon standard de résoudre les conflits est de faire confiance aux sources les plus fiables. Nous proposons une définition de la fiabilité basée sur les informations disponibles. Cette définition peut être utile pour évaluer la fiabilité d'un agent dans un système multi-agents ou sur un réseau social, mais aussi d'une source sur le web, dans un journal, etc.

Plus précisément, nous considérons un ensemble de sources qui nous fournissent des informations (que nous appellerons *faits*) sur différentes questions (que nous appellerons *objets*). Notre objectif est d'évaluer à la fois la fiabilité des sources et la fiabilité des faits, ce qui nous permet ensuite de trouver les réponses correctes aux différentes questions (objets).

Il existe des travaux antérieurs qui utilisent la même structure (sources/faits/objets), mais leur objectif est uniquement de trouver la vérité parmi les faits [33, 28].

Pour trouver cette vraie information, nous nous appuyons sur l'idée du théorème du jury de Condorcet [6], selon lequel il est plus probable que la majorité des individus choisissent la bonne solution. L'intuition est la suivante : supposons que parmi 10 sources de fiabilité égale, 8 vous disent que la *Capitale de l'Australie* est *Canberra*, et 2 vous disent que c'est *Sydney*. Suivre ce que dit la majorité est le moyen le plus sûr de trouver la vérité. Le théorème du jury de Condorcet requiert un grand nombre d'hypothèses (toutes les sources ont la même fiabilité, elles sont toutes fiables (c'est-à-dire qu'elles ont plus de 50% de chances de trouver la vérité), elles sont indépendantes et le choix

ne porte que sur deux possibilités). Cependant, toutes ces hypothèses peuvent être plus ou moins assouplies [17, 35, 23, 4, 10, 12, 2, 5, 3, 19, 21, 34, 14, 15, 9].

Dans cet article, nous supposons qu’au départ, nous n’avons aucune information sur la fiabilité des sources, et nous définissons une procédure itérative pour déterminer leur fiabilité. Au début, nous attribuons la même fiabilité à toutes les sources, puis nous comparons les réponses aux différentes questions et nous utilisons cet argument du "théorème du jury de Condorcet" pour récompenser les sources qui fournissent des informations (*faits*) confirmées par d’autres, et qui sont donc plus susceptibles d’être vraies. Ensuite, nous itérons le processus avec la nouvelle fiabilité des sources jusqu’à convergence.

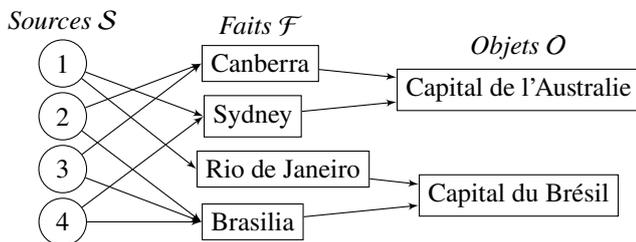


FIGURE 1 – Sources, faits et objets

Pour illustrer ce processus, prenons l’exemple de la figure 1, où quatre sources donnent des informations sur deux objets : *Capitale du Brésil* et *Capitale de l’Australie*. Notons qu’il y a d’abord égalité pour *Capitale de l’Australie*, deux sources donnant *Canberra* et deux sources donnant *Sydney*. Mais nous pouvons utiliser le résultat de l’autre objet. Il y a une majorité pour *Brasilia*, donc *Brasilia* est considéré comme le bon fait, et les sources qui donnent ce fait sont favorisées par rapport à celles qui donnent *Rio de Janeiro*. Et, à la prochaine itération, nous pourrions les départager sur *Capitale de l’Australie* puisque des sources plus fiables affirment *Canberra*.

Plus précisément, à chaque itération, les sources donnent une certaine force aux faits qu’elles affirment sur les différents objets. Cette force est la fiabilité (actuelle) de la source. Ainsi, on peut classer pour chaque objet les faits correspondants du plus fiable au moins fiable, en utilisant simplement la somme des forces obtenues. Ensuite, pour chaque objet, un vote est organisé pour récompenser les sources. Nous utilisons des méthodes de vote par scores (*scoring-voting rules*) afin d’associer un nombre à chaque rang de faits. La plus simple est la règle de la majorité simple, où seuls les faits les plus fiables rapportent un score de 1 aux sources correspondantes, et tous les autres n’obtiennent rien (0). La nouvelle fiabilité de chaque source est calculée en combinant tous ces scores. Nous devons faire un choix entre deux normalisations pour cette étape : une qui favorise les sources qui fournissent le plus d’informations, et une qui favorise les sources qui sont plus prudentes et qui font peu

d’erreurs. Ensuite, une nouvelle itération commence avec la fiabilité actualisée de chaque source.

L’évaluation de la fiabilité peut être utilisée pour décider de faire confiance ou non à une source. Il s’agit d’un élément important pour évaluer la confiance (*trust*) directe (la confiance peut être divisée en confiance directe – obtenue par interaction directe avec un agent, et en confiance indirecte – évaluation obtenue par un tiers [20]). De nombreux travaux s’intéressent à la prise en compte de la confiance et de la réputation dans les décisions et les processus de raisonnement [31, 1, 26, 29, 24, 30, 11]. Mais il existe peu de travaux sur la manière d’évaluer cette confiance directe à partir de preuves, et notre travail peut être très utile à cet égard.

Après avoir présenté nos méthodes S&F (pour *Sources & Faits*), nous discutons des propriétés logiques pour caractériser les méthodes intéressantes qui visent à évaluer la fiabilité des sources et des faits. Nous passons en revue les propriétés qui ont été proposées par [28], et nous expliquons pourquoi certaines d’entre elles ne sont pas adaptées à ce contexte. Nous proposons également de nouvelles propriétés requises pour toutes les méthodes, ainsi que certaines propriétés qui caractérisent des sous-classes intéressantes. Nous vérifions ensuite quelles propriétés sont satisfaites par nos méthodes.

Outre cette évaluation formelle, nous proposons également une étude expérimentale. L’idée est de tester si nous pouvons atteindre cet objectif d’évaluation de la fiabilité des sources et des faits en pratique. Il n’y a pas beaucoup de jeux de données réels qui peuvent être utilisés pour cette tâche, mais nous testons nos méthodes sur deux de ces jeux de données. Ensuite, nous testons également nos méthodes sur des jeux de données que nous avons générés, ce qui nous permet de tester beaucoup plus de paramètres.

Et les résultats sont bons. Nous montrons que pour les tâches liées à la recherche des faits réels, nous sommes meilleurs que les méthodes existantes. Mais, contrairement aux méthodes existantes, nous pouvons également donner une bonne évaluation de la fiabilité des sources.

2 Préliminaires

Nous considérons trois ensembles S , \mathcal{F} et O respectivement appelés *Sources*, *Faits* et *Objets*. Les *Sources* représentent les agents (humains ou artificiels) qui fournissent les informations. Les *Objets* sont les questions sur lesquelles nous aimerions obtenir des informations et les *Faits* sont les réponses possibles. Sur chaque objet, les faits sont distincts et exclusifs : chaque source ne peut affirmer qu’un seul fait par objet.

Ces objets+faits peuvent donc être considérés comme des questions+réponses ou comme des variables+valeurs. Il s’agit ici d’une question de vocabulaire. Nous utilisons celui utilisé dans les travaux précédents [33, 28].

