



La Lutte contre la Fraude

- Outils Mathématiques -

1. Présentation

La lutte contre la fraude utilise des outils mathématiques depuis longtemps : comme le rapporte Vitruve, Archimède a détecté une fraude dans la composition d'une couronne pour le roi Hiéron de Syracuse ; voir le livre de Bernard Beauzamy "Archimedes' Modern Works" [AMW].

Aujourd'hui, la notion de fraude apparaît davantage de nature juridique : telle personne a ou n'a pas droit à tel document (passeport, carte Vitale, etc.), à telle prestation (chômage, retraite, etc.). On peut donc s'attendre à ce que toute vérification soit nécessairement individuelle et porte sur l'authenticité des documents présentés par la personne concernée. Ce sujet, en apparence, ne relève pas des mathématiques.

Mais il se trouve que des milliers d'organismes et des millions de personnes sont concernés : on ne peut pas vérifier à la main toutes les demandes une à une. Il est donc indispensable de disposer d'outils, essentiellement mathématiques, qui vont caractériser de manière simple, rapide et robuste, les "profils" susceptibles de représenter des fraudes ; ces profils seront alors l'objet de vérifications plus poussées. Ces outils ont ainsi un rôle préliminaire : ils permettent un "prétraitement" rapide. Ils doivent être les plus efficaces possible : détecter les fraudeurs, mais aussi ne pas désigner comme tels des gens qui n'en sont pas !

Les outils mathématiques mis en œuvre permettent deux sortes d'investigations :

- Cohérence interne des informations fournies par le demandeur

Il s'agit de vérifier que toutes les informations fournies sont cohérentes entre elles : telle date, qui doit être consécutive à telle autre, l'est effectivement ; les différents lieux mentionnés sont compatibles entre eux, etc.

- Cohérence des informations fournies par le demandeur avec les autres demandes

Il s'agit de vérifier qu'il n'y a pas des centaines de Jean Dupont, faisant la même demande, tous provenant du même village de l'Aveyron, ou bien tous avec la même consommation d'électricité, ou bien quantité de gens avec la même photo, etc.

La recherche de doublons (champs identiques) dans une base de données ne contrevient en rien à la Loi Informatique et Libertés. Il s'agit simplement de recenser les situations identiques, selon certains critères.

Les outils que nous venons de mentionner sont purement déterministes : ils fonctionnent par tri informatique. On peut en outre les compléter par des outils probabilistes, qui permettent de caractériser les "profils moyens" des demandeurs, en fonction de tous les critères que l'on souhaite inclure (âge, domicile, lieu de naissance, époque de l'année, etc.) On pourra alors détecter toute personne qui s'écartera trop du profil moyen, selon un seuil que l'on fixera. Voir notre livre [MPPR] pour les outils correspondants.

La mise en œuvre d'outils déterministes est assez simple ; les outils probabilistes sont plus délicats à manipuler, si l'on veut une action efficace. A l'intérieur d'un organisme donné, ils ne devraient pas être confiés aux statisticiens propres à l'organisme, pour deux raisons essentielles :

- Les statistiques usuelles de l'organisme incorporent la fraude, qui peut exister depuis très longtemps ;
- Les statisticiens "maison" voient cette détection comme une remise en cause de leur travail.

La lutte contre la fraude relève donc nécessairement de la Direction Générale, ou bien, s'il en existe un, du corps des Inspecteurs ou Contrôleurs. C'est une bonne illustration du vieil adage "on ne peut être à la fois juge et partie" : la vérification du bon fonctionnement des outils ne doit pas être laissée à ceux qui les ont conçus.

2. Références

[AMW] Bernard Beauzamy : Archimedes' Modern Works (en anglais), SCM SA, ISBN 978-2-9521458-7-9, ISSN 1767-1175, relié, 224 pages. Août 2012.

http://scmsa.eu/archives/SCM_AMW_order.htm

[MPPR] Bernard Beauzamy : Méthodes Probabilistes pour l'étude des phénomènes réels. SCM SA, ISBN 2-9521458-0-6. ISSN 1767-1175, broché, 369 pages. Mars 2004. Seconde Edition, juin 2016.

http://scmsa.eu/archives/SCM_MPPR_order.htm

3. Fiches associées

Qualité des données : détection de données aberrantes, reconstitution de données manquantes

http://scmsa.eu/fiches/SCM_Qualite_Information.pdf

Analyses de vulnérabilités et Contre-Mesures

http://scmsa.eu/fiches/SCM_vulnerabilites.pdf

4. Contrats récents

Ils concernent également la recherche de "vulnérabilités" dans les systèmes ; c'est une démarche logique : la fraude résulte d'une vulnérabilité que quelqu'un a découverte.

- Service des Programmes Navals, DGA, 2004-2009 : Le Bouclier Naval ; protection d'un navire de surface contre une attaque terroriste à courte portée. En cotraitance avec Thales Naval France, Thales Systèmes Aéroportés et TDA Armements, la SCM étant responsable de la faisabilité du système complet
- Laboratoire Pharmaceutique, 2006 : Détection des importations parallèles
- CDC-Arkhinéo, 2006 : Analyse des vulnérabilités d'un système d'archivage électronique
- Agence Nationale des Titres Sécurisés, 2008 et 2013 : Analyse des vulnérabilités dans le processus de délivrance d'un passeport biométrique
- Monceau Assurances, 2013-2014 : Mise en évidence de situations de fraudes par des méthodes probabilistes
- Monceau Assurances, 2016 et 2017 : Etude des "sursinistralités"