

Le téléphone du futur : plus intelligent pour une exploitation optimale des fréquences

Mots clés

Radio cognitive,
Téléphonie mobile,
Intelligence artificielle,
Réseaux flexibles

■ Romain COUILLET^{1,2}, Mérouane DEBBAH²
ST-Ericsson, Sophia Antipolis¹, Alcatel-Lucent Chair, Supélec²

Agacé de voir vos communications téléphoniques interrompues pour faute de couverture réseau alors que dans le même temps, votre point d'accès Internet télécharge pour vous une vidéo à haut débit ? Agacé de voir également que, dans la rue, à quelques mètres de vous, un promeneur, le téléphone à l'oreille, semble ne pas rencontrer les mêmes soucis que vous ? Agacé de manière générale de vous dire que ce téléphone, que vous avez payé pourtant si cher, n'est pas capable de comprendre de lui-même que partout autour de vous des flots immenses d'information circulent librement et qu'il serait si simple d'en profiter ?

Les chercheurs du monde des télécommunications se posent aujourd'hui les mêmes questions et parviennent peu à peu à des réponses concrètes. Ces questions ont amené il y a maintenant dix ans à repenser totalement le mode opératoire des protocoles de communications pour passer progressivement du monde de la radio classique au monde de la radio cognitive. Demain, votre téléphone portable sera capable de comprendre qu'il existe autour de lui de multiples canaux de communications et qu'il suffit alors de profiter du canal le plus intéressant pour lui. Pour ce faire, votre téléphone devra être doté d'une intelligence avancée, qui lui permettra par exemple (i) d'intégrer de l'information concernant son environnement, de sorte qu'il comprenne par lui-même le monde qui l'entoure, (ii) de formuler des requêtes aux appareils intelligents qui l'entourent, etc.

Comment rendre un téléphone intelligent ? Tout simplement en le faisant raisonner de la manière la plus avancée que nous connaissons : comme le cerveau humain. Votre téléphone sera un être électronique à part entière adapté à votre usage personnel, qui naîtra lors de son achat et vieillira avec sagesse, en comprenant au fil du temps l'environnement qui l'entoure et plus généralement vos propres habitudes. Ce téléphone idéal n'existe pas encore, faute de pouvoir effectuer la quantité de calculs que cette intelligence requiert et faute de pouvoir stocker la quantité colossale d'information qu'il lui faut intégrer >

L'ESSENTIEL

Au jour où les communications mobiles demandent de plus en plus de vitesse de transfert de larges volumes de données à destination d'utilisateurs de plus en plus nombreux, il apparaît que les limites physiques des protocoles de communication sont bientôt atteintes. Une révolution technologique est ainsi nécessaire et sur le point d'éclorre : celle-ci passe par la mise en place de systèmes de communications opportunistes, coopératifs, autonomes et idéalement suffisamment intelligents pour servir au mieux les requêtes de l'utilisateur. Ces différents aspects, certains d'ores et déjà d'actualité, d'autres à l'état embryonnaire sont discutés successivement dans cette étude.

SYNOPSIS

Mobile communications require increasingly fast transmissions of voluminous data for an increasingly large audience of users, to the extent that the physical limits of communication are wearing out. A technological revolution is therefore necessary and on the verge to arise: the latter requires autonomous, opportunistic and cooperative communication bodies, ideally able to meet exactly the individual user's requests. These aspects, some already planned for future communication networks, and some still in their infancy, are introduced and motivated in this article

> à tout instant. Néanmoins, ces problèmes sont d'ordre purement techniques et sauront être dépassés dans les générations à venir, alors que la théorie mathématique qui permet de créer cette intelligence artificielle dans le téléphone, existe bel et bien.

1. Un peu d'histoire

Le domaine des télécommunications a ceci de très spécifique, comparé à la plupart des importants domaines de recherche actuels (histoire, médecine, philosophie par exemple), qu'en 1948, par l'intermédiaire d'une unique publication de 55 pages, Claude Elwood Shannon [1] a radicalement bouleversé la vision des télécommunications modernes en inventant une toute nouvelle théorie, connue sous le nom de théorie de l'information. Dans cette publication, Shannon apporte en particulier la réponse à un problème philosophique fondamental en termes de communication : quelle quantité d'information peut être véhiculée entre deux interlocuteurs communiquant dans un environnement et en un temps donnés, de sorte que chacun comprenne sans erreur l'information qui lui est destinée ? Depuis que cette réponse a été formulée par Shannon, nous voyons se développer autour de nous des téléphones, cartes WiFi, etc... capables de transmettre de plus en plus de données à la seconde. Cependant, nous ne devons pas nous leurrer, et c'est en ceci que le travail de Shannon est si fondamental : le débit d'information transmissible sans erreur est naturellement limité par l'environnement de communication, la bande de fréquences (appelée plus couramment bande passante) utilisée, et la puissance des signaux émis. Ainsi, lorsque que chacune de ces trois ressources fondamentales sera épuisée, nous ne pourrons plus transmettre d'information à un débit supérieur. Cela n'est heureusement pas prêt d'arriver pensez-vous ? Détrompez-vous, cette situation est d'ores et déjà d'actualité. Devant cet état de fait, Joseph Mitola, il y a dix ans de cela, comprit qu'une révolution dans le mode d'évolution des télécommunications devait avoir lieu de toute urgence. La principale conclusion de Mitola est que l'évolution rapide et relativement mal contrôlée de l'utilisation des ressources de télécommunication (en particulier de la bande passante) a mené à un gaspillage énorme de ces ressources. Le plus simple des exemples de gaspillage est le mode de fonctionnement du standard GSM (aussi connu sous la dénomination 2G) : ce dernier permet à huit utilisateurs du réseau téléphonique mobile de se connecter simultanément à une station de base de leur fournisseur d'accès. Lorsqu'un seul utilisateur est connecté à la station de base, cet utilisateur n'utilise cependant qu'un huitième des ressources, et non la totalité qui est pourtant gratuitement disponible. Le caractère statique des protocoles actuels de communication amène à se demander comment rendre le domaine radio plus « flexible ». De cette importante réalisation concernant directement la pérennité des télécommunications modernes, est donc né le domaine des radios flexibles ; il tend à rendre les appareils de communication plus

autonomes, capables de décider quelles ressources utiliser et comment les utiliser efficacement. Si aucune couverture réseau GSM n'est possible dans une pièce de la maison, pourquoi ne pas profiter de la borne WiFi par exemple ? Il est fortement désirable qu'un système de communication sache prendre ce type de décision de manière autonome. Il est même extrêmement souhaitable que nos appareils de télécommunication puissent effectuer exactement ce raisonnement de manière intelligente, en réfléchissant tel que le ferait un être humain confronté à une telle situation.

1.1. Claude Shannon et la théorie de l'information

L'histoire des télécommunications modernes est divisée en deux ères : l'avant et l'après Shannon. Avant 1948 et la publication de la *Théorie Mathématique de la Communication* par Claude Elwood Shannon, les principales technologies de télécommunications ne consistaient qu'en des essais successifs de méthodes *ad-hoc* permettant de transmettre toujours plus d'information de manière fiable. Néanmoins, les limitations physiques de tels systèmes n'étaient pas comprises : la question du débit auquel l'information peut être transmise avec assurance d'une réception sans erreur semblait ne pas avoir de réponse. Comment peut-on en effet imaginer une réponse à un problème à la fois physique et philosophique si avancé ? Par ailleurs l'hypothèse initiale (et, il est vrai, relativement naturelle), que la meilleure façon de transmettre efficacement plus d'information est d'augmenter la puissance d'émission est en vérité totalement fautive.

En effet, prenons la situation d'une conversation entre vous-même et l'un de vos amis. Si le niveau de bruit ambiant, généré typiquement par d'autres conversations voisines, est trop élevé, votre ami, qui reçoit le signal acoustique que vous émettez, ne pourra peut-être pas décoder de manière fiable le message qui lui est destiné. Il semble relativement intuitif, et c'est ce que nous ferions, faute de meilleure méthode, d'élever la voix pour nous faire comprendre. Shannon a démontré que cette méthode, qui permet en effet de rendre la communication plus fiable, est cependant très inefficace, ou plus justement qu'elle a un rendement très faible puisque, pour simplement doubler le débit d'information transmissible avec fiabilité, il faut quadrupler la puissance émise, et cela coûte très cher ! Par contre, si, au lieu de quadrupler la puissance de notre voix, nous pouvions élargir la gamme de fréquences à laquelle nos cordes vocales peuvent vibrer et si notre oreille interne est, elle aussi, capable de recevoir ces signaux, Shannon montre que nous pourrions transmettre deux fois plus d'information fiable en utilisant simple-

ment deux fois plus de bande passante, i.e. une gamme de fréquences double. Notez par exemple que, lorsque vous êtes enrhumé, certaines gammes de fréquences (les sons aigus typiquement) ne sont plus transmissibles ; il vous faut alors produire un certain effort cérébral pour changer votre méthode de transmission de parole pour l'adapter à la nouvelle gamme de fréquences que vous pouvez utiliser, et la quantité d'information transmissible est alors significativement réduite. Alternativement, si vous décidez de ne pas altérer votre voix pour mieux vous faire comprendre, certains sons ne sont pas produits, et il vous faut répéter ou modifier la tournure de vos phrases pour vous faire comprendre, auquel cas, à nouveau, le débit d'information utile transmise est fortement réduit.