Définition 1 Soit $G = (V, E)$ un graphe orienté avec $V = \mathcal{S} \cup \mathcal{F} \cup \mathcal{O}$ et $E \subseteq (\mathcal{S} \times \mathcal{F}) \cup (\mathcal{F} \times \mathcal{O})$, t.q. :

- Pour chaque fait $f \in \mathcal{F}$ il y a un objet unique $o \in \mathcal{O}$ avec $(f, o) \in E$.
- Une source $s \in \mathcal{S}$ peut affirmer au plus un fait par objet $o \in \mathcal{O}$ (c'est-à-dire $\forall s \in \mathcal{S}$ il n'y a pas de $f_1 \in \mathcal{F}, f_2 \in \mathcal{F}$ t.q. $\{(f_1, o), (f_2, o), (s, f_1), (s, f_2)\} \subseteq E$). $(s, f) \in E$ signifie que la source s affirme que le fait f est la réponse correcte pour l'objet correspondant. Il est possible qu'un fait ne soit affirmé par aucune source.

Pour une meilleure lisibilité, nous utiliserons les notations suivantes : $src(f) = \{s \in \mathcal{S} : (s, f) \in E\}$, $fct(s) = \{f \in \mathcal{F} : (s, f) \in E\}$, $fct(o) = \{f \in \mathcal{F} : (f, o) \in E\}$, $obj(f) = \{o \in \mathcal{O} : (f, o) \in E\}$.

Lorsque ce n'est pas évident (par exemple lorsque nous avons plus d'un graphe), nous pouvons spécifier le graphe, et nous écrivons, pour le graphe G , $src_G(f)$, $fct_G(s)$, $fct_G(o)$ et $obj_G(f)$ à la place des notations ci-dessus.

Nous notons $r_G(s) \in [0, 1]$ la fiabilité d'une source s dans le graphe G . Nous notons $c_G(f) \in \mathbb{R}^+$ la confiance d'un fait f dans le graphe G .¹

3 Travaux Connexes

Il existe dans la littérature des algorithmes de *découverte de la vérité* qui visent à identifier les faits réels.

Truth Finder[33] est un algorithme itératif qui met à jour le score des sources et des faits à chaque itération. Cette méthode se concentre sur la confiance des faits pour trouver la vérité. Avec *Truth Finder*, la fiabilité d'une source est la confiance moyenne dans les faits affirmés par cette source. Pour la confiance d'un fait, les auteurs supposent que les faits peuvent se soutenir mutuellement, auquel cas la confiance augmente, mais diminue si les faits se contredisent. *Truth Finder* permet également de prendre en compte la similarité entre les faits. Nous n'examinons pas cette possibilité dans ce travail, nous laissons donc cet aspect de côté (voir [32] pour une discussion sur les méthodes de découverte de la vérité, où il y a d'autres paramètres qui peuvent être pris en compte pour calculer la fiabilité).

Hubs and Authorities [13] est une méthode définie pour classer les pages web mais peut aussi être utilisée pour la recherche de la vérité. Il s'agit également d'une méthode itérative, qui définit deux scores différents pour une page. *Hub* (qui correspond aux sources) favorise les pages qui pointent vers de nombreuses autres pages, et *authority* (qui correspond aux faits) favorise les pages qui sont pointées par de nombreux hubs différents.

Sums [22] est basé sur *Hubs and Authorities*. La principale différence réside dans la manière dont la fiabilité

des sources et des faits est normalisée. Avec *Sums*, la fiabilité d'une source est la somme de la confiance des faits qu'elle affirme ($r(s) = \sum_{f \in fct(s)} c(f)$), et la confiance d'un fait f est obtenue comme la somme de la fiabilité des sources qui l'affirment ($c(f) = \sum_{s \in src(f)} r(s)$). Ces résultats sont ensuite normalisés par la valeur maximale obtenue (par $\max(r(s)) \in \mathcal{S}$ pour $r(s)$ et par $\max(c(f)) \in \mathcal{F}$ pour $c(f)$).

Booth et Singleton ont été les premiers à proposer une approche axiomatique du problème de la *découverte de la vérité* dans [27] et [28]. Ils proposent également une nouvelle méthode, appelée *Unbounded-Sums*, qui est basée sur *Sums*, mais où ils ne normalisent pas le score.

Nous comparerons les résultats expérimentaux de nos méthodes aux résultats des algorithmes *Hubs and Authorities*, *Truth Finder*, *Sums* et *Unbounded-Sums*. Nous nous comparerons également avec *Voting*, la méthode naïve qui choisit le fait avec le plus de soutien (arêtes dans le graphe) sur chaque objet.

4 Méthodes S&F

Dans cette section, nous présentons nos méthodes, composées de deux étapes. Tout d'abord, la confiance des faits est calculée. Ensuite, une règle de vote est utilisée pour évaluer la fiabilité des sources. Nous utilisons une méthode itérative pour calculer ces deux étapes.

À chaque itération, la confiance des faits (*evaluate-facts*(\mathcal{F}) – voir section 4.1) est évaluée et un classement des faits de chaque objet est obtenu. La méthode de vote est ensuite utilisée pour récompenser les faits par rapport au classement (*vote*(\mathcal{F}) – voir section 4.2). La fiabilité des sources est déduite des récompenses des faits qu'elles affirment (*evaluate-sources*(\mathcal{S}) – voir l'équation 2) et est normalisée (*normalize*(\mathcal{S}) – voir la section 4.3).

L'algorithme s'arrête lorsque le processus converge, c'est-à-dire lorsque la distance euclidienne entre la fiabilité des sources de la dernière itération et celle de l'itération actuelle est inférieure à ϵ avec $\epsilon = 0,001$.

Une méthode de cette famille est caractérisée par le choix d'une règle de vote et d'une fonction de normalisation.

4.1 Confiance des faits

La confiance dans un fait f est simplement calculée en additionnant la fiabilité des sources qui l'ont affirmé. Plus les sources qui l'affirment sont fiables, plus la confiance en ce fait est grande.

$$c(f) = \sum_{s \in src(f)} r(s) \quad (1)$$

1. Nous notons simplement $r(s)$ et $c(f)$, sans l'indice, lorsqu'il n'y a aucune ambiguïté sur le graphe utilisé.

Algorithm 1 S&F Algorithm**Input** : Un graphe $G = (\mathcal{S} \cup \mathcal{F} \cup \mathcal{O}, E)$.**Output** : La fiabilité des sources $r(s)$ et la fiabilité des faits $c(f)$.

- 1: $r(s) = 1 \forall s \in \mathcal{S}$ #fiabilité initiale des sources
 - 2: $c(f) = 0 \forall f \in \mathcal{F}$ #fiabilité initiale des faits
ts (resp. ts^{-1}) est le vecteur de fiabilité des sources pendant l'itération actuelle (resp. dernière) t.q. $ts = \langle r(s) : \forall s \in \mathcal{S} \rangle$.
 - 3: **while** not convergence(ts, ts^{-1}) **do**
 - 4: evaluate-facts(\mathcal{F})
 - 5: vote(\mathcal{F})
 - 6: evaluate-sources(\mathcal{S})
 - 7: normalize(\mathcal{S})
 - 8: **end while**
-

4.2 Fiabilité des sources

L'évaluation de la fiabilité des sources est obtenue par un vote des objets. Chaque objet récompense les sources qui affirment les faits les plus plausibles. Nous utiliserons des règles de vote par scores pour calculer cette fiabilité.

Pour chaque objet, nous classons les faits correspondants du plus fiable au moins fiable, et la règle de vote par scores associe un score à chaque rang.

La nouvelle fiabilité des sources est alors la somme de ces scores, comme il est d'usage dans les règles de vote par scores.

Mais nous devons procéder à trois ajustements : Le premier est que les règles de vote par scores sont définies pour des ordres linéaires, alors que nous obtenons des pré-ordres totaux : certains faits peuvent avoir le même rang. Nous utiliserons donc la moyenne des scores des faits avec le même classement. Le second est que le nombre d'options (faits) n'est pas le même pour tous les objets, de sorte que nous devons choisir comment normaliser ces scores sur différentes échelles. Le troisième est que, lorsque les scores sont reçus par les sources, nous les normalisons afin d'avoir un résultat compris dans $[0, 1]$.

