Loi sur le renseignement : Les coulisse d’un algorithme intrusif | Le Net Expert Informatique

Un chaton travaillant sur un algorithme (Mr Thinktank/Flickr/CC)	Loi sur le renseignement : Les coulisse d’un algorithme intrusif

On a demandé à des spécialistes en informatique s'il était possible de concevoir un programme répondant aux attentes du gouvernement en matière de renseignement. Résultat : techniquement, c'est très foireux.

Vous ne savez sans doute pas de quoi il s'agit. Pour être francs, nous non plus, nos élus non plus, et même nos contacts les plus calés en informatique nous répondent que ce domaine est trop pointu pour eux.
Pourtant, ce sujet est l'un des points les plus controversés du projet de loi sur le renseignement, discuté à l'Assemblée nationale depuis lundi : l'algorithme que le gouvernement, à la demande des services secrets, souhaite faire tourner au cœur de l'Internet français.

Programmé au sein des fameuses « boîtes noires » que l'exécutif veut installer sur les tuyaux des opérateurs (Orange, Free, Numericable…) et des hébergeurs (Google, Facebook, et autres), cet algorithme a pour but de détecter, avant même la commission d'actes terroristes, d'éventuels suspects.

On a déjà beaucoup parlé des similitudes entre cette ambition et la science-fiction. Mais concrètement, comment fonctionnera cet « algorithme » que tous les députés, tous les ministres, tous les conseillers, et donc tous les médias, ont à la bouche ces derniers jours ?
Nous sommes allés poser la question à des chercheurs en informatique, qui réfléchissent à la question de la vie privée, du stockage des données, ou bien encore à l'intelligence artificielle.
Croisées avec les rares explications fournies par le gouvernement (secret-défense oblige), les réflexions de nos interlocuteurs nous permettent d'affirmer qu'en l'état, quelle que soit la forme de l'algorithme choisie, le dispositif sera coûteux, intrusif et inefficace.

1. Un algorithme, c'est d'abord des humains
Ce n'est pas une formule magique, mais du code informatique créé par des êtres humains
« Quand les gens du gouvernement en parlent, on a l'impression qu'il s'agit d'une formule d'Harry Potter ! »
A l'instar de ce docteur en intelligence artificielle (IA), qui a souhaité garder l'anonymat, les personnes qui bossent quotidiennement sur des algorithmes sont aujourd'hui un peu désolées de voir leur outil de travail autant malmené.
Car si le terme revient souvent ces derniers temps, il est très rarement défini. Et devient l'objet d'une fascination béate, ou, à l'inverse, d'une peur irrationnelle.

Ni magique, ni diabolique : une recette de cuisine
Instant définition donc, histoire de lever le brouillard. Comme nous l'explique notre interlocuteur :
• « *Un algorithme, c'est simplement une suite d'opérations définies très strictement, que l'ordinateur, parfaitement stupide, exécute.* »
Un simple bout de code informatique (voir exemple ci-dessous), parfois court, parfois très long, qui vise à accomplir quelque chose. Pour cette raison, on compare souvent les algorithmes à une recette de cuisine : une série d'ingrédients précis qui aboutissent à un plat. L'analogie est plutôt bonne. Car si certaines recettes peuvent facilement être déduites de l'assiette posée devant soi (par exemple, un croque-monsieur), d'autres sont bien plus difficiles à cerner.

Un algorithme, ça ressemble en partie à cela. Extrait de Scikit-learn, qui donne des outils de data-mining (Scikit-learn)
Comme l'explique Gilles Dowek, chercheur à l'Institut national de recherche en informatique et en automatique (Inria), par e-mail :
• « Vous avez sans doute déjà mangé dans un restaurant un plat qui vous a plu que vous avez tenté de reproduire dans votre cuisine en essayant d'imaginer la recette qui y a conduit (pour ma part, j'essaie souvent avec un succès inégal). »