C'est en ce premier constat que le résultat de Shannon a initialement révolutionné le monde des télécommunications modernes : utiliser toujours plus de puissance est une méthode inefficace pour gagner en débit de transmission fiable alors qu'utiliser plus de bande passante est au contraire une méthode simple, efficace et qui, à l'époque de Shannon, était d'autant plus une solution bon marché, les fréquences de transmission constituant une ressource quasi inexploitée. Comprenez par exemple que, sans ce résultat de Shannon, aucune communication par satellite ne serait possible de nos jours : comment en effet transmettre d'un satellite positionné à 40 000 km au-dessus de nos têtes, une énergie suffisante pour couvrir le niveau de bruit électromagnétique ambiant (dû en particulier à tous nos produits électroménagers et au réseau électrique local) et assurer une réception haut débit fiable ? Les communications satellites utilisent de nos jours des bandes passantes larges parfois de 100 MHz (ce chiffre peut être typiquement comparé aux 20 MHz, soit 5 fois moins de bande passante, consacrés à l'intégralité de la bande passante de la radio FM française qui elle est transmise par voies terrestres).

Le deuxième et le plus fondamental des résultats de Shannon réside dans l'expression mathématique théorique du débit d'information qui peut être transmis puis décodé de manière fiable d'un émetteur à un récepteur. Cette valeur de débit est connue sous le terme « capacité », s'exprime mathématiquement sous la forme :

$$= F \log_2 \left(1 + \frac{P}{B} \right)$$

et a pour unité les bits par seconde. Dans cette expression, P est la puissance du signal émis, B , la puissance du bruit ambiant (dû aux interférences électromagnétiques) et F est la bande de fréquences, ou bande passante, dédiée à la communication. La fonction mathématique \log_2 a la propriété de croître moins rapidement qu'une racine carrée, moins rapidement qu'une racine cubique etc... ainsi, nous vérifions bien ici que lorsque la puissance P augmente, le terme augmente également, mais de manière très très lente, de sorte qu'augmenter la puissance d'émission n'apporte qu'un gain mineur en capacité. Par contre, augmenter F , la bande passante, augmente d'autant la valeur de la capacité C . Le premier constat de Shannon est aisément confirmé.

Shannon ne nous dit pas, par contre, par quelles techniques mettre en œuvre pour atteindre cette valeur de capacité : ce travail est laissé à l'ingénieur en télécommunications. Cependant, 60 ans après la publication des résultats de Shannon, nous connaissons beaucoup de techniques capables de s'approcher extrêmement près de cette limite.

1.2. Les systèmes multi-antennes, une nouvelle dimension ?

En 1999, Emre Telatar [2] se rend compte que la capacité de Shannon n'est en fait valide que pour une paire transmetteur-récepteur ne possédant chacun qu'une seule antenne. Utiliser plusieurs antennes sur chaque appareil tout en conservant la même puissance totale d'émission P (avec 2 antennes, la puissance moyenne émise par chaque antenne serait alors $P/2$) ne semble apporter *a priori* aucune performance en termes de capacité, pensez-vous ? C'est en vérité, et de manière très surprenante, une fausse conception : multiplier par deux le nombre d'antennes en émission et en réception peut effectivement doubler la capacité de transmission tout en gardant la même puissance totale à l'émission. La raison fondamentale réside dans la diversité des canaux de transmission que peuvent rencontrer les ondes transmises par chaque antenne. En effet, avec deux antennes en émission et en réception, il existe quatre chemins de propagation, au lieu d'un seul : un chemin de la première antenne d'émission à la première antenne de réception, un autre de la première antenne d'émission à la seconde antenne de réception, et de même pour la seconde antenne d'émission. Telatar a alors remarqué que, si ces chemins de propagation sont très différents (le terme consacré est « indépendants »), alors le récepteur sera capable de démêler les combinaisons des deux signaux émis qui s'interfèrent à la fois au niveau de sa première antenne et au niveau de sa deuxième antenne. Ainsi, le récepteur sera en possession d'un premier signal (celui émis par la première antenne du transmetteur) reçu à une puissance $P/2$ et d'un deuxième signal totalement décorrélé du premier (celui émis par la deuxième antenne du transmetteur) et de puissance $P/2$ également. Le débit accessible à la transmission est alors la somme de deux débits indépendants (c'est un peu comme si l'émetteur avait utilisé deux bandes de fréquence différentes) de signaux de puissance $P/2$. Shannon nous rappelle alors qu'envoyer un signal de puissance P sur une bande passante réduite est moins efficace qu'envoyer un signal sur deux fréquences distinctes avec une puissance $P/2$ pour chacun : la capacité d'un système avec deux antennes est donc supérieure à celle d'un système mono-antenne, sous la même contrainte de puissance d'émission totale. Il peut même être démontré que cette capacité, lorsque le rapport P/B est grand, égale deux fois la capacité du système mono-antenne.

Il semble dès lors que l'utilisation de large bandes passantes ne soit plus nécessaire en fin de compte puisqu'alors deux antennes peuvent remplacer l'utilisation de deux fois plus de fréquences. En lieu et place de large bandes passantes

tes, pourquoi ne pas alors utiliser des multitudes d'antennes au niveau du transmetteur et du récepteur ? Le problème de la surexploitation de la bande passante, que nous avons mentionné dans l'introduction et qui sera développé dans le prochain paragraphe, n'aurait-il donc pas lieu d'être ? La réalité est malheureusement encore une fois plus cruelle que la théorie : nous avons évoqué que les systèmes multi-antennes génèrent d'importants gains de capacités sous deux conditions, quelque peu masquées dans notre texte :

1. les canaux de propagation entre tous les couples d'antennes émettrices/réceptrices doivent être extrêmement différents (indépendants) et
2. le rapport signal sur bruit ambiant P/B doit être grand.

Or, dans la pratique, ces deux conditions sont très rarement rencontrées. La plus problématique de ces conditions étant la contrainte d'indépendance des voies de propagation. Pour assurer cette indépendance, deux conditions fondamentales sont requises à nouveau : l'espacement entre chaque antenne d'émission et de réception doit être au moins de l'ordre de la longueur d'onde des signaux émis, et l'environnement de propagation doit être composé de nombreux réflecteurs : cette dernière condition assure que toutes les ondes transmises sont réfléchies de manières très différentes. Dans le cas des communications de téléphonie mobile, l'ordre des longueurs d'onde transmises environne les 10 cm : ainsi, un téléphone portable, de taille également 10 cm, ne peut contenir qu'un maximum de deux à quatre antennes pour assurer un gain de capacité théorique significatif par rapport au cas mono-antenne. Par ailleurs, en dehors des zones urbaines où le nombre d'immeubles et plus généralement le nombre de réflecteurs dans l'environnement est faible, le gain en capacité est fortement réduit comparé aux gains potentiels prévus par Telatar. Le problème reste donc entier ! Même si nous voyons de plus en plus apparaître sur nos points d'accès Internet ou nos téléphones cellulaires récents de multiples antennes (parfois dissimulées dans le boîtier), les résultats de Telatar ne s'appliquent qu'à un nombre limité de situations pratiques. Par ailleurs, le besoin d'un rapport P/B important est également un problème majeur à l'heure actuelle dans les réseaux mobiles composés de multiples cellules de couverture ; au bord de ces cellules (lorsque l'utilisateur du réseau mobile est approximativement à équidistance entre deux stations de base), l'utilisateur connecté à une station de base est fortement interféré par les signaux provenant de la station de base voisine, générant ainsi un niveau de bruit B important. Posséder de nombreuses antennes dans ce type de situation n'apporte également aucun gain en capacité. Heureusement, cette situation d'interférence en bordure de cellule peut être fortement évitée si la cellule voisine, au lieu de se comporter en ennemi producteur d'interférence, se comporte en allié producteur de signal utile à l'utilisateur. Ainsi, depuis soixante ans, il est de notoriété que l'usage de puissances importantes, par ailleurs dangereuses pour la santé, est une stratégie inefficace dans l'objectif de transmettre de plus en plus d'information à la

seconde. Malgré les résultats prometteurs de Telatar cinquante ans après Shannon, le constat final est simple : pour gagner en débit, il faut toujours plus de bande passante. Paradoxalement, alors que de nombreux produits de la vie courante demandent toujours plus d'énergie et que les ressources fossiles ou nucléaires semblent former la seule ressource indispensable à l'évolution de notre société technologique moderne, le monde des télécommunications apporte une affirmation totalement orthogonale : l'unique ressource fiable de transfert d'information est totalement immatérielle et à première vue gratuite car inépuisable, les fréquences radio.