Définition 2 Soit M un entier et e une suite d'entiers non décroissants avec $e_1 \geq e_2 \geq \dots \geq e_M$ t.q. $e_1 > e_M$. Une règle de vote par scores v est une fonction qui, à chaque ordre linéaire $>$ sur un ensemble d'au plus M faits et à chaque fait f , associe un entier positif t.q. si le fait f est classé à la i ème position dans l'ordre linéaire $>$, alors $v(>, f) = e_i$.

Lorsque $e_1 = 1$ et $e_2 = e_M = 0$, la règle est appelée majorité simple. Lorsque $e_1 = M - 1, e_2 = M - 2, \dots, e_M = 0$, il s'agit de la règle de Borda.

Pour les procédures de vote standard, les électeurs votent sur un ensemble fixe de candidats, et M est le nombre de candidats. Dans notre cas, les objets sont liés à différents nombres de faits. Nous devons donc procéder à une

(première) normalisation en fonction du nombre maximal de faits. Nous définissons donc $M = \text{best_score}(\mathcal{F}) = \max(|\text{fct}(o)|), \forall o \in \mathcal{O}$. Cela signifie que pour tous les objets, le fait d'être le plus plausible donne toujours le même score (e_1), quel que soit le nombre de faits liés à l'objet.

De plus, contrairement aux hypothèses classiques des règles de vote standard, le classement associé à un objet n'est pas un ordre linéaire, mais un pré-ordre total (certains faits peuvent avoir la même confiance). Nous devons ajuster les règles de vote par scores pour pouvoir gérer de manière adéquate les égalités possibles. Dans ce cas, nous donnons la note moyenne, c'est-à-dire la moyenne des notes qu'ils étaient censés recevoir, comme dans [25].

Un pré-ordre total (une relation réflexive, transitive et totale)² peut être considéré comme un ensemble de strates. Un élément x appartient à une strate T_{\geq} composée d'un ensemble d'éléments équivalents $\{y | x \approx y\}$. T_{\geq}^i est la i^{th} strate du pré-ordre si $\exists x_1, \dots, x_{i-1}$ t.q. $x_1 > \dots > x_{i-1} > y$ avec $y \in T_{\geq}^i$. S'il n'existe pas de $x_1 > y$ avec $y \in T_{\geq}^i$ alors $i = 1$.

Un ordre linéaire $>$ est dit compatible avec un pré-ordre \geq si $\forall x \in T^i, \forall y \in T^j, i < j \Rightarrow x > y$.

Définition 3 Pour chaque fait f , considérons l'objet correspondant o . Soit $\mathcal{P}(o)$ le pré-ordre donné par la confiance des faits (c'est-à-dire $f_1 \geq_{\mathcal{P}(o)} f_2$ ssi $c(f_1) \geq c(f_2)$) et m le nombre de strates dans $\mathcal{P}(o)$. On a $\mathcal{P}(o) = \{T^1(o), T^2(o), \dots, T^m(o)\}$, où $T^k(o)$ est la k^{th} strate dans $\mathcal{P}(o)$. Le score attribué à f pour le pré-ordre $\mathcal{P}(o)$ et la règle de pondération v est défini comme suit (où $>_o$ est tout ordre linéaire compatible avec $\mathcal{P}(o)$) :

$$V_v(f) = \frac{\sum_{\{y \in T^i(o) | f \in T^i(o)\}} v(>_o, y)}{|T^i(o)|}$$

Exemple 1 Soit G un graphe avec deux objets. Soit v la règle de Borda. Le premier objet $o1$ et le second objet $o2$ ont respectivement 6 et 9 faits qui leur sont liés. On a $\text{best_score}(\mathcal{F}) = 9$. Supposons que $\mathcal{P}(o1) = \{T^1(o1) = \{f_1, f_2\}, T^2(o1) = \{f_3\}, T^3(o1) = \{f_4, f_5, f_6\}\}$, c'est-à-dire que $o1$ a classé 2 faits en premier (dans la première strate), 1 fait en deuxième et 3 faits en troisième. Le score de f_1 et f_2 est donc $V_{\text{Borda}}(f_i) = \frac{(9-1)+(9-2)}{2} = 7.5$, le score de f_3 est $V_{\text{Borda}}(f_3) = (9-3) = 6$ et le score de f_4, f_5 et f_6 est $V_{\text{Borda}}(f_i) = \frac{(9-4)+(9-5)+(9-6)}{3} = 4$.

Et, comme d'habitude pour les règles de vote par scores, le score d'une source est simplement la somme des scores donnés par chaque votant (objet) :

Définition 4 La fiabilité initiale d'une source (avant normalisation) est :

2. À partir de tout pré-ordre total \geq , nous définissons l'ordre strict correspondant $>$ comme $x > y$ ssi $x \geq y$ et $y \not\approx x$, et la relation d'équivalence correspondante \approx comme $x \approx y$ ssi $x \geq y$ et $y \approx x$.

$$r^I(s) = \sum_{f \in \text{fct}(s)} V_v(f) \quad (2)$$

Voyons maintenant le dernier ajustement nécessaire des scores.

4.3 Normalisations A and C

Nous souhaitons donner une estimation de la fiabilité d'une source, c'est-à-dire de la probabilité qu'étant donné un objet, cette source trouve le fait réel correspondant. Nous devons donc normaliser la fiabilité des sources pour nous assurer que cette fiabilité soit comprise entre 0 et 1. Il existe (au moins) deux façons raisonnables de normaliser la fiabilité. La première favorise les sources qui fournissent beaucoup d'informations plausibles. La seconde se concentre sur la qualité et ensuite sur la proportion des informations fournies par la source.

Nous appelons la première normalisation A (pour *All objects*). La fiabilité des sources est divisée par le nombre d'objets dans le graphe. Si une source a un score proche de 1, nous savons que la source s'exprime correctement pour presque tous les objets. Si la fiabilité d'une source est faible, cela signifie soit que la source commet beaucoup d'erreurs et perd des votes, soit que la source ne s'exprime pas beaucoup.

Définition 5 La fiabilité d'une source après la normalisation A est :

$$r^A(s) = \frac{r^I(s)}{\text{best_score}(\mathcal{F}) * |\mathcal{O}|} \quad (3)$$

La seconde normalisation est appelée C (pour *Claimed facts*). La fiabilité d'une source est divisée par le nombre d'objets sur lesquels elle affirme un fait. Contrairement à la normalisation précédente, si une source a un score proche de 1, nous savons que la source s'exprime correctement mais nous n'avons pas idée du nombre d'objets sur laquelle elle s'exprime.

Définition 6 La fiabilité d'une source après la normalisation C est :

$$r^C(s) = \frac{r^I(s)}{\text{best_score}(\mathcal{F}) * |\text{obj}(s)|} \quad (4)$$

Où $\text{obj}(s) = \{o \in \mathcal{O} : \exists f \in \mathcal{F} : (s, f), (f, o) \in E\}$.

Cette normalisation favorise donc les sources qui s'expriment correctement, tandis que la précédente favorise les sources qui s'expriment beaucoup (et de manière correcte).

Notons que le score le plus élevé qu'une source puisse obtenir est $\text{best_score}(\mathcal{F})$, il faut donc multiplier le dénominateur par cette valeur. Dans le cas d'un graphe complet, les deux normalisations sont identiques. On note $r(s)$ la fiabilité normalisée d'une source lorsqu'il n'y a pas d'ambiguïté sur la normalisation utilisée.

4.4 Exemple

Nous allons maintenant voir un exemple afin de montrer les différences entre les deux normalisations pour la majorité simple.

L'impact de la normalisation est illustré dans la Table 1 et dans la Table 2 : la source 1 est l'une des meilleures sources avec la normalisation C mais l'une des pires avec la normalisation A. De même, le fait a_2 devient l'un des meilleurs sur l'objet A avec la normalisation C, alors qu'il est l'un des pires avec la normalisation A. En changeant la normalisation, nous avons des changements significatifs dans les résultats de l'algorithme, comme la source 1 et le fait a_2 dans l'exemple. Les itérations de l'algorithme sont détaillées pour la méthode avec la majorité simple et la normalisation A dans la Table 1 et dans la Table 2 pour la normalisation C.