Avant, après : il y a des êtres humains
Autre corolaire de cette définition : si l'ordinateur exécute, c'est bien l'être humain qui définit ce qu'il doit exécuter. Et ce qu'il attend de cette opération. Notre spécialiste de l'IA explique :
• « En informatique, il y a toujours une entrée et une sortie. Au milieu, il y a une boîte, dans laquelle on entre une série d'opérations à faire, pour lesquelles on attend un résultat. »

Pour ce chercheur, les limites de l'opération sont déjà nettes :
• « La sortie attendue ici n'est pas très claire : il s'agit de dégager des comportements atypiques de la population qui seraient aussi typiques du terrorisme. »

Problème : comment définir des comportements atypiques ? Et typiques ? Le fait d'aller regarder une vidéo de décapitation de l'organisation Etat islamique est-il déjà un acte suspect ? On vous renvoie à la lecture de cet entretien très éclairant avec la chercheuse Antoinette Rouvroy.

Gilles Dowek pousse la démonstration un peu plus loin, en imaginant un système s'appuyant sur une liste de mots utilisés par des terroristes :
• « Que faire par exemple, si on s'aperçoit que cette liste contient le mot "banane" ? Cela signifie que statistiquement, les criminels utilisent fréquemment le mot "banane". Doit-on supprimer ce mot qui, manifestement, n'est pas suspect ? Ou alors considérer comme suspecte toute personne qui utilise ce mot ? »

Pour notre spécialiste de l'intelligence artificielle, on demande ici à un ordinateur une tâche bien trop fine : celle de catégoriser des êtres humains.
• « Or, ce système d'étude est hyper compliqué. Les ordinateurs n'ont qu'un modèle simplifié de l'humain. Par exemple, un humain vu par Amazon sera l'ensemble des bouquins qu'il a achetés sur un an. »

Or, à la différence d'Amazon et de tous les autres géants du Web, l'algorithme voulu par ce projet de loi ne recommandera pas des livres ou des sites internet, mais des humains.

Le gouvernement a beau jeu de dire que ces acteurs appliquent déjà, sur nous et avec notre accord, ces mêmes règles. Ce n'est pas tout à fait la même chose. Et par ailleurs, comme le note Isabelle Falque-Pierrotin, présidente de la Commission nationale de l'informatique et des libertés (Cnil), la puissance publique a d'autres prérogatives (et responsabilités) que les entreprises privées.

2. Quel que soit l'algorithme choisi, la surveillance est massive
Manuel Valls et Bernard Cazeneuve ont beau répéter que ce n'est pas la cas, la technique les fait mentir.
Passées ces précisions, entrons dans le dur : comment les services vont-ils s'y prendre avec cet algorithme ? Ou plus précisément, avec ces algorithmes qui, enchaînés entre eux, aboutiront au résultat voulu ?
« Il y a des tas de façons de faire », nous rétorquent l'ensemble des chercheurs que nous avons interrogés. Néanmoins, trois options se dégagent nettement. Et chacune, pour être efficace, nécessite de surveiller tout le monde.

La méthode déjà possible : les relations sociales
L'une des options possibles est de s'appuyer sur un objet mathématique bien connu, le graphe. Concrètement, il s'agirait de regarder les relations des « 3 000 personnes engagées de près ou de loin dans la mouvance terroriste », dont parlait le Premier ministre lundi.
De cette façon, on établit rapidement une cartographie de potentiels terroristes. En présumant, bien sûr, que le simple fait d'être en relation avec un terroriste identifié suffit à faire un suspect. Or, même en étant un véritable proche d'un terroriste, on peut ne rien savoir de ses activités. Il suffit de prendre le cas d'Amedy Coulibaly, en rupture avec certaines de ses sœurs.
Avec cet outil, la surveillance est massive. Peut-être connaissez vous la théorie des six degrés de séparation, selon laquelle chaque personne sur Terre est à six relations d'une autre ? Avec Internet, ce chiffre serait passé à quatre. Et la NSA, par exemple, va étudier jusqu'à deux degrés de séparation… Et espionne donc, grosso modo, des centaines de millions d'individus.
Néanmoins, on voit mal pourquoi les services français auraient attendu ce texte pour mettre en place ce dispositif, il suffit qu'il l'appliquent sur les 3 000 personnes déjà connues. Et il ne leur permettra pas de repérer les futurs éventuels coupables, ce qui est le but de cette loi.