1.3. De Shannon à nos jours : une raréfaction des ressources

Malheureusement, il est également de notoriété que cette dernière affirmation est fautive : les bandes de fréquences allouées aux fournisseurs de services téléphoniques, Internet et autres, sont loin d'être gratuites. Et pourtant, les fréquences *a priori* utilisables s'étalent de zéro Hertz à une infinité de Hertz. En vérité, deux problèmes majeurs sont rencontrés lorsque des signaux sont transmis à haute fréquence : techniquement il est très difficile et très coûteux de générer des signaux stables de très haute fréquence et d'autre part, physiquement, les signaux haute fréquence s'atténuent fortement lorsqu'ils se propagent dans l'atmosphère et sont ainsi très impropres à la transmission de données à longue distance. Ainsi, seuls quelques gigahertz du spectre total de fréquences sont à ce jour utilisables pour les transmissions à longue distance en particulier. Ce qui fait de la bande passante une ressource immatérielle soit, mais rare et donc chère.

De part la croissance exceptionnelle ces dernières années des services de communication sans fil à usages public ou privé, le spectre des fréquences utilisables tend à saturer. L'exemple le plus frappant est donné par la carte d'allocation de ressources spectrales aux États-Unis (<http://www.ntia.doc.gov/osmhome/allochrt.pdf>) qui présente une utilisation quasi complète des ressources allant de 9 KHz à 300 GHz. Un autre exemple d'actualité est celui de la télévision analogique publique française qui, sur le point de disparaître, va se voir désallouer sa bande passante dans la gamme de 47 MHz à 862 MHz au profit de services alternatifs, tels que la téléphonie mobile privée dont les principaux acteurs se disputent l'allocation de cette bande. Dès lors, augmenter le débit de transmission d'une technologie sans fil par l'usage d'un spectre de fréquences plus large devient financièrement inaccessible aux fournisseurs de services. L'ère Shannon est désormais dépassée, et des solutions innovantes pour pallier au problème de raréfaction des ressources fréquentielles sont nécessaires. C'est dans cet état d'esprit que Joseph Mitola introduisit en 1999 [3] la philosophie des radios flexibles et des radios cognitives.

1.4. Joseph Mitola et la radio cognitive

Le constat initial de Mitola a été de réaliser que nos téléphones portables, capables par exemple aujourd'hui d'utiliser les réseaux de communication mobile 2G, 3G, ou de se connecter à l'Internet via un réseau WiFi local, ne sont en vérité que des instruments rudimentaires destinés à remplir les tâches spécifiques pour lesquelles ils ont été programmés. Par exemple, lorsque votre téléphone portable se trouve dans une situation très particulière, telle que dans un hélicoptère en déplacement rapide à travers une zone montagneuse, et que cette situation n'a pas été envisagée par les concepteurs du service téléphonique, il vous sera très certainement difficile d'engager une communication téléphonique alors même que le téléphone pourrait se rendre compte de l'environnement spécial à travers lequel les ondes se propagent ; mais il ne se posera pas ce genre de questions, n'étant pas programmé pour cela. Selon les termes de Mitola, le téléphone « ne sait pas ce qu'il connaît ». Beaucoup d'informations accessibles au téléphone pourraient être transmises au réseau de sorte que ce dernier s'adapte à cette situation si particulière.

Revenons maintenant à l'utilisation intensive des bandes passantes de communication. Un autre constat de Mitola et des chercheurs du domaine des radios cognitives est que le spectre de fréquences, pourtant si cher, est utilisé de manière très inefficace : rappelons par exemple le cas des communications 2G qui peuvent n'utiliser parfois qu'un huitième de la bande passante disponible. Lorsque de nouveaux utilisateurs de protocoles différents (WiFi, 3G etc.) se trouvent dans des situations où leur réseau ne leur permet pas d'atteindre leur requête de communication, pourquoi ne pas profiter de l'excédent de bande passante du réseau 2G ? Le téléphone pourrait même idéalement devenir une entité capable de se connecter à tout réseau alentour possible ; il pourrait même ne pas savoir à la base utiliser quelque réseau de communication que ce soit mais pourrait apprendre ; pour ceci, il lui suffirait (i) d'observer qu'une station lui offre un certain service (qu'il ne connaît pas initialement) dans son voisinage, (ii) si le service proposé semble répondre à sa requête, de demander à la station de lui transmettre le « logiciel » lui permettant d'implémenter ce service, (iii) après téléchargement et installation automatique, une communication peut alors être opérée. Cette idée d'un appareil capable de télécharger et apprendre des nouveaux services de communication est un sujet important du monde des radios cognitives, connu sous le nom de radio logicielle (*software defined radio* en anglais).

La rationalisation des différents protocoles de communication peut même être poussée plus loin : lorsqu'une bande passante a un coût d'exploitation très important et qu'il arrive que le réseau se trouve parfois totalement inutilisé, pourquoi ne pas réguler également le coût d'accès au réseau ? Ainsi, lorsqu'un utilisateur requiert une importante garantie de service (par exemple, s'il souhaite effectuer un appel urgent) sur une certaine bande passante,

il paiera cher l'exploitation de cette dernière, alors que lorsque l'utilisateur ne désire pas une importante qualité de service (par exemple s'il accède à une page web mais qu'il est prêt à attendre quelques secondes pour que la page se charge), il pourra obtenir ce service à un prix modéré ; l'intérêt de l'opérateur dans cette démarche sera l'assurance d'une occupation plus efficace de la bande passante qui lui est allouée : lorsque beaucoup d'utilisateurs demandent un accès simultané, les premiers servis seront les utilisateurs qui paieront le prix fort, alors que, lorsqu'aucune requête urgente n'est demandée sur le réseau, ce dernier pourra être exploité à un moindre coût pour les utilisateurs. Ainsi, le gaspillage de bande passante se verra fortement limité.

Le gaspillage de bande passante apparaît également du fait que la capacité de Shannon d'un certain canal de communication dépasse parfois largement le débit moyen requis par un utilisateur (par exemple, l'utilisateur peut utiliser un très fort débit à un instant, puis un débit beaucoup plus faible l'instant suivant). Dans ce cas, la différence « capacité de Shannon moins débit moyen » peut être large et génère ainsi un énorme gaspillage. Dans les cas d'utilisation modérée du réseau, d'autres utilisateurs, opportunistes, peuvent profiter de l'excès de débit accessible. De ce scénario simple est né le concept des réseaux à utilisateurs primaires (ceux qui paient pour une garantie de service) et à utilisateurs secondaires (ceux qui profitent à un faible coût des opportunités du réseau).

En somme, l'idée générale de Mitola est de tendre vers une unification des réseaux de communication qui, s'étant développés de manières indépendante et inefficace, gaspillent beaucoup la ressource essentielle qu'est la bande passante, à un tel point qu'aujourd'hui la carte des fréquences est totalement saturée. La rationalisation de l'exploitation des ressources semble pouvoir permettre de renverser ce processus de gaspillage en un processus vertueux qui verra nos réseaux de télécommunication futurs évoluer intelligemment et en harmonie.

2. De l'être humain à la machine

Cependant, ces dernières idées d'harmonisation des réseaux et d'exploitation plus dynamique des ressources, ne révèlent qu'en filigrane la dimension « intelligente » des télécommunications du futur. En effet et à nouveau, les téléphones du futur seront capables selon Mitola de scruter les opportunités de bande passante et de communiquer de manière plus fiable, mais cette intelligence additionnelle repose encore sur des algorithmes programmés par des ingénieurs. Ces nouvelles technologies radios ne seront donc pas réellement cognitives mais plutôt flexibles. Est-ce qu'une machine, un téléphone portable dans notre cas, peut réellement être doué d'intelligence ? Sous quelle forme doit alors apparaître cette intelligence ? A supposer que cela soit possible, est-ce qu'un téléphone portable peut devenir plus intelligent qu'un être humain ? Toutes ces

questions ont été débattues depuis de longs siècles, d'un point de vue philosophique tout d'abord puis technique par la suite. Dans ce qui suit, les auteurs proposent leur point de vue personnel, qui peut fortement diverger de l'avis d'autres chercheurs, et donneront des exemples de cas pratiques pour lesquels ils ont créé une intelligence virtuelle pour des appareils de télécommunications mobiles.