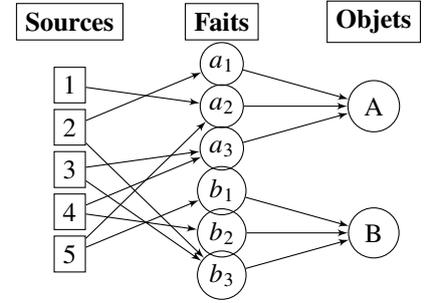


FIGURE 2 – Graphe

sources		1	2	3	4	5	/
It1	fiabilité	1	1	1	1	1	/
It2	fiabilité	0.5	0.5	1	0.5	0.5	/
It3	fiabilité	0	0.5	1	0.5	0	/
faits		a1	a2	a3	b1	b2	b3
It1	confiance	1	2	2	1	1	2
	$V_f(v)$	0	1	1	0	0	1
It2	confiance	0.5	1	1.5	0.5	0.5	1.5
	$V_f(v)$	0	0	1	0	0	1
It3	confiance	0.5	0	1.5	0	0.5	1.5
	$V_f(v)$	0	0	1	0	0	1

TABLE 1 – Itérations avec majorité simple A

sources		1	2	3	4	5	/
It1	fiabilité	1	1	1	1	1	/
It2	fiabilité	1	0.5	1	0.5	0.5	/
faits		a1	a2	a3	b1	b2	b3
It1	confiance	1	2	2	1	1	2
	$V_f(v)$	0	1	1	0	0	1
It2	confiance	0.5	1.5	1.5	0.5	0.5	1.5
	$V_f(v)$	0	1	1	0	0	1

TABLE 2 – Itérations avec majorité simple C

Si nous entrons dans les détails, avec la normalisation A,

nous pouvons voir à la première itération que la confiance des faits a_2 et a_3 est la même, puisqu'ils sont tous deux affirmés par deux sources. Mais la fiabilité de ces sources est ajustée à la fin de cette itération, et pour la deuxième itération, a_3 devient plus crédible. Notez qu'à la troisième itération, a_2 devient encore moins crédible que a_1 puisqu'il n'est affirmé que par des sources non fiables à cette itération.

Notons également que la méthode avec la normalisation C s'arrête en deux itérations mais que la normalisation A a besoin d'une itération supplémentaire pour s'arrêter.

5 Propriétés

Cette section comporte deux volets. Tout d'abord, nous souhaitons abstraire le problème et nous demander quelles propriétés devraient être satisfaites par les méthodes qui visent à évaluer la fiabilité des sources et la confiance dans les faits. Nous rappelons certaines propriétés proposées dans [28, 27] et les discutons, et nous en proposons de nouvelles. En particulier, nous proposons un ensemble de propriétés (les propriétés de base) que toute méthode devrait satisfaire, ainsi que d'autres propriétés intéressantes pour caractériser des sous-classes intéressantes de méthodes. Le second objectif de cette section est d'évaluer nos méthodes par rapport à cet ensemble de propriétés.

Commençons par donner quelques définitions utilisées par les propriétés.

Définition 7 Nous notons $B(o)$ l'ensemble des faits classés premiers pour l'objet o : $B(o) = \{f \in \text{fct}(o) \mid \forall f' \in \text{fct}(o), c(f) \geq c(f')\}$.

Définition 8 Nous notons $B(\mathcal{F})$ l'ensemble des faits classés premiers pour un objet du graphe : $B(\mathcal{F}) = \{f \in \mathcal{F} \mid \exists o \in \mathcal{O}, f \in B(o)\}$

Nous avons besoin d'une définition venant de [28] pour la notion de "est moins plausible que" utilisée dans leurs propriétés.

Définition 9 Soit $Y, Y' \subseteq \mathcal{F}$, Y est moins plausible que Y' s'il y a une bijection $\phi : Y \rightarrow Y'$ t.q. $c(f) \leq c(\phi(f))$ pour chaque $f \in Y$ et $c(f') < c(\phi(f'))$ pour au moins un fait $f' \in Y$. Soit $X, X' \subseteq \mathcal{S}$, X est moins fiable que X' est défini de façon similaire.

5.1 Propriétés de base

Nous présentons les propriétés que toutes méthodes qui souhaitent estimer correctement la fiabilité des sources et trouver la vérité parmi les faits doivent satisfaire.

Si la fiabilité d'une source est égale à 1 (le score le plus élevé pour une source), cela signifie que tous ses faits sont les plus plausibles, c'est-à-dire qu'ils ont la plus grande confiance sur leur objet :

P1 (Best) Soit $s \in \mathcal{S}$, si $r(s) = 1$ alors $\text{fct}(s) \subseteq B(\mathcal{F})$.

Si une source n'affirme aucun fait, sa fiabilité doit être la plus faible :

P2 (Null Player) Soit $s \in \mathcal{S}$, si $\text{fct}(s) = \emptyset$ alors $r(s) = 0$.

Nous rappelons quatre propriétés et une définition de [28].

Si un fait n'est affirmé par aucune source, sa confiance est inférieure ou égale à celle de tous les autres faits :

P3 (Groundedness) Nous supposons que $\text{src}(f) = \emptyset$ pour $f \in \mathcal{F}$. Alors pour tous les autres $g \in \mathcal{F}$, $c(f) \leq c(g)$.

Si un fait est affirmé par toutes les sources, sa confiance sera la plus élevée :

P4 (Unanimity) Nous supposons que $\text{src}(f) = \mathcal{S}$ pour $f \in \mathcal{F}$. Alors pour tous les autres $g \in \mathcal{F}$, $c(f) \geq c(g)$.

Définition 10 Deux graphes G et G' sont équivalents s'il existe un graphe isomorphe π entre eux qui préserve les sources, faits et objets t.q. $\pi(s) \in \mathcal{S}'$, $\pi(f) \in \mathcal{F}'$ et $\pi(o) \in \mathcal{O}'$ pour tous les $s \in \mathcal{S}$, $f \in \mathcal{F}$ et $o \in \mathcal{O}$.

Les valeurs calculées pour la fiabilité des sources et la confiance des faits dépendent uniquement du graphe et non de leur nom. Cette propriété indique donc que toutes les sources et tous les faits sont traités de la même manière (le nom original de cette propriété est "Symmetry" dans [28]) :

P5 (Neutrality) Si G et $G' = \pi(G)$ sont des graphes équivalents, alors $(r_G(s1) \geq r_G(s2) \text{ ssi } r_{G'}(\pi(s1)) \geq r_{G'}(\pi(s2)))$ et $(c_G(f1) \geq c_G(f2) \text{ ssi } c_{G'}(\pi(f1)) \geq c_{G'}(\pi(f2)))$.

Le classement des éléments d'une composante connexe n'est pas influencé par les éléments extérieurs à la composante (le nom original de cette propriété est "PCI" dans [28]) :

Définition 11 Soit $G = (V, E)$ et $G' = (V', E')$ deux graphes. On dit que G et G' sont indépendants lorsqu'il n'y a aucun liens qui relient les éléments du graphe G et les éléments du graphe G' , i.e. $V \cap V' = \emptyset$.

P6 (Independence) Soit $G = (V, E)$, $G1 = (V1, E1)$, $G2 = (V2, E2)$ trois graphes t.q. G et Gi ($i \in \{1, 2\}$) sont des graphes indépendants. Alors le classement des sources et des faits du graphe G doit être le même pour $G \cup G1$ et pour $G \cup G2$: $\forall s1, s2 \in \mathcal{S}_G$ nous avons $r_{G \cup G1}(s1) \geq r_{G \cup G1}(s2) \text{ ssi } r_{G \cup G2}(s1) \geq r_{G \cup G2}(s2)$. Et $\forall f1, f2 \in \mathcal{F}_G$ nous avons $c_{G \cup G1}(f1) \geq c_{G \cup G1}(f2) \text{ ssi } c_{G \cup G2}(f1) \geq c_{G \cup G2}(f2)$.