La méthode basique : définir un ensemble de règles
Une autre option serait de demander à l'ordinateur de signaler les internautes qui font un ensemble de choses sur Internet considérées comme suspectes. Se connecter à tel et tel site djihadiste, utiliser régulièrement sur Internet une série de mots (par exemple « bombe », « arme », « explosion », « tirer », « tuer »), vérifier qu'une vidéo de propagande a bien été mise en ligne.
Ce scénario est l'un des plus crédibles. Il correspond aux rares exemples fournis par les ministres et leurs conseillers. Et a le caractère « rustique » que dépeignait, à l'occasion d'une conférence le 9 avril sur le sujet, l'un des conseillers de Manuel Valls, Renaud Vedel.
Néanmoins, là encore, pour être efficace, cette méthode nécessite de scruter l'ensemble des communications internet pour repérer ce que l'on recherche.

La méthode plus fine : un algorithme qui apprend
Ici, l'ordinateur apprend d'un jeu de données — on parle d'apprentissage automatique ou statistique.
Concrètement, la méthode consisterait, pour les services, à soumettre à la machine les habitudes de navigation des 3 000 personnes qui sont aujourd'hui dans leur radar. A partir de ces données, l'ordinateur détecte des particularités (des motifs), qui lui permettront par la suite de dire si oui ou non, tel ou tel internaute correspond à un profil de suspect.

Deux chercheurs de l'Inria, là encore sous couvert d'anonymat, nous alertent alors sur un point précis :
• « Pour détecter les motifs, l'algorithme a besoin des données d'un ensemble d'individus ayant le profil recherché que l'algorithme analyse (apprentissage) au regard des données d'individus ne correspondant pas au profil (pris au hasard). »

Traduction : là encore, pour que la méthode fonctionne, il faut surveiller non seulement des gens dont on ne sait pas s'ils sont suspects, mais dont on est certain qu'ils ne le sont pas. Non seulement cela confine à l'absurde, mais signifie que tout le monde peut être surveillé.

3. Quel que soit l'algorithme choisi, il sera inefficace
Faux positifs, faible nombre de suspects, limites du programme…

Une quantité astronomique de faux positifs
Les chercheurs sont également unanimes sur ce point : même si l'algorithme concocté par les services est hyper-balèze, il ne pourra échapper à une quantité considérable de faux positifs (en l'occurrence, des gens identifiés comme potentiellement suspects et qui se révèlent non coupables).

Pire, comme le dit notre spécialiste de l'intelligence artificielle :
• « Même avec un système d'une performance extrêmement élevée, il y aura toujours beaucoup plus d'innocents que de coupables accusés. »

Nos interlocuteurs de l'Inria confirment. Et déploient une démonstration implacable :
• « Supposons un algorithme d'une super-qualité qui n'a qu'une chance sur 100 de se tromper. Sur 60 millions de personnes, ça fait 600 000 personnes détectées à tort, plus les 1 000 "vrais positifs" qu'on a bien détectés. Donc l'algorithme détecte 601 000 personnes, parmi lesquelles en réalité 1 000 seulement sont de vrais terroristes. L'algorithme détecte alors les terroristes avec une probabilité de 1 000/601 000, soit 1/600, soit 0,02%… Tout ça pour ça ?

Par e-mail, Marc Schoenauer, directeur de recherche à l'Inria, évoque par exemple la possibilité de prendre dans ces filets "les créatifs d'avant-garde" (les gens qui ont des comportements "anormaux"). On peut aussi penser à tous ceux qui vont voir les vidéos de l'Etat islamique. Aux journalistes, aux chercheurs qui travaillent sur ces sujets.