2.1. Une radio plus intelligente que l'Homme ?

Il peut sembler difficile de quantifier les capacités d'un système conçu par l'Homme à devenir intelligent. En vérité, les limites de cette intelligence sont claires : d'un point de vue philosophique, il n'est pas possible pour l'Homme de concevoir une intelligence qui lui soit supérieure. Raisonnons par l'absurde : si l'Homme est un jour capable de mettre en place un mode de raisonnement supérieur à sa propre intelligence, alors il doit à un moment de la création de cette nouvelle forme d'intelligence imaginer une méthode de raisonnement intellectuellement inaccessible aux moyens de raisonnement (ceux de l'Homme) qui ont établi cette nouvelle méthode. Ce qui est une contradiction. Il existe donc une limite philosophique infranchissable à l'intelligence artificielle et celle-ci est donnée par l'intelligence du créateur de cette intelligence artificielle. D'ailleurs, si l'Homme est amené à comprendre un jour l'intégralité des mécanismes physiologiques qui sont à la base de son activité cérébrale, alors il ne sera pas difficile (du moins en théorie) de mettre en place une machine qui mime ces mécanismes et ainsi de « créer » un être humain artificiel. Nos téléphones portables ne seront donc jamais plus intelligents que nous ne le sommes. Ils seront vraisemblablement capables de réfléchir plus vite (ce qui est déjà le cas, grâce à l'impressionnant nombre de calculs qu'un ordinateur peut effectuer en une seconde), de stocker plus d'information et de transmettre cette information plus vite, mais ils ne pourront jamais être amenés à raisonner de manière incompréhensible pour l'Homme.

La plupart des machines actuelles, telles que nos ordinateurs, téléphones portables et autres machines du quotidien, sont souvent qualifiées de machines élémentaires en particulier pour une raison fondamentale : elles ne fonctionnent que suivant une logique binaire d'opérations pour lesquelles elles ont été programmées. Ainsi, lorsqu'une de ces machines est prise au dépourvu par une information manquante, elle ne peut en général pas réagir. En réalité, et c'est ce qui se passe souvent dans nos téléphones portables, lorsqu'une information cruciale est manquante (par exemple si le téléphone ne connaît pas du tout le canal à travers lequel l'information qu'il reçoit est transmise), le téléphone va le plus souvent lancer l'application « recherche de canal » qui est enregistrée dans sa mémoire. Ce programme va retourner une valeur empirique pour ce canal, en fonction des signaux récents reçus par le téléphone. Ce dernier sera alors à même de poursuivre son opération de communication. Cependant, tout ce mécanisme, aussi bien huilé soit-il, est fortement statique et bi-

naire : (i) le téléphone requiert une information en lançant un programme, (ii) ce programme retourne une réponse, qui peut potentiellement être fortement inexacte, que le téléphone va alors utiliser. Jamais dans cette suite d'action le téléphone ne se demande si les erreurs potentielles dans l'estimation du canal par le programme peuvent conduire à des soucis dans la qualité de la communication ressentie par l'utilisateur. Idéalement, nous souhaiterions que ce téléphone soit un peu plus sceptique concernant l'information que le programme en mémoire lui fournit : il serait en effet bon que le téléphone réalise s'il met en danger la communication, s'il ne peut pas faire mieux que ce que le programme lui propose etc... De manière générale, il serait essentiel que le téléphone puisse agir en fonction de son degré de confiance sur l'information qui lui est fournie, parce que c'est exactement ce que l'Homme fait au quotidien lorsqu'il reçoit une information. Si nous autres, êtres humains étions amenés à croire aveuglément toute information qui nous est donnée, quitte à changer radicalement d'opinion lorsqu'une information opposée nous est apportée, l'échange même d'information (que nous appelons conversation) n'aurait que fort peu d'intérêt.

Tout appareil de communication, ou de manière générale tout appareil électronique, est en vérité capable en théorie de raisonner en toute honnêteté vis-à-vis de l'information qu'il possède ou qu'il acquiert. Nous allons tenter de donner dans le prochain paragraphe les grandes lignes de la méthodologie mathématique qui permet aux machines de devenir conscientes, avant d'expliquer en dernière section de cet article comment les machines peuvent évoluer et tendre vers la sagesse.

2.2. Honnêteté et inquisition

Les deux outils mathématiques fondamentaux [4] qui recourent la notion philosophique d'honnêteté vis-à-vis de l'information connue sont (i) les probabilités bayésiennes et (ii) le principe d'entropie maximale, que nous allons aborder tour à tour.

2.2.1 Degrés de croyance et probabilités

Lorsque l'on parle de probabilités, on pense immédiatement au jet de la pièce de monnaie qui a une chance sur deux de tomber pile et une chance sur deux de tomber face ou au jet du dé non pipé qui a une chance sur six de tomber sur chacune de ses faces. Cette vision des probabilités est en vérité malheureuse et peut prêter à confusion. Selon les probabilistes, dits orthodoxes, cette vision est cependant correcte : en effet, selon eux, toute pièce ou tout dé contient en lui-même un aléa, une force métaphysique pourrait-on penser, qui tend à générer cette chance sur deux ou sur six de tomber sur une face donnée. Selon les probabilistes bayésiens (de l'illustre mathématicien Bayes), la nature n'a rien d'aléatoire, mais est régie par des lois physiques que chacun est capable de comprendre. Imaginons qu'un être humain très vif de pensée observe

un jet de pièce : plus exactement, disons que cet être humain est capable d'évaluer de manière exacte la force avec laquelle la pièce est initialement lancée, et la direction du lancer. Connaissant les paramètres physiques de l'atmosphère et de la table sur laquelle la pièce va achever son vol, si cet être humain est réellement très vif d'esprit, il pourra calculer de manière certaine, avant même la chute de la pièce, le résultat du lancer. Et donc il pourra anticiper avec une « probabilité de 100 % » le résultat du lancer. Suivant ce point de vue, il n'y a pas d'aléa dans le résultat du lancer, et la probabilité du résultat « pile » ne sera pas une chance sur deux selon cet observateur. Par contre, son voisin, qui lui n'a pas pris le temps d'observer le jet de la pièce, s'en tiendra à dire, en toute honnêteté, que les chances du résultat pile sont bien d'une sur deux. En réalité, personne n'est humainement capable d'évaluer avec précision la vitesse et la direction initiales du lancer de la pièce, et donc le calcul du résultat devra considérer toutes les hypothèses de vitesses et directions et leurs degrés de vraisemblance respectifs (selon l'observateur toujours). C'est en cela que les probabilités bayésiennes expriment, non pas les aléas de la réalisation d'un processus¹, mais le degré de croyance qu'un observateur a sur la réalisation de ce processus. Si l'observateur possède de bonnes informations *a priori* sur cette réalisation, la distribution de probabilité associée à tous les événements possibles tendra à montrer une très forte probabilité pour l'événement qui va réellement se passer, et une très faible probabilité pour les événements qui n'auront vraisemblablement pas lieu (à moins que l'informateur de cet observateur lui ait fourni de mauvaises informations). Dans le cas contraire où l'observateur ne connaît que peu de choses *a priori*, la distribution de probabilités n'exclura que peu d'événements non probables.

Dès lors, l'âge des auteurs de cet article, le poids de la tour Eiffel, le temps qu'il fera demain sont tout autant de variables pour lesquelles chacun d'entre vous, lecteurs, possédez des degrés de croyance et donc des probabilités *a priori*. Lorsque nos appareils de communication reçoivent une information concernant le canal de propagation des signaux, la puissance de la station émettrice, le nombre d'utilisateurs voisins interférant etc., il serait intéressant, au lieu de prendre une décision *ad hoc* potentiellement erronée fournie par un programme annexe concernant toutes ces variables, que ces appareils soient capables d'établir des degrés de confiance sur les résultats et agissent en conséquence. Disons que vous souhaitez organiser un pique-nique pour ce week end mais que vous n'avez aucune idée des conditions météorologiques prévues (soit une chance sur deux pour un temps pluvieux et une chance sur deux pour un temps ensoleillé). Si une cartomancienne vous indique que les cartes prédisent une météo excellente, il serait souhaitable que votre degré de confiance, initialement d'une chance sur deux pour un temps excellent,

ne saute pas immédiatement à une confiance absolue pour l'hypothèse d'un temps excellent. La cartomancienne ou le logiciel qui estiment un paramètre pour le téléphone ne participent qu'à l'altération de vos degrés de croyance.