Les faits affirmés par des sources moins fiables sont forcément moins crédibles.

P7 (Fact Coherence) Si $\text{src}(f1)$ est moins fiable que $\text{src}(f2)$ alors $c(f1) < c(f2)$.

Certaines propriétés supplémentaires semblent également souhaitables.

Définition 12 Soit $G = (V, E)$ un graphe. Nous notons $\text{dupS}(G, s, n)$ le graphe avec lequel l'on copie la source s ainsi que tous ses liens n fois. $\text{dupS}(G, s, n) = (V', E')$ où $V' = V \cup \{s_1, s_2, \dots, s_n\}$ et $E' = E \cup \{(s_i, f) | f \in \text{fct}_G(s), i = 1, \dots, n\}$.

Si une opinion est suffisamment populaire, elle doit être considérée comme la vérité. Ainsi, si une source est dupliquée suffisamment de fois, ses faits doivent devenir les plus plausibles.

P8 (Majority) Soit $G = (V, E)$ un graphe et $s \in \mathcal{S}_G$. $\exists n > 0$ t.q. $\text{fct}_{G'}(s) \subseteq B_{G'}(\mathcal{F})$ avec $G' = \text{dupS}(G, s, n)$.

Nous allons maintenant voir un cas particulier où un graphe n'a qu'un seul objet.

Définition 13 Notons \mathcal{O}_1 le graphe $G = (\mathcal{S} \cup \mathcal{F} \cup \mathcal{O}, E)$ avec un seul objet (t.q. $|\mathcal{O}| = 1$).

Lorsqu'il n'y a qu'un seul objet dans un graphe, un fait qui est affirmé par un plus grand nombre de sources aura une meilleure confiance qu'un autre fait étant moins affirmé. Cette propriété est importante, car elle indique que la force de base d'un fait est donnée par le nombre d'affirmations. Mais avec plus d'un objet, les informations recueillies sur d'autres objets peuvent être utilisées pour prendre de meilleures décisions. Cette propriété n'est donc pas souhaitable pour plus d'un objet, puisque dans ce cas, nous voulons prendre en compte à la fois le nombre d'affirmations et la performance des sources sur d'autres objets.

P9 (Claims) Si $c_{\mathcal{O}_1}(f) > c_{\mathcal{O}_1}(f')$ alors $|\text{src}_{\mathcal{O}_1}(f)| > |\text{src}_{\mathcal{O}_1}(f')|$

5.2 Propriétés supplémentaires

Les propriétés présentées dans la section précédente sont celles que toute méthode doit satisfaire. Dans cette section, nous donnons des propriétés supplémentaires, qui ne sont pas nécessaires pour toutes les méthodes, mais qui caractérisent des comportements intéressants de certaines méthodes.

Les deux propriétés suivantes sont liées à la propriété (Best). Une source doit (correctement) affirmer tous les faits si elle veut obtenir le score le plus élevé :

P10 (Best A) Soit $s \in \mathcal{S}$, $r(s) = 1$ ssi $\text{fct}(s) = B(\mathcal{F})$.

Une alternative consiste à considérer qu'une source est la plus fiable (fiabilité égale à 1) si elle trouve toujours le fait le plus plausible (sans avoir à s'exprimer sur tous les objets) :

P11 (Best C) Soit $s \in \mathcal{S}$, $r(s) = 1$ ssi $\text{fct}(s) \subseteq B(\mathcal{F})$.

Notons que (Best A) et (Best C) impliquent tous deux la propriété (Best).

Si la fiabilité d'une source est égale à 0 (le score le plus bas), cela signifie qu'aucun de ses faits n'est plausible (sur l'objet correspondant) :

P12 (Worst) Soit $s \in \mathcal{S}$, $r(s) = 0$ ssi $\text{fct}(s) \subseteq \mathcal{F} \setminus B(\mathcal{F})$.

Si une source affirme des faits plus crédibles qu'une autre source, la fiabilité de la première source sera meilleure :

P13 (Source Dominance) Soit deux sources s et s' , si $|B(\mathcal{F}) \cap \text{fct}(s)| > |B(\mathcal{F}) \cap \text{fct}(s')|$ alors $r(s) > r(s')$.

Lorsqu'une source s affirme un fait avec une confiance supérieure à celle d'une autre source s' pour chaque objet, alors la fiabilité de s doit être meilleure :

P14 (Pareto) Soit $G = (V, E)$ un graphe complet et $s, s' \in \mathcal{S}$. Si $c(f) > c(f')$ avec $f \neq f'$, $f \in \text{fct}(s)$, $f' \in \text{fct}(s')$ et $\text{obj}(f) \cap \text{obj}(f') = \{o\} \forall o \in \mathcal{O}$ alors $r(s) > r(s')$.

5.3 Propriétés discutables

Nous présentons dans cette section certaines propriétés de [28] que nous considérons comme discutables pour toutes les méthodes et nous expliquons pourquoi nous pensons qu'elles ne sont pas satisfaisantes.

La première propriété stipule que les sources qui affirment des faits plus crédibles doivent être plus fiables :

P15 (Source Coherence) Si $\text{fct}(s1)$ est moins plausible que $\text{fct}(s2)$ alors $r(s1) < r(s2)$.

Notons que la notion de "moins plausible" ne nécessite pas que les faits portent sur les mêmes objets. Le problème de cette propriété est que nous comparons des faits qui concernent (potentiellement) des objets différents, alors que l'évaluation des faits est faite pour chaque objet. Par exemple, deux faits peuvent avoir la même confiance, mais l'un est le plus plausible pour son objet, tandis que l'autre est le moins plausible pour un autre objet.

La deuxième propriété stipule que, lorsqu'un fait reçoit un nouveau soutien, son classement doit être strictement meilleur :

P16 (Monotonicity) Soit G un graphe, $s \in \mathcal{S}$, $f \in \mathcal{F} \setminus \text{fct}(s)$. Nous écrivons E pour les arêtes de G , et G' un graphe avec les arêtes $E' = \{(s, f)\} \cup E \setminus \{(s, g) : g \neq f, \text{obj}(g) = \text{obj}(f)\}$. Alors pour tout $g \neq f$, $c_G(g) \leq c_G(f)$ implique $c_{G'}(g) < c_{G'}(f)$.

Cette propriété ne tient pas compte du reste du graphe et des changements qui peuvent survenir lorsqu'une arête est modifiée. Cette propriété semble être associée à une vision locale du problème, où l'évaluation des faits correspondant à un objet est indépendante des autres objets. Mais

il est important de connaître les performances des sources sur d'autres objets afin de prendre une décision sur un objet donné, et la modification d'une arête sur un objet peut changer la crédibilité des faits sur d'autres objets et la fiabilité de nombreuses sources. Enfin, l'évaluation de l'objet sur lequel le changement a été effectué donnera un résultat différent.

Une autre propriété discutable stipule que la confiance des faits ne doit dépendre que de l'objet auquel ils sont liés. Notons que les auteurs [28] classent également cette propriété comme discutable :

P17 (POI) Soit G, G' deux graphes et $o \in O$. Nous supposons que $fct_G(o) = fct_{G'}(o)$ et $src_G(f) = src_{G'}(f)$ pour chaque $f \in fct_G(o)$. Alors $c_G(f1) \leq c_G(f2)$ ssi $c_{G'}(f1) \leq c_{G'}(f2)$ pour tous $f1, f2 \in fct_G(o)$.