Le problème des signaux faibles, c'est qu'on ne les voit pas
A cause des faux positifs, et parce que rapporté à des dizaines de millions de personnes, les terroristes sont très rares, les algorithmes auront du mal à les détecter.
Là encore, les chercheurs sont formels : cela revient à chercher une goutte dans l'océan que vous ciblez. A étendre la taille de la meule de foin dans laquelle vous recherchez l'aiguille, pour reprendre une expression de Pierre Lellouche, élu UMP opposé au texte.
Or, si des techniques permettent de repérer ce genre de signaux sur Internet, les chercheurs estiment que ces derniers ne sont pas assez fiables en l'espèce. Encore loupé.

La solution du gouvernement : encore moins efficace
Quand on leur demande, les conseillers du gouvernement sont formels : l'algorithme en question n'adaptera pas seul ses paramètres. C'est en effet une possibilité technique : certains algorithmes, à partir des données de départ, évoluent, apprennent en fonction des nouveaux usages observés.
Or, selon l'exécutif, chaque modification du code source de l'algorithme sera soumis au contrôle de la commission prévue à cet effet, la CNCTR. Ce qui est très inquiétant en termes de garanties pour les citoyens (on voit mal en effet comment la commission pourrait contrôler effectivement un algorithme qui change sans cesse), mais qui rend le dispositif bien précaire.

Si le but est de détecter de nouveaux terroristes, et qu'il faut modifier, à la main, le code de l'algorithme à chaque fois qu'une nouvelle pratique propre aux mouvements terroristes est détectée sur Internet, on voit mal comment on pourra les identifier à l'avance.

Ou, comme le résume notre expert en intelligence artificielle :
• "J'ai la crainte que quelque chose comme ça soit toujours en retard d'une guerre."

4. Un stockage incontournable, un anonymat tout relatif

Le stockage de nos données : nécessaire et faisable
De l'aveu même du gouvernement, les données observées par ce dispositif seront stockées, quelque part en France. Néanmoins, il assure que seules les informations intéressantes (les données qui correspondraient à un profil suspect) seront entreposées dans des disques durs.
Qu'il soit partiel ou intègre l'ensemble du trafic internet français, ce stockage pose déjà un énorme enjeu de sécurité. Sans mettre en doute la bonne foi des espions, le risque d'intrusion informatique existe.
Par ailleurs, nos interlocuteurs doutent de la possibilité de ne pas stocker du tout, même brièvement, les données de tout le monde.
Certes, des systèmes existent pour se débarrasser des informations parasites, du bruit, que peuvent brasser d'énormes détecteurs de données. C'est par exemple le cas au Cern, le fameux accélérateur de particules, qui se fiche bien (comme les services, on le suppose), de stocker toutes les données de toutes les particules. Mais comme le dit notre spécialiste de l'intelligence artificielle :
• "Même s'il y a une élimination du signal, ça veut tout de même dire qu'à un moment, ils ont les données."

Les deux chercheurs de l'Inria nous font par ailleurs remarquer que l'argument selon lequel stocker toutes ces données serait très difficile, du fait du volume que cela représenterait, n'est pas valable :
• "Le volume n'est pas si conséquent que ça : par exemple, la liste des sites web (juste l'adresse du site, pas le contenu, soit de l'ordre de 100 octets par site) visités par jour, avec éventuellement le temps resté sur chaque page. Le nombre de sites différents visités par jour n'est pas si important par personne (disons 100 par jour)". En imaginant qu'on trace 60 000 000 de personnes, […] ça ferait soit moins d'un terra octet… Ça tient sur un disque dur et c'est faisable de les traiter. Donc en gros, sur un disque dur, nous avons l'ensemble des métadonnées françaises pour la journée (qu'on pourrait même compresser)."

Anonymat des données : illusoire
Là encore, grand scepticisme. Le gouvernement assure que l'anonymat des données collectées selon ce dispositif ne sera levé qu'après avis de la commission de contrôle.
Sauf que pour être efficace, l'algorithme devra savoir que telle ou telle donnée correspond à la même personne. Pour nos deux experts de l'Inria :
• "C'est un contresens juridique de définir la possibilité de retirer l'anonymat : l'anonymat est le fait que rien ni personne ne puisse nous identifier, quelque soit les mesures mises en œuvre (loi informatique et libertés)."