La loi mathématique essentielle qui permet cette altération des croyances et qui représente parfaitement notre méthode humaine de raisonnement est la loi dite de Bayes. Considérons l'événement, dénoté A , « la météo sera bonne ce week end » et l'événement, dénoté B , « la cartomancienne prévoit un temps ensoleillé ». Dénotons $P(A)$ et $P(B)$ la probabilité, ou le degré de confiance, attachés aux événements A et B , $P(A|B)$ la probabilité « la météo sera bonne, sachant que la cartomancienne a prévu un temps ensoleillé » et enfin $P(B|A)$ la probabilité « la cartomancienne aura prévu un temps ensoleillé s'il s'avère qu'il fera beau ». La connaissance *a priori* du pique-niqueur est $P(A)$ qui vaut une chance sur deux, donc $P(A)=1/2$. La règle de Bayes dit alors, qu'après avoir parlé à la cartomancienne, la nouvelle probabilité qu'il fasse beau ce week end sera :

$$P(A|B) = \frac{P(A) \times P(A|B)}{P(B)} = \frac{1}{2} \frac{P(A|B)}{P(B)}$$

après avoir remplacé $P(A)$ par sa valeur $1/2$.

Ainsi, supposons que l'on pense que la cartomancienne est honnête, à savoir, elle n'a aucune raison avant de tirer les cartes de décréter qu'il fera beau ou non ce week end ; alors $P(B)$, la probabilité qu'elle prévienne un temps ensoleillé sans avoir tiré les cartes au préalable, vaut également $1/2$. Et donc la probabilité qu'il fasse beau après avoir parlé à la cartomancienne vaut la probabilité que la cartomancienne annonce effectivement un temps ensoleillé lorsque ce sera effectivement le cas. En bref, notre degré de croyance pour un temps ensoleillé ce week end dépendra exclusivement de notre degré de confiance en la perspicacité de la cartomancienne. Si selon vous, les cartes ne sont pas fiables, alors $P(B|A)$ vaut également une chance sur deux, et donc votre confiance en un temps ensoleillé ce week end n'évoluera pas de sa position initiale : $P(A|B) = P(A) = 1/2$. Par contre, si vous êtes convaincu à 99 % de la fiabilité des cartes, alors $P(A|B)=P(B|A) = 99\%$. Néanmoins, cartes fiables ou non, vous pouvez initialement penser que, pour vous faire plaisir, la cartomancienne a tenté avant de tirer les cartes d'annoncer un temps clément, et ainsi $P(B)$ augmente, diminuant $P(A|B)$, au point que, si vous êtes convaincu qu'elle annonce de toute façon du beau temps, cartes ou pas cartes, $P(B|A) = 100\%$ et $P(B) = 100\%$. Ainsi de nouveau $P(A|B) = P(A)$ et votre situation n'a pas évolué.

Ainsi, les mathématiques des probabilités bayésiennes nous permettent de représenter le comportement humain qui régit toutes nos réflexions critiques et de les transposer au monde des machines qui, dès lors, sont capables d'exprimer des degrés de confiance relatifs à l'information qu'elles reçoivent. Cependant, si nous autres humains, modérons nos croyances *a priori* (ou les propos de la

¹ Notez que selon les probabilistes bayésiens, la notion même d'aléa n'a aucun lieu d'exister.

cartomancienne par exemple) de manière plus ou moins naturelle (tout cela dépend en vérité du tempérament de chacun), comment les machines peuvent-elles, de manière systématique, décider de valeurs *a priori*, telles que $P(A)$ et $P(B)$ dans l'exemple précédent ? Pourquoi croire ou ne pas croire en la cartomancienne, pourquoi supposer un *a priori* un demi sur l'ensoleillement du week end ? Cette pièce manquante à la théorie des probabilités bayésiennes fut longtemps l'objet de critiques par la communauté orthodoxe, avant de trouver une réponse appropriée et extraordinairement naturelle : le principe d'entropie maximale.

2.2.2. Le principe d'entropie maximale

Lorsque nous, ou les machines, sommes confrontés à des situations où peu d'information est disponible, par exemple quand vous devez estimer l'âge des auteurs de cet article sans même avoir accès à une photographie, la meilleure manière de raisonner pour établir des degrés de croyance fiables, est de (i) prendre en compte toute l'information qui nous est disponible (par exemple, le fait qu'aucun être humain sur Terre a plus de 130 ans, le fait que la plupart des chercheurs et enseignants chercheurs ont plus de 18 ans et moins de 60 ans etc.) et d'être le plus honnête possible concernant l'information qui vous fait défaut (par exemple, vous ne supposerez pas, sans le savoir, qu'un des auteurs a des cheveux blancs). Ainsi, très vraisemblablement, la courbe de degrés de confiance en l'âge des auteurs, formera un lobe centré vraisemblablement autour de 35 ans qui s'atténue très doucement autour de 30 à 40 ans puis s'atténuera bien plus fortement dans les zones inférieures à 20 ans et supérieures à 50 ans. La question est maintenant de savoir comment, de manière systématique, introduire cette notion d'honnêteté dans les machines ? Cette question se réduit finalement à la découverte d'un modèle mathématique témoignant de cette honnêteté. Ce modèle existe depuis maintenant 50 ans, après avoir été longtemps utilisé par les physiciens à leur insu, et après avoir été réanimé par Claude Shannon : l'entropie. Le principal instigateur de ce modèle décrivant l'honnêteté des machines fut Edwin Thompson Jaynes qui consacra la majeure partie de son existence à développer le principe d'entropie maximale.

Le cœur mathématique du principe d'entropie maximale repose sur quatre axiomes, que Jaynes préférerait dénommer *desiderata*, indispensables pour associer à toute variable ou à tout événement une distribution de probabilité tenant compte fidèlement de l'information *a priori* et qui soit la plus honnête possible concernant ce que la machine ignore. Appelons A la variable pour laquelle la machine cherche une distribution de probabilités honnête, I l'information que la machine connaît initialement, et $P(A|I)$ la probabilité de A quand I est pris en compte. Les quatre requêtes imposées à la probabilité $P(A|I)$ sont les suivantes :

1. Il existe une fonction capable de mesurer numériquement le niveau d'ignorance concernant A .

2. Si l'information I change légèrement (par exemple, on vous informe que les deux auteurs sont détenteurs du baccalauréat), alors le niveau d'ignorance sur A doit changer très légèrement également. En effet, la détention du baccalauréat par les auteurs étaient presque une évidence qui ne change pas significativement l'ignorance du lecteur. Cette hypothèse est dite hypothèse de continuité.
3. Si le nombre de possibilités augmente (par exemple, si on vous demande, en plus de l'âge, le mois de naissance des auteurs), alors l'hypothèse d'honnêteté demande que l'ignorance du lecteur sur A augmente également.
4. Enfin, si l'on regroupe les hypothèses en paquets isolés (par exemple, si l'on sépare l'échelle des âges par paquets de dix ans), l'ignorance du lecteur ne change pas s'il établit des probabilités sur l'échelle complète des âges ou s'il décide premièrement d'établir des probabilités sur chaque paquet de dix ans, puis à l'intérieur de chaque paquet d'établir d'autres probabilités. Aucune information n'étant ajoutée, l'ignorance sur A n'est pas changée par cette manipulation.

De manière très surprenante, il a été démontré par Shannon puis par Jaynes, qu'il n'existe qu'une seule fonction mathématique capable de tenir compte de ces quatre critères ; ainsi une seule fonction est réellement capable d'exprimer l'ignorance sur A : cette fonction s'appelle l'entropie de la distribution $P(A|I)$. Alors, le principe d'entropie maximale peut être clairement établi : la distribution $P(A|I)$ que l'hypothèse d'honnêteté demande d'associer à A est la distribution qui :

1. Est consistante avec le contenu de l'information *a priori* I . Dans notre exemple, cela veut dire en particulier que la probabilité que les auteurs aient 150 ans reste égale à 0.
2. Est celle qui, parmi ces fonctions qui respectent l'hypothèse ci-dessus, maximise l'ignorance du lecteur vis-à-vis de l'information qu'il ne possède pas.

Comme cette fonction d'ignorance a été démontrée comme étant mathématiquement modélisée par l'entropie, la distribution la plus honnête pour A sachant l'information initiale I est la distribution qui maximise l'entropie de A tout en respectant les contraintes imposées par I : c'est la distribution à entropie maximale.