Cette propriété pose un problème similaire à la précédente. Il est important d'évaluer les performances des sources sur d'autres objets afin de prendre une décision sur un objet donné, comme illustré dans l'exemple de l'introduction, où l'évaluation des performances sur *Capitale du Brésil* nous aide à prendre une décision sur *Capitale de l'Australie*.

5.4 Propriétés des méthodes S&F

Dans les tableaux et les figures, PIA et PIC correspondent respectivement à la méthode S&F avec la majorité simple et à la normalisation A ou C. BoA et BoC correspondent aux méthodes avec la règle de Borda et les normalisation A et C.

Vérifions quelles sont les propriétés satisfaites par nos méthodes. Nous nous concentrons sur les deux normalisations (C et A), ainsi que sur la majorité simple et la règle de Borda.

Proposition 1 PIA satisfait (P1-P9), (P10), (P12), (P13) et (P14). Elle ne satisfait pas (P11), (P15-P17).

Proposition 2 PIC satisfait (P1-P9), (P11), et (P12). Elle ne satisfait pas (P10), (P13), (P14) et (P15-P17).

Proposition 3 BoA satisfait (P1-P9), (P10) et (P14). Elle ne satisfait pas (P11-P13) et (P15-P17).

Proposition 4 BoC satisfait (P1-P9), (P11) et (P14). Elle ne satisfait pas (P10), (P12), (P13) et (P15-P17).

Les résultats sont résumés dans la Table 3³. Tout d'abord, il est important de noter que nos méthodes satisfont toutes les propriétés de base, c'est-à-dire les propriétés attendues pour toutes les méthodes. Il est intéressant de discuter des propriétés qui ne sont satisfaites que par certaines méthodes, afin d'illustrer la différence dans leurs comportements. Tout

	PIA	PIC	BordaA	BordaC	
P1 Best	✓	✓	✓	✓	Nécessaire
P2 Null Player	✓	✓	✓	✓	
P3 Groundedness	✓	✓	✓	✓	
P4 Unanimity	✓	✓	✓	✓	
P5 Neutrality	✓	✓	✓	✓	
P6 Independence	✓	✓	✓	✓	
P7 Fact Coherence	✓	✓	✓	✓	
P8 Majority	✓	✓	✓	✓	
P9 Claims	✓	✓	✓	✓	
P10 Best A	✓	✗	✓	✗	Optionnelle
P11 Best C	✗	✓	✗	✓	
P12 Worst	✓	✓	✗	✗	
P13 Source Dominance	✓	✗	✗	✗	
P14 Pareto	✗	✗	✓	✓	
P15 Source Coherence	✗	✗	✗	✗	
P16 Monotonicity	✗	✗	✗	✗	
P17 POI	✗	✗	✗	✗	Indésirable

TABLE 3 – Propriétés satisfaites par les méthodes S&F

d'abord, notons que *Best C* implique d'utiliser la normalisation C pour nos méthodes, alors que *Best A* correspond à la normalisation A. La propriété *Worst* correspond au comportement de la majorité simple, avec d'autres règles de pondération, elle ne sera pas satisfaite. Inversement, la propriété *Pareto* est liée à la règle de Borda, et n'est pas satisfaite par la majorité simple, qui effectue une évaluation plus drastique des faits. Enfin, *Source dominance* n'est satisfaite que par la majorité simple et la normalisation A.

6 Étude Expérimentale

Outre l'évaluation théorique de nos méthodes, nous avons également procédé à une évaluation expérimentale de leurs performances en matière d'identification des faits réels et d'évaluation de la fiabilité des sources. Nous avons mené des expériences à la fois sur des ensembles de données réelles et des ensembles de données synthétiques.

6.1 Données réelles

Nous évaluons nos méthodes sur deux ensembles de données provenant de <http://lunadong.com/fusionDataSets.htm>, à savoir l'ensemble de données *Book* [8, 33] et le jeu de données *Flight* [16, 8].

Nous abrégons TF pour Truth Finder ([33]), H&A pour Hubs and Authorities ([13]), Usums pour Unbounded-Sums ([28]) et Sums ([22]). P représente la métrique *Precision*, A pour *Accuracy*, R pour *Recall* et C pour *CSI* (Critical Success Index), voir [18, 7] pour plus de détails sur ces mesures.

Book. La difficulté avec cet ensemble de données a été de créer le graphe, car les données nécessitent un traitement de texte. Après le nettoyage des données, le graphe se compose

3. Les lignes grises de la Table 3 correspondent aux propriétés de [28].

de 876 sources, 5685 faits et 1263 objets. Le *ground truth* est composé de 100 objets avec un fait réel connu. Nous voyons dans la Table 4 que la méthode S&F avec la majorité simple et la normalisation A est la meilleure méthode avec cet ensemble de données.

	PIA	PIC	BoA	BoC	TF	H&A	Sums	Usums
P	78.00	76.00	71.00	76.00	72.00	74.00	74.00	72.00
A	90.98	90.16	88.11	90.16	88.52	89.34	89.34	88.52
R	78.00	76.00	71.00	76.00	72.00	74.00	74.00	72.00
C	63.93	61.29	55.04	61.29	56.25	58.73	58.73	56.25

TABLE 4 – Résultat pour le jeu de données *Book*

	PIA	PIC	BoA	BoC	TF	H&A	Sums	Usums
P	91.35	82.34	83.82	81.91	80.36	82.21	82.21	82.79
A	91.49	82.61	84.06	82.18	81.72	82.48	82.48	83.05
R	91.35	82.34	83.82	81.91	83.22	82.21	82.21	82.79
C	84.08	69.98	72.14	69.36	69.15	69.80	69.80	70.63

TABLE 5 – Résultat pour le jeu de données *Flight*

Flight. Pour nettoyer cet ensemble de données, nous avons mis toutes les dates et heures dans le même format. Nous avons supprimé le terminal de la porte car il n’apparaît que quelques fois. Après le nettoyage des données, le graphe se compose de 38 sources, 399 506 faits et 207 912 objets. Le *ground truth* est composé de 16 089 objets avec un fait vrai connu. Nous voyons dans la Table 5 que la méthode avec la majorité simple et la normalisation A est également la meilleure méthode avec cet ensemble de données. Notre méthode surpasse les autres méthodes parce qu’elle parvient à trouver la vérité sur les objets même lorsque la majorité des sources n’affirment pas le vrai fait.

Nous voyons donc sur ces deux ensembles de données réelles que notre méthode S&F surpasse toutes les méthodes existantes de la littérature pour trouver les faits réels pour toutes les mesures de performance (P, A, R, C).

6.2 Données synthétiques

Le nombre limité d’ensembles de données réelles disponibles ne nous permet pas d’évaluer les performances des méthodes dans de nombreuses situations différentes. Nous avons généré des ensembles de données synthétiques pour pouvoir effectuer cette évaluation plus précise.

Tous les graphes générés sont composés de 10 objets et de 4 faits par objet. Pour chaque objet, nous choisissons au hasard l’un des quatre faits comme étant la vérité pour cet objet. Ce sera notre *ground truth* pour évaluer nos méthodes avec les métriques.

Pour chaque source, nous choisissons aléatoirement un nombre d’objets entre 1 et $|O|$ sur lequel cette source affirmera un fait. Pour générer les liens entre les sources et les faits, nous attribuons à chaque source une probabilité p (entre 0.1 et 0.9) de choisir un vrai fait sur chaque objet. Les faux faits ont la probabilité $1 - p$, uniformément distribuée, d’être choisis. Les graphes générés peuvent ne pas être complets, c’est-à-dire que les sources peuvent ne pas affirmer un fait sur chaque objet. Après la génération,

nous connaissons la probabilité *a posteriori* de choisir un vrai fait pour toutes les sources. Cette valeur représente la fiabilité réelle de ces sources.