Au passage, cet enjeu pose une autre difficulté : comment les services vont-ils faire pour savoir qu'une même personne se connecte sur un site suspect de chez elle, sur un autre site suspect depuis son téléphone ou depuis un cyber-café ? A l'heure des écrans multiples, des bornes wifi, comment repérer un seul et même individu ?

5. Un contrôle délicat
Il faut des moyens humains et financiers à la hauteur du défi
On l'a déjà vu, en fonction du type d'algorithme choisi par les services, le contrôle prévu dans le projet de loi sera plus ou moins effectif. Ainsi, si l'algorithme évolue sans cesse, on voit mal comment les experts pourront aller vérifier qu'il fonctionne bel et bien uniquement pour détecter d'éventuels terroristes.

De même, certains algorithmes sont par nature très opaques : on parle alors de "boîtes noires". Eh oui ! L'expression utilisée par des conseillers gouvernementaux renvoie aussi à un type d'algorithme très précis. Dans ces cas-là, un peu comme avec un plat très élaboré dont il n'est pas évident de reproduire la recette, ou avec notre réseau de neurones, on sait que ça marche, mais on ne sait pas bien comment…

Le gouvernement rassure en affirmant que le code source de l'algorithme sera remis à la commission de contrôle. Un conseiller parlant même, dans un sourire, "de logiciel libre dans un monde de secret-défense".

L'initiative est louable, mais même en ayant écrit le code source, il arrive que les chercheurs n'arrivent pas à comprendre comment l'algorithme aboutit à un résultat précis. Ce n'est pas donc pas forcément suffisant !

Par ailleurs, le contrôle de cet algorithme sera de toute manière très complexe. Et lourd. Comme le confie notre docteur en intelligence artificielle :
• "Relire le code écrit par quelqu'un d'autre, croyez-moi, c'est l'enfer !"

A l'en croire, des théories mathématiques existent aujourd'hui pour vérifier qu'un algorithme ne sorte pas de son domaine. Problème : elles s'appliquent sur des codes assez limités, comme sur un avion de ligne.

• "C'est valable dans l'aviation, mais le code d'un Airbus est petit par rapport à ce qu'il y a sur votre Windows !" »

La qualité de contrôle de l'algorithme dépendra donc de la quantité et de la qualité des données à dispositions des experts, des moyens humains et financiers à leur disposition, du délai dont ils disposeront… Le tout pour trancher si oui ou non, pour citer le texte, ces données reflètent une réelle menace terroriste.
• « Oppenheimer devant la bombe atomique »

La responsabilité est donc colossale. Et renvoie, selon les chercheurs, toujours au même problème : la question fondamentale n'est pas un enjeu technique mais un enjeu social. Comme le dit Gilles Dowek :

• « Acceptons-nous ou non d'être observés en permanence afin que quelques criminels soient arrêtés au moment où ils en sont encore à préparer un crime ? »

Colin de la Higuera, membre du laboratoire informatique de l'université de Nantes, regrette pour sa part que le sujet, aux « vraies répercussions pour la société », ne fasse pas l'objet d'un débat public avec les chercheurs compétents.

De son côté, notre spécialiste de l'intelligence artificielle se définit « comme Oppenheimer devant la bombe atomique ! ». Et ajoute :

• « J'ai l'impression que les politiques viennent me raconter mon boulot alors qu'ils ne le connaissent pas mieux que moi. […] Ils parlent d'algorithme avec un grand A, comme s'il s'agissait d'un archange tombé du ciel pour arrêter les méchants… Non ! Un algorithme ne sort que des cerveaux humains. Et je vous en conjure : méfiez-vous de ce qui sort de mon cerveau ! »

Après cette lecture, quel est votre avis ?
Cliquez et laissez-nous un commentaire…

Source : http://rue89.nouvelobs.com/2015/04/15/lalgorithme-gouvernement-sera-intrusif-inefficace-prouve-258672
Par Andréa Fradin