Dès lors, les machines peuvent raisonner comme des êtres humains lorsqu'elles reçoivent une information, et ce en étant le plus honnête possible. Par exemple, lorsque votre téléphone voudra estimer la puissance du bruit électromagnétique ambiant (qui a été dénoté B dans les parties précédentes), il pourra prendre en compte toute l'information qu'il connaît issue du sondage des signaux reçus, mais devra également demeurer sceptique sur la fiabilité de ces données qui sont en général impropres à la découverte immédiate du niveau de bruit ambiant. Cependant, lorsque vraiment peu d'information est disponible, de sorte qu'aucune décision fiable ne peut être entreprise avec si peu de données, d'autres remèdes doivent être découverts. Par exemple, si vous êtes amené à reporter l'âge des auteurs à une tierce personne, vous ne pourrez en aucun

cas prendre une position stricte sur ces variables. Pour pallier à cette difficulté, l'être humain recourt le plus souvent au principe de requête ou d'inquisition d'information supplémentaire, dans l'espoir qu'une partie de cette nouvelle information sera cruciale et permettra une décision univoque par la suite. Typiquement, si vous envoyez un message électronique au premier auteur cité pour demander l'âge du second auteur cité, vous aurez une information importante qui ne sera peut-être pas exacte (si le premier auteur ne connaît pas l'âge exact du second) mais qui permettra de reporter cette valeur avec une erreur d'estimation minimale. Ce dispositif d'inquisition d'information n'existe pas non plus dans la majeure partie des appareils de communication à présent. Ou si elle existe, ce n'est que sous la forme d'un nouveau programme déclenché de manière automatiquement et qui a pour mission de requérir une information statique, en général non adaptée au problème spécifique de l'appareil.

2.2.3. Inquisition

Les modèles mathématiques régissant les idées de questionnement, requête d'information, et également de sélection de la « bonne question » à formuler sont plus variés et sujets à contestation (de la même manière que l'est toujours la théorie des probabilités bayésiennes). Nous allons nous attarder quelque peu sur le modèle initialement proposé par Richard Threlkeld Cox [5], particulièrement simple et intéressant sous de nombreux aspects. Selon Cox, une question peut être définie mathématiquement comme un ensemble (disons un récipient) contenant toutes les réponses possibles à cette question. Par exemple, à la question directe « fera-t-il beau ce week end ? », nous associons l'objet mathématique {oui il fera beau, non il ne fera pas beau} qui est un ensemble constitué de deux éléments répondant possiblement à cette question. L'intérêt de la machine, ou de l'Homme, lorsqu'elle ou il est en besoin d'information, n'est pas véritablement de choisir la « bonne » question à formuler mais plutôt de choisir la question qui porte un bon espoir d'apporter une réponse utile. De ce point de vue, le principe d'entropie maximale est assez similaire en nature au principe d'inquisition. Grâce à la distribution d'entropie maximale, nous sommes capables de donner une représentation honnête de nos degrés de confiance en la valeur de toute variable ou en la véracité de toute affirmation. L'intérêt de sélectionner la question qui apporte le meilleur espoir de réponse utile est similaire : connaissant la distribution de probabilité *a priori* des réponses possibles à chacune des questions, nous choisirons la question qui n'est pas la plus honnête (l'honnêteté d'une question n'ayant aucun intérêt particulier ici) mais la question qui représente le plus grand espoir de gain d'information utile.

Malheureusement, si le principe d'entropie maximale est simple en nature, il est parfois difficile à mettre en œuvre car demandeur de beaucoup de calculs. La théorie de l'inquisition repose sur un étage encore supérieur, en

ce sens que chaque réponse à chaque question doit être soumise à un test d'entropie maximale avant de pouvoir sélectionner la question dont les réponses sont les plus potentiellement porteuses d'information utile. Les applications pratiques de cette théorie sont par conséquent peu nombreuses à l'heure actuelle. Néanmoins, la théorie existe et ne demande que des moyens plus rapides, des capacités de mémoires plus importantes ou alternativement des méthodes de calculs approximatifs simples pour pouvoir émerger dans les appareils de communication de notre quotidien.

Considérons un exemple d'application simple qui concerne le mode de fonctionnement de nos téléphones portables et de nos cartes WiFi en particulier. Pour pouvoir décoder l'information qui arrive à son ou ses antenne(s), le téléphone portable doit déterminer les chemins de propagation des signaux, ou plus exactement doit déterminer l'atténuation dans l'air et le déphasage des ondes reçues. Pour cela, un moyen très simple existe : la station de base émettrice de l'information transmet des séquences de symboles (chaque symbole contient un certain nombre de bits d'information) dites pilotes, qui sont connues du téléphone. Ce dernier peut alors comparer les symboles reçus (et donc atténués et déphasés par le canal de propagation) à ceux qui ont été émis, et ainsi déterminer les atténuation et phase qui ont affecté chaque symbole. Seulement, lorsque la station de base envoie des signaux pilotes, elle gaspille des ressources (du temps et des fréquences) qui pourraient être alternativement utilisées pour transmettre de l'information utile. Si le téléphone pouvait, de lui même, requérir l'information manquante lui permettant de connaître le canal plus précisément, la station de base gaspillerait moins de ressource à transmettre des informations non nécessairement utiles. Tout ceci suit bien l'esprit de flexibilité avancée par Mitola : le téléphone est capable seul de comprendre son environnement et de solliciter plus d'information (i) si, et seulement si, il en a réellement besoin et (ii) sous la forme qui lui sera le plus profitable, en posant la question la plus utile.

2.3. Cas pratiques : découvrir son environnement et détecter des points d'accès

Le tout premier moyen d'augmenter les capacités cognitives de tout appareil de communication est de les rendre capable de détecter des sources de transmission avec lesquelles une communication peut être envisagée. En particulier, dans tout réseau cellulaire mobile, les communications s'effectuent entre le téléphone portable et une ou plusieurs stations de base environnantes. Lorsque le téléphone est mis en marche, il doit, et cela le plus vite possible, détecter toutes les stations de base environnantes. Classiquement, des algorithmes de détection de signal existent qui supposent la connaissance *a priori* de beaucoup d'informations telles que la puissance du bruit ambiant B . Or, le téléphone venant d'être activé, il lui est impossible de connaître ce niveau de bruit avec précision. Il va donc utili-

ser à nouveau un programme annexe qui lui retournera une valeur pour B en laquelle le téléphone devra faire confiance, faute de savoir faire preuve de scepticisme.

L'approche cognitive est fortement différente et pourtant simple car systématique. En effet, les méthodes *ad hoc* classiques nécessitent que les ingénieurs trouvent une solution propre répondant spécifiquement à chaque problème. La méthode cognitive consiste en une recette que chacun peut appliquer à tout problème, à la condition que toutes les informations *a priori* sur le problème en question sont proprement rassemblées. Cette méthodologie se base sur les trois piliers de la radio cognitive précédemment évoqués : les probabilités bayésiennes, le principe d'entropie maximale et les requêtes d'information. Avant de retourner à notre cas d'étude, voyons comment répondre à un problème quelconque en appliquant une démarche cognitive. Disons que nous cherchons à avoir accès à une information ou à une valeur A . La démarche d'apprentissage est la suivante :

1. Toute l'information *a priori* concernant A de près ou de loin est rassemblée : nous dénotons cette information initiale I .
2. A partir de cette information, nous utilisons le principe d'entropie maximale et déterminons ainsi la distribution *a priori* de notre confiance en chaque valeur de A . Dit autrement, nous établissons la distribution de probabilité $P(A|I)$ la plus honnête respectant les contraintes contenues dans I . Notons cette probabilité plus simplement sous la forme $P(A)$.
3. Une première évaluation de cette distribution peut alors être effectuée : si une valeur x pour A en particulier émerge fortement de l'ensemble des valeurs possibles, alors il n'est plus besoin d'aller plus loin, la valeur choisie sera $A = x$. Par contre, si aucune valeur particulière ne montre un niveau de confiance élevé, plus d'information est nécessaire. A ce moment, de deux choses l'une :
 - Soit une source d'information fournit naturellement un nouvel élément B (ce qui est le cas d'une station de base qui émet continûment des signaux dans l'environnement), alors cette information doit être utilisée (par honnêteté, toute information doit être prise en compte) comme indice supplémentaire à la découverte de A . Nous sommes en mesure, grâce à la loi de Bayes, de calculer $P(A|B)$, le degré de confiance en A lorsque B est connu.
 - Soit aucune source d'information n'apporte d'information utile à la détermination de A , auquel cas nous devons procéder à une requête d'information et donc être capable de sélectionner une question qui nous apportera probablement une réponse utile. La réponse à cette question est une nouvelle information, que l'on peut également dénoter B . La règle de Bayes s'applique identiquement au cas précédent dans cette situation et nous sommes en mesure de déterminer $P(A|B)$.
4. $P(A|B)$ nous étant donné, nous évaluons à nouveau la qualité de l'estimation de A apportée par cette nouvelle

information : si une valeur particulière x montre un degré fort de probabilité alors nous assignons $A = x$ et le processus s'achève. Si au contraire, et à nouveau, B n'a pas apporté d'information suffisante, le même processus d'obtention naturelle ou artificielle d'information supplémentaire peut être effectué jusqu'à ce qu'une valeur pour A montre un haut niveau de confiance.