Dans les tests, nous classons les expériences en fonction de la fiabilité moyenne des sources. Nous pouvons voir ce qui se passe lorsque les sources sont globalement plus ou moins fiables. Dans les graphiques, une fiabilité moyenne de $x\%$ signifie qu’il y a $x\%$ de liens entre les sources et les vrais faits (et $(100 - x)\%$ de liens entre les sources et les faux faits). Chaque point sur les graphiques correspond à la moyenne obtenue avec la génération de 1000 graphes.

Nous comparons les résultats de nos méthodes aux méthodes de la littérature (Truth Finder, Hubs and Authorities, Sums, Unbounded sums) et Voting.⁴

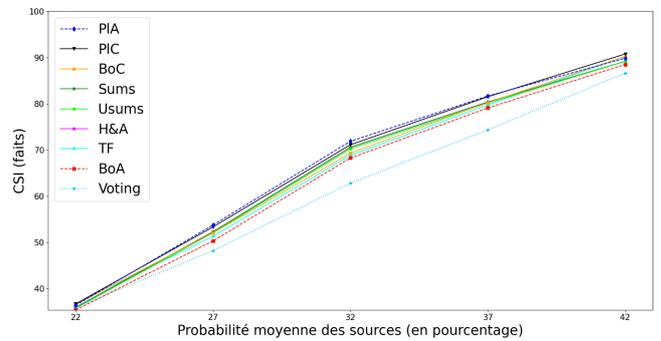


FIGURE 3 – CSI - 10 sources

Confiance des faits - Truth Discovery. Nous voyons (figure 3) que les méthodes S&F avec la majorité simple, et les deux normalisations, sont meilleures pour la métrique *CSI* que les autres méthodes de la littérature lorsque la fiabilité moyenne est supérieure à 27%. Par rapport aux autres méthodes, la méthode de la majorité simple trouve la vérité plus souvent lorsqu’il y a un nombre égal de sources affirmant le vrai et le faux fait. Elle trouve également la vérité lorsqu’une minorité de sources affirme le vrai fait. Il est intéressant de noter que les méthodes donnent de très bons résultats même lorsque la fiabilité moyenne est faible. Nos méthodes obtiennent de bons résultats avec les deux normalisations.

Toutes les méthodes trouvent la vérité lorsque la fiabilité moyenne est supérieure à 57%. Entre 42 et 57%, les résultats sont pratiquement les mêmes pour toutes les méthodes, c’est pourquoi nous ne montrons pas l’ensemble des résultats pour une meilleure lisibilité sur le graphique.

Fiabilité des sources.

Nous avons effectué des expériences avec plusieurs mesures (nombre de swaps, distance euclidienne, etc.), avec des résultats très convaincants, mais nous n’avons pas assez d’espace pour décrire toutes ces mesures, nous nous

4. Nous ne mettons ici que les figures pour *CSI*, mais nous obtenons des résultats identique pour *Precision*.

s	Probabilité	PIA	BoA	Voting	Sums	TF
s1	0.11	0.131	0.35	0.197	0.28	0.75
s2	0.17	0.187	0.39	0.249	0.35	0.77
s3	0.21	0.233	0.42	0.288	0.41	0.78
s4	0.27	0.296	0.46	0.341	0.47	0.81
s5	0.33	0.341	0.49	0.386	0.52	0.82
s6	0.39	0.403	0.53	0.438	0.58	0.85
s7	0.47	0.476	0.58	0.503	0.66	0.86
s8	0.53	0.528	0.61	0.549	0.70	0.88
s9	0.57	0.564	0.63	0.58	0.74	0.89
s10	0.61	0.591	0.64	0.603	0.75	0.90

TABLE 6 – Fiabilité des sources - Fiabilité moyenne de 37%

concentrerons donc sur la différence moyenne : nous calculons la différence moyenne entre la fiabilité calculée et la probabilité (*a posteriori*) de choisir le vrai fait pour chaque objet. Cette distance mesure donc à quel point la fiabilité estimée des sources est proche de la vraie fiabilité (la probabilité *a posteriori*).

Sur la Table 6, nous comparons la fiabilité estimée obtenue avec la fiabilité réelle (probabilité *a posteriori*), pour le cas où la fiabilité moyenne est de 37%. On peut voir que les estimations fournies par la majorité simple sont très proches de la probabilité réelle (rappelons que les résultats sont une moyenne sur 1000 expériences).

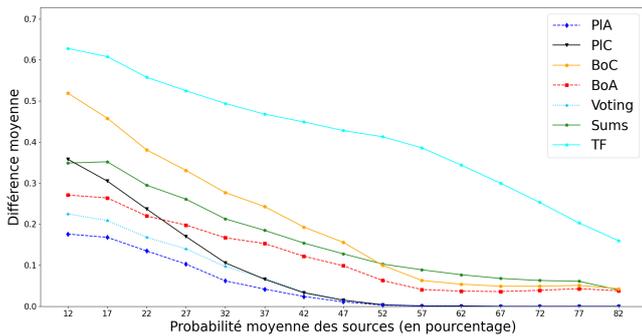


FIGURE 4 – Fiabilité des sources - Différence moyenne - 10 sources

Pour voir les résultats plus globalement, la figure 4 montre l'évolution de la différence moyenne avec différentes fiabilités moyennes pour les sources. La fiabilité estimée des sources est plus proche de la fiabilité réelle lorsque la fiabilité moyenne des sources augmente. Nous voyons que nous obtenons exactement la fiabilité réelle lorsque la fiabilité moyenne est meilleure que 57% pour la méthode utilisant la majorité simple. Avec la règle de Borda, on attribue des points à toutes les sources. C'est pourquoi la fiabilité n'est pas identique à la probabilité *a posteriori*, puisque les sources obtiendront également des points pour les fausses affirmations. Cependant, lorsque la fiabilité moyenne des sources augmente, la différence tend tout de même vers 0.

Nous comparons nos méthodes à celles de la littérature lorsque le score calculé en tant que *fiabilité des sources* est compris entre 0 et 1. Pour *Voting*, nous définissons la fiabilité d'une source comme la proportion d'objets pour

lesquels la source affirme le choix de la majorité. Nous ne comparons pas ici le résultat à Unbounded-Sums car le score augmente toujours pour cette méthode. *Voting* donne de bons résultats lorsque les sources sont fiables (fiabilité moyenne supérieure à 57%), mais avant cela, notre méthode avec la majorité simple est meilleure. La méthode itérative permet de trouver les faits réels même lorsque les sources ne sont pas vraiment fiables par rapport à l'utilisation d'une méthode de vote basique.

6.3 Convergence

Nous n'avons pas de preuve de convergence de nos méthodes, mais lors de nos tests, nous avons généré des millions de graphes avec différents paramètres et notre algorithme s'est toujours arrêté rapidement. La Figure 5 donne le nombre maximal d'itérations pour différentes fiabilités moyennes. Ainsi, le nombre maximal d'itérations que nous avons obtenues est de 14, et en moyenne la convergence est obtenue autour de 4 itérations.

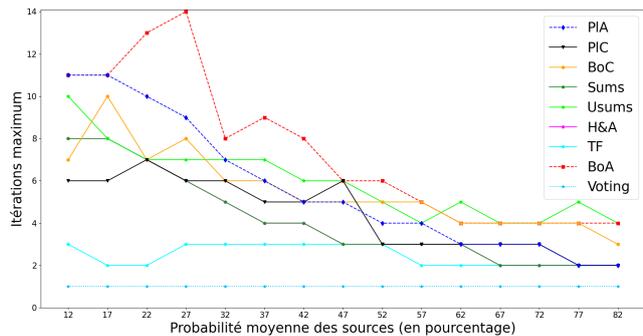


FIGURE 5 – Nombre d'itérations maximum - 10 sources

7 Conclusion

Dans cet article, nous avons présenté les méthodes S&F pour évaluer la fiabilité des sources conjointement à la confiance des faits dans un système multi-agents basé sur l'information. Nous avons proposé et discuté des propriétés que ces méthodes devraient ou pourraient satisfaire. Nous avons vérifié quelles propriétés sont satisfaites par nos méthodes. Nous avons également effectué quelques évaluations expérimentales. Tout d'abord, nous montrons que nos méthodes (en particulier avec la majorité simple) sont plus performantes que les méthodes de la littérature pour identifier les vrais faits sur des données réelles et générées. Mais nous montrons aussi que nos méthodes permettent d'estimer correctement la fiabilité des sources.

