Renseignement : pour l'Inria, les boîtes noires seront inefficaces

Dans une note publiée par *Le Monde* en fin de semaine dernière et datée du 30 avril, l'Inria ne se prive pas de dire tout le mal qu'il pense des boîtes noires, ces dispositifs d'écoute qui doivent être placés au cœur des réseaux des opérateurs, FAI et hébergeurs. Rappelons qu'il s'agit là d'**une des dispositions les plus controversées du projet de loi sur le renseignement**, adopté par l'Assemblée nationale mais qui doit encore être examiné par le Sénat, début juin. L'institut de recherche dédié au numérique examine les limites techniques du dispositif imaginé par le gouvernement Valls pour repérer, via les communications sur Internet, la préparation d'actes terroristes. Rappelons que le mécanisme est fondé sur l'analyse en temps réel des métadonnées 'anonymisées', afin de repérer des comportements suspects. Et d'aboutir, dans un second temps, à l'identification des internautes dont le profil se rapprocherait le plus de celui d'un apprenti terroriste, après une opération de 'dés-anonymisation'.

Pour l'Inria, voir une loi manier les notions de donnée anonyme ou anonymisée est déjà plus que discutable. Tout simplement parce qu'il « *n'existe pas aujourd'hui de technique d'anonymisation fiable* ». L'Inria conseille de parler de « *données pseudo-anonymes ou encore de données personnelles* ».

Suspects protégés, innocents accusés ?

Passé ce premier tacle juridico-technique, l'institut de recherche se penche sur ce qui, pour les techniciens, reste la principale hypothèque pesant sur les boîtes noires : leur efficacité. D'abord l'Inria remarque que **la collecte de données peut être facilement contournée**, via plusieurs techniques. Et de citer le chiffrement, via un VPN, « *vers le serveur d'un opérateur ou d'un hébergeur extérieur à la juridiction française* », mais aussi des techniques (spams, utilisation de botnets) permettant de multiplier les connexions sur les sites surveillés afin de protéger l'identité des réels apprentis terroristes. Bref, les personnes que visent réellement les boîtes noires ont à leur disposition des moyens simples à mettre en œuvre pour échapper aux écoutes.

Tandis que d'authentiques innocents pourraient eux être pris dans les mailles des filets des services de renseignement. C'est **le paradoxe des faux-positifs**, bien connu en statistiques et que l'Inria s'emploie à illustrer via un exemple. « *Tout algorithme de détection a une marge d'erreur c'est-à-dire va identifier des personnes sans intention terroriste (des 'faux-positifs'). Si la marge d'erreur est de 1%, ce qui est considéré à ce jour comme très faible, l'algorithme identifiera quelques 600 000 personnes sur une population totale de 60 millions de personnes. Si le nombre de vrais terroristes est par exemple de 60, ces vrais terroristes ne représenteront que 0,01% de la population identifiée* ». Bref, les 'faux-positifs' vont noyer sous le nombre les vrais comportements suspects, rendant le dispositif inopérant en l'état.

L'Inria remarque qu'il est possible d'améliorer la probabilité de détecter de réels suspects grâce à des croisements de données. Si ces techniques sont aujourd'hui « *efficaces* », selon l'institut, elles semblent difficilement utilisables dans le cadre du projet de loi. Tout simplement parce que les données nécessaires « *sont acquises et stockées quasi-exclusivement en dehors du territoire ou de la*

juridiction française par Google, Bing, Facebook, Twitter, Amazon… ». Or, on voit mal ces sociétés américaines, qui tentent déjà de rassurer leurs utilisateurs après le scandale né de leur collaboration active avec les programmes d'écoute de la NSA, accorder aux services français un accès sans limitation à leurs bases de données.

Manque criant de compétences techniques

Enfin, l'Inria relève la faiblesse des compétences techniques de la future Commission nationale de contrôle des techniques de renseignement (CNCTR), chargée de superviser les pratiques des services de renseignement et créée par le projet de loi. En l'état actuel du texte, ladite commission ne doit comporter qu'**un seul membre avec un profil de technicien**, nommé par l'Arcep, le régulateur des télécoms. L'Inria recommande « *une représentation équilibrée entre les compétences numériques et juridiques* », avec des membres nommés par l'Arcep, la Cnil mais aussi Allistene, l'alliance des organismes, universités et écoles en sciences et technologies du numérique (dont font partie l'Inria, le CNRS ou encore le CEA).

Ce n'est pas la première fois que l'absence de maîtrise des sujets techniques est pointée du doigt lors de l'examen de ce projet de loi. Le sujet avait déjà largement occupé le devant de la scène lors des débats à l'Assemblée Nationale, les défenseurs du texte (le ministre Bernard Cazeneuve et le rapporteur Jean-Jacques Urvoas) montrant à plusieurs reprises leurs limites sur ce terrain. Rappelons que les hébergeurs, concernés par le dispositif des boîtes noires, ont un temps menacé de déménager leurs activités hors de France avant d'obtenir un amendement cousu main (même si les décrets d'applications devront valider les aménagements que disent avoir obtenus les hébergeurs). La semaine dernière, lors des premières phases de l'examen du projet de loi au Sénat, le sénateur UMP Claude Malhuret remarquait que les élus étaient dans « *l'incapacité* » de comprendre les algorithmes que vont embarquer ces fameuses boîtes noires. Ce qui n'a pas empêché la Commission des affaires étrangères, de la défense et des forces armées, chargée de l'examen préliminaire du texte, d'estimer in fine le dispositif « *équilibré* ».

A lire aussi :

Loi sur le renseignement : OVH satisfait des concessions du gouvernement
Renseignement : le gouvernement tente de repeindre en rose ses boîtes noires
Projet de loi sur le renseignement : les 5 sujets qui fâchent
Loi sur le renseignement : le casse-tête des boîtes noires pour l'Internet français

Crédit photo © kurhan- shutterstock