Notez à nouveau que ce mécanisme artificiel se base sur une logique humaine de la quête d'information. Si on vous demande de déterminer si la Terre est plate ou sphérique, initialement aucune hypothèse ne prévaut, jusqu'à ce que vous ouvriez les yeux, observiez et commencez à imaginer que la Terre est en effet plus probablement sphérique que plate. Mais vous n'en êtes convaincu que lorsque d'autres personnes attestent ou que des ouvrages d'astronomie démontrent que cette hypothèse est en effet la plus probable. Néanmoins, toutes ces hypothèses font-elles que la Terre est réellement sphérique avec une certitude de 100 % ? En vérité, il reste une infime possibilité que, depuis votre tendre enfance, votre entourage a comploté pour vous amener à croire que la Terre est effectivement sphérique. Il est très rare, que ce soit dans le monde réel ou dans le monde des machines qu'une hypothèse puisse être retenue avec une probabilité de 100 %. Eliminer les hypothèses alternatives, peu viables, peut en effet avoir des conséquences néfastes : si votre téléphone mobile est programmé pour penser que jamais vous ne volerez dans un hélicoptère au-dessus d'une zone montagneuse, alors vous aurez toute les peines du monde à établir une communication fiable dans ce milieu. De la même manière, éliminer l'hypothèse, absurde pour certains, qu'un téléphone intelligent puisse voir le jour, amène à la possibilité dangereuse que jamais ce téléphone ne puisse apparaître, faute de croire en son existence potentielle.

Revenons maintenant au problème de détection de sources de puissance à l'allumage de votre téléphone portable et tentons d'appliquer, sans rentrer dans les détails, la méthodologie proposée ci-dessus. Pour déterminer avec efficacité une source de puissance à l'aide d'un dispositif cellulaire, il faut idéalement connaître :

- la puissance du bruit ambiant, dû à toute source électromagnétique d'interférence,
- les conditions de propagation (hypothétique) des signaux, i.e. le canal à travers lequel l'information se propage.

A l'aide de ces deux éléments, il est en général facile de déterminer la présence ou l'absence d'une source de signal et même d'en déterminer la puissance. Disons que nous souhaitons juste établir la présence ou l'absence de cette source de signal. La manière de procéder est la suivante :

1. A partir de l'information *a priori* qui nous est disponible (par exemple, si nous avons affaire à une communication WiFi, nous savons qu'avec de grandes chances nous sommes dans un milieu intérieur avec de nombreux réflecteurs avoisinants : tables, chaises, murs etc.), nous déterminons à la fois une distribution de degrés de confiance *a priori* sur l'environnement de propagation et sur la puissance du bruit ambiant. Dans

le cas de la téléphonie mobile par exemple, il est à supposer qu'avec très peu de chance, nous avons la station émettrice de puissance en ligne de mire. Ceci affecte fortement notre *a priori* sur le modèle du canal. Nous devons d'autre part intégrer l'information que plus de 99 % du territoire est couvert par un réseau de communication mobile.

2. Cette information est bien sûr quasi-inutile en soit puisque rien ne prouve à cette étape qu'une source de puissance existe ou non. Ainsi le récepteur du signal issu de l'hypothétique station de base va analyser les signaux à l'entrée de ses antennes (qui peuvent bien sûr n'être que des contributions de bruit ambiant). Ces signaux apportent une nouvelle information au récepteur qui sera intégrée à l'information initiale pour former une nouvelle distribution de la probabilité de présence d'une station de base émettrice.
3. Si ces signaux ne constituent pas de preuves suffisantes à une décision sans risque quant à la présence d'une source de puissance, le téléphone enverra une requête spécifique à tout terminal de communication alentour pour obtenir plus d'information concernant directement la présence d'un émetteur ou alternativement quelque information supplémentaire concernant le canal typique de propagation des ondes dans l'environnement qui l'entoure.
4. Si une preuve suffisante de la présence ou l'absence de cet émetteur de puissance est disponible, l'algorithme s'arrête et une décision peut être prise. Dans le cas contraire, de nouvelles questions ou l'analyse de nouveaux signaux reçus par les antennes réceptrices seront envisagées.

Malheureusement, entre philosophie, théorie et pratique, certains vides doivent être comblés, à cause de difficultés d'ordres technique, mathématique, théorique ou même politique. Un bref état des lieux du monde des communications sans fil et de leur avenir est dressé dans la dernière partie de cet article.

3. Les défis de demain

La première difficulté qu'un chercheur en télécommunications modernes rencontre lorsqu'il découvre une nouvelle technologie telle que les systèmes multi-antennes ou l'intégration de capacités cognitives, est de prime abord la justification d'un besoin : devons-nous vraiment rendre nos téléphones portables si intelligents ? La radio flexible, qui permet à un appareil unique de se connecter à différents services mobiles, semble être très accessible et peu coûteuse dans ce sens où elle ne nécessite aucun supplément d'infrastructure dans le réseau existant et surtout aucune modification radicale des protocoles de communications mis en place : son but essentiel est de minimiser le gaspillage engendré par les réseaux établis, sans modification radicale de leur nature. Au contraire, implémenter de l'intelligence dans les dispositifs cellulaires confère à ces derniers la possibilité de prendre des décisions autonomes, de solliciter

d'autres appareils et même peut-être de créer naturellement des réseaux « sociaux » d'entités de communication. Cette hypothèse vient en rupture totale des protocoles actuels qui ne permettent qu'un degré de liberté très limité aux appareils mobiles, toute l'intelligence du réseau étant concentrée dans l'infrastructure fixe (i.e. les stations de base ou autres points d'accès). Passer d'une intelligence centralisée concentrée dans l'infrastructure fixe à une intelligence décentralisée (telle que dans l'Internet moderne) est un défi important des télécommunications du futur.

3.1. Une unification des appareils de communication ?

La possibilité de l'unification future des systèmes de communication est le leitmotiv des chercheurs en radio logicielle. En effet, notre téléphone portable pourrait accomplir la plupart des tâches de communications utilisées au quotidien : appareil de téléphonie mobile, système d'accès à l'Internet, utilitaire de paiement sans fil se substituant à la carte de crédit, système de stockage de données sans fil remplaçant la clé USB etc... La première question est de savoir si cette solution est politiquement viable. En vérité, la réponse ne peut être donnée que par un choix de la société. D'après les auteurs, il ne semble exister aucun frein à la volonté de chacun de rassembler toutes nos données dans un dispositif unique, en lieu et place de multiples appareils encombrant nos poches. Le problème de la substitution du téléphone portable à la carte de crédit n'est en soit critique que dans le contexte de sécurité absolue qui est demandée au transfert sans fil d'informations bancaires. Ce problème est techniquement résolu, comme l'a démontré à nouveau Claude Shannon un demi-siècle plus tôt.

Si tous nos besoins de communications sans fil peuvent être remplis par un unique dispositif, peut-être alors pouvons-nous à l'avenir rassembler tous les protocoles de communications, qui utilisent chacun une bande de fréquences spécifique, en un seul protocole large bande, extrêmement flexible ? De cette manière, idéalement, tout service utilisera la même bande de fréquences qui sera suffisamment large pour ne jamais pouvoir être exploitée à son maximum. Cependant, si le téléphone comme module de communication unique est une solution viable et désirable, une bande passante unique pour servir tout type de communication est beaucoup moins profitable à chaque service considéré individuellement. En effet, la physique des propagations de signaux et les différents débits demandés par chacun de ces services viennent ici nous prévenir qu'une rationalisation des services de communications ne peut se faire par une unification triviale des bandes de fréquences. Les deux problèmes majeurs sont les suivants :

- Les ondes électromagnétiques s'atténuent fortement lorsqu'elles se propagent à l'air libre et cela d'autant plus que leur fréquence est élevée. Ainsi, des ondes de très hautes fréquences sont impropres aux réseaux cel-

lulaires dont la zone de couverture peut atteindre des rayons autour des stations émettrices de plusieurs kilomètres : la dissipation importante d'énergie dans l'atmosphère rend ici les communications sans fil à longue portée très inefficaces. Au contraire, les ondes de basses fréquences se propagent beaucoup plus facilement et peuvent atteindre des portées significatives pour une perte d'énergie minimale. Un exemple de vie quotidienne illustrant bien cette idée est celui de la propagation des ondes acoustiques qui, à basse fréquence, traversent aisément les murs, tandis que les ondes à fréquence élevée sont bien plus atténuées : ce qui explique pourquoi, de la musique écoutée par vos voisins, vous n'entendez qu'un lancinant « boum, boum » privé de sons de voix. Notez que les pertes de puissance lors de la traversée des ondes électromagnétiques dans l'atmosphère sont souvent modélisées par un paramètre d'atténuation allant de 2, i.e. la puissance de l'onde diminue proportionnellement au carré de la distance parcourue, à 5, i.e. l'onde perd en intensité de manière proportionnelle à la distance parcourue élevée à la puissance 5. Ainsi, après un kilomètre de propagation, lorsque l'onde de basse fréquence a perdu en intensité d'un facteur de l'ordre de 1 000 000, l'onde de haute fréquence a, elle, diminué en puissance d'un ordre de 1 000 000 000 000 000, ce qui rend son intensité à la réception illusoire pour des communications fiables (à moins bien sûr qu'une importante bande passante soit utilisée).