Il existe de nombreuses pistes pour les travaux à venir. Les plus directes sont de permettre une certaine similarité (ou dépendance) entre les objets, mais nous pourrions aussi utiliser des sujets (*topics*) différents pour nos objets. On peut aussi tenter de prendre en compte des informations a priori sur la fiabilité des sources.

Remerciements

Ce travail a bénéficié du support de la Chaire IA BE4musIA (ANR-20-CHIA-0028).

Références

- [1] Artz, Donovan et Yolanda Gil: *A survey of trust in computer science and the semantic web*. Journal of Web Semantics, 5(2) :58–71, 2007.
- [2] Austen-Smith, David et Jeffrey Banks: *Information Aggregation, Rationality, and the Condorcet Jury Theorem*. American Political Science Review, 90 :34–45, 1996.
- [3] Ben-Yashar, Ruth et Jacob Paroush: *A nonasymptotic Condorcet jury theorem*. Social Choice and Welfare, 17(2) :189–199, 2000.
- [4] Ben-Yashar, Ruth et Mor Zahavi: *The Condorcet jury theorem and extension of the franchise with rationally ignorant voters*. Public Choice, 148(3/4) :435–443, 2011.
- [5] Berend, Daniel et Jacob Paroush: *When is Condorcet’s Jury Theorem valid?* Social Choice and Welfare, 15(4) :481–488, 1998.
- [6] Condorcet, Marquis de: *Essai sur l’application de l’analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie royale Paris, 1785.
- [7] Donaldson, Ralph, Rosemary Dyer et Michael Kraus: *An objective evaluator of techniques for predicting severe weather events*. Dans *Preprints, Ninth Conf. on Severe Local Storms, Norman, OK, Amer. Meteor. Soc.*, tome 321326, 1975.
- [8] Dong, Xin Luna, Barna Saha et Divesh Srivastava: *Less is more : Selecting sources wisely for integration*. Proceedings of the VLDB Endowment, 6(2) :37–48, 2012.
- [9] Estlund, David: *Opinion leaders, independence, and Condorcet’s Jury Theorem*. Theory and Decision, 36(2) :131–162, 1994.
- [10] Everaere, Patricia, Sébastien Konieczny et Pierre Marquis: *The epistemic view of belief merging : can we track the truth?* Dans *Nineteenth European Conference on Artificial Intelligence (ECAI’10)*, pages 621–626, 2010.
- [11] Hammer, Stephan, Michael Wissner et Elisabeth André: *Trust-based decision-making for smart and adaptive environments*. User Modeling and User-Adapted Interaction, 25 :267 – 293, 2015.
- [12] Hummel, Patrick: *Jury theorems with multiple alternatives*. Social Choice and Welfare, 34(1) :65–103, 2010.
- [13] Kleinberg, Jon M: *Authoritative sources in a hyper-linked environment*. Journal of the ACM (JACM), 46(5) :604–632, 1999.
- [14] Ladha, Krishna K.: *The Condorcet Jury Theorem, Free Speech, and Correlated Votes*. American Journal of Political Science, 36(3) :617–634, 1992.
- [15] Ladha, Krishna K: *Information pooling through majority-rule voting : Condorcet’s jury theorem with correlated votes*. Journal of Economic Behavior & Organization, 26(3) :353 – 372, 1995.
- [16] Li, Xian, Xin Luna Dong, Kenneth Lyons, Weiyi Meng et Divesh Srivastava: *Truth finding on the deep web : Is the problem solved?* Proceedings of the VLDB Endowment, 6(2) :97–108, 2012.
- [17] List, Christian et Robert E. Goodin: *Epistemic Democracy : Generalizing the Condorcet Jury Theorem*. Journal of Political Philosophy, 9(3) :277–306, 2001.
- [18] Olson, David L et Dursun Delen: *Advanced data mining techniques*. Springer Science & Business Media, 2008, ISBN 978-3-540-76916-3.
- [19] Owen, Guillermo, Bernard Grofman et Scott L. Feld: *Proving a distribution-free generalization of the Condorcet Jury Theorem*. Mathematical Social Sciences, 17(1) :1 – 16, 1989.
- [20] Parhizkar, Elham, Mohammad Hossein Nikravan, Robert C. Holte et Sandra Zilles: *Combining Direct Trust and Indirect Trust in Multi-Agent Systems*. Dans *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, (IJCAI’20)*, pages 311–317, 2020.
- [21] Paroush, Jacob: *Stay away from fair coins : A Condorcet jury theorem*. Social Choice and Welfare, 15(1) :15–20, 1998.
- [22] Pasternack, Jeff et Dan Roth: *Knowing what to believe (when you already know something)*. Dans *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pages 877–885, 2010.
- [23] Peleg, Bezalel et Shmuel Zamir: *Extending the Condorcet Jury Theorem to a general dependent jury*. Social Choice and Welfare, 39(1) :91–125, 2012.
- [24] Pinyol, Isaac et Jordi Sabater-Mir: *Computational trust and reputation models for open multi-agent systems : a review*. Artificial Intelligence Review, 40 :1–25, 2013.
- [25] Rast, Erich: *Theory of Value Structure : From Values to Decisions*. Lexington Books, 2022, ISBN 9781793616951. <https://books.google.fr/books?id=BGRoEAAAQBAJ>.
- [26] Sabater, Jordi et Carles Sierra: *Review on computational trust and reputation models*. Artificial intelligence review, 24(1) :33–60, 2005.

- [27] Singleton, Joseph et Richard Booth: *An axiomatic approach to truth discovery*. Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '20, pages 2011–2013, 2020.
- [28] Singleton, Joseph et Richard Booth: *Towards an axiomatic approach to truth discovery*. Journal of Autonomous Agents and Multi-Agent Systems, 36(2) :1–49, 2022.
- [29] Siqueira Braga, Diego de, Marco Niemann, Bernd Hellgrath et Fernando Buarque de Lima-Neto: *Survey on Computational Trust and Reputation Models*. ACM Computing Surveys (CSUR), 51 :1 – 40, 2018.
- [30] Škorić, Boris, Sebastiaan de Hoogh et Nicola Zannone: *Flow-based reputation with uncertainty : evidence-based subjective logic*. International Journal of Information Security, 15 :381–402, 2014.
- [31] Vu, Quang Anh Nguyen, Salima Hassas, Frederic Armetta, Benoit Gaudou et Richard Canal: *Combining trust and self-organization for robust maintaining of information coherence in disturbed MAS*. Dans *IEEE Fifth International Conference on Self-Adaptive and Self-Organizing Systems(SASO'11)*, pages 178–187, 2011.
- [32] Waguih, Dalia Attia et Laure Berti-Equille: *Truth discovery algorithms : An experimental evaluation*. arXiv preprint arXiv :1409.6428, 2014.
- [33] Yin, Xiaoxin, Jiawei Han et Philip S. Yu: *Truth Discovery with Multiple Conflicting Information Providers on the Web*. IEEE Transactions on Knowledge and Data Engineering, 20(6) :796–808, 2008.
- [34] Young, Hobart Peyton: *Condorcet's Theory of Voting*. The American Political Science Review, 82(4) :1231–1244, 1988.
- [35] Young, Hobart Peyton et Arthur Levenglick: *A Consistent Extension of Condorcet's Election Principle*. SIAM Journal on Applied Mathematics, 35(2) :285–300, 1978.