- Chaque application spécifique de communications sans fil demande des débits propres qui ne peuvent être atteints efficacement que pour des distances de transmission limitées. Typiquement, la prochaine génération de nos clés USB, l'USB 3.0, seront sans fil et permettront une vitesse de transfert dix fois supérieure à nos clés USB 2.0. De tels débits de transfert requièrent une bande passante très large. Les fréquences sollicitées appartiennent donc à des gammes allant de 3 GHz à 10 GHz, pour lesquelles les puissances des ondes irradiées se dissipent très rapidement dans l'air. Les débits espérés pourront atteindre jusqu'à 480 Mb/s à un mètre de votre ordinateur, et seulement 110 Mb/s à une distance de dix mètres. Au contraire, les communications cellulaires ne requièrent pas de tels débits, même si la possibilité d'accéder à l'Internet à haut débit par les récents réseaux 3 G sollicitent désormais de forts débits. Les distances acceptables de couverture diminuent donc avec l'introduction de ces nouvelles demandes de services mais demeurent de l'ordre de quelques centaines de mètres à plusieurs kilomètres.

A chaque service sans fil est ainsi attaché une heuristique entre taille des zones de couverture et demande de débits minimum. L'unification de tous les services, capables de « piocher » dans une bande unique leurs besoins en bande passante, n'est donc pas envisageable. Idéalement, le spectre fréquentiel doit être divisé en plusieurs secteurs, chaque secteur correspondant à un cas d'utilisation distance-débit spécifique. Typiquement, trois scénarios se découpent suivant leur distance de couverture :

- Les services cellulaires extérieurs, acceptant des distances de l'ordre de quelques centaines de mètres à plusieurs kilomètres, et des débits actuellement de l'ordre de quelques mégabits par seconde dans la pratique.
- Les services cellulaires à faible portée tels que les solutions WiFi à l'intérieur d'immeubles ou maisons, dont la couverture est de l'ordre de la dizaine de mètres et dont les débits peuvent atteindre plusieurs mégabits par seconde.
- Les services sans fil de proximité, tels que mémoires USB sans fil, qui n'acceptent qu'une zone de couverture de quelques mètres mais des débits atteignant jusqu'à plusieurs centaines de mégabits par seconde.

Intégrer tous ces services sans fil dans un contexte unique de radio flexible apparaît comme un défi accessible, à nouveau grâce à l'existence de l'infrastructure et des services au-dessus desquels la radio flexible constituera une nouvelle couche de régulation. Qu'en est-il cependant de la radio cognitive, qui elle demande une restructuration complète des infrastructures mais également de la philosophie sous-jacente au monde des communications sans fil ?

3.2. Un téléphone social qui évolue ?

Autoriser les appareils mobiles à réfléchir et prendre des décisions de manière autonome permettrait de transiter d'une infrastructure fixe dans laquelle les terminaux mobiles sont soumis aux décisions du réseau à une infrastructure dynamique dans laquelle les terminaux mobiles se trouvent être des agents intelligents permettant d'accroître les capacités de communication du réseau tout entier. Un exemple simple consiste en l'utilisation de téléphones portables comme relais de l'information transmise par d'autres appareils se trouvant dans des zones de couverture limitée. Dans un environnement urbain dense en appareils mobiles, certaines stations fixes utilisées comme seuls transmetteurs à l'heure actuelle pourraient même être éliminées du réseau, ce dernier devenant une toile de terminaux mobiles autonomes. Chaque terminal étant doté de capacités cognitives, ce réseau de terminaux forme une véritable société d'individus capables de raisonnement critique dont le lien social est une nécessité collective d'augmenter la qualité de service de chacun. Une nouvelle difficulté, qui n'a pas été introduite dans les sections précédentes semble apparaître ici : si l'individu de cette société est intelligent, est-ce que l'ensemble des terminaux peut se comporter collectivement de manière intelligente ? Et si oui, comment ? A nouveau, des modèles mathématiques existent qui rendent compte des interactions entre individus d'une société qui ont (i) soit des motivations individualistes (par exemple, augmenter leur propre débit), (ii) soit des motivations collectives (par exemple, augmenter la somme totale des débits individuels). La théorie mathématique en question est connue sous le nom de théorie des jeux, développée il y a cinquante ans par John Von Neumann et John Nash.

D'un point de vue individuel, chaque entité de ce réseau cognitif est capable d'apprendre et de prendre des

décisions en respectant une contrainte d'honnêteté. Ainsi, toujours par honnêteté, toute nouvelle information sur l'environnement qui l'entoure doit être intégrée dans la mémoire de cette entité ; cela permet à cette dernière de devenir de plus en plus consciente du monde qui l'entoure. Ainsi, lors de la première utilisation d'un appareil cognitif, celui-ci devra tout apprendre et sera donc moins efficace que quelques années plus tard lorsqu'il aura appris et compris le type d'environnement auquel il est le plus fréquemment confronté. En d'autres mots, votre téléphone portable serait capable de comprendre que le matin son environnement évolue pour, avec forte probabilité devenir l'environnement de votre lieu de travail alors que le soir son environnement évolue à nouveau pour devenir celui de votre résidence. Avec le temps, ce dernier deviendra peut-être même totalement personnalisé à vos besoins qu'il aura pris la peine d'analyser et comprendre.

4. Conclusion

Nous sommes peut-être aujourd'hui à l'aube d'une toute nouvelle ère dans le domaine des télécommunications. Soixante ans après l'extraordinaire contribution de Shannon qui a bouleversé une première fois le monde des télécommunications modernes, un nouveau paradigme doit être mis en place pour répondre au problème de rarefaction des ressources fréquentielles pour les systèmes de communications futurs. La radio flexible est envisagée comme une solution à court terme pour limiter les gaspillages importants des dispositifs de communication actuels, en réutilisant efficacement les ressources des protocoles préétablis. Cependant cette heuristique n'est une fois encore qu'un remède à l'impact temporel vraisemblablement limité et une solution totalement révolutionnaire devra un jour être mise en place pour permettre une exploitation intelligente et optimale des ressources. Cette solution peut être envisagée si tout appareil participe à la réussite du réseau de communications, qu'il soit un nœud fixe du réseau ou un terminal mobile. Pour ce faire, les entités du réseau devront agir individuellement et collectivement de manière intelligente, devenant ainsi le cœur de la prochaine évolution des communications sans fil, la radio cognitive.

Références

- [1] C. SHANNON, "A Mathematical Theory of Communications", Bell System Technical Journal, vol. 27, no. 7, pp. 379-423, 1948.
- [2] I.E. TELATAR, "Capacity of Multi-Antenna Gaussian Channels", Bell Laboratories, Lucent Technologies, Technical Memorandum, October 2005.
- [3] J. MITOLA, G. Q. MAGUIRE, "Cognitive radio: Making Software Radios More Personal", Personal Communications, IEEE [voir également IEEE Wireless Communications] 6(4), 13-18, 1999.
- [4] E. T. JAYNES, "Probability Theory: The Logic of Science", Cambridge University Press, 2003.
- [5] R. T. COX, "Of Inference and Inquiry, An Essay in Inductive Logic", The Maximum Entropy Formalism, pp. 119-167, edited by Raphael D. Levine & Myron Tribus ; MIT Press, USA, 1979.

Les auteurs

Romain Couillet est né à Abbeville en France. Il a reçu son Master of Science en systèmes de télécommunications (STIC) à l'institut Eurecom en 2007. Il a reçu son diplôme d'ingénieur (spécialisation en communications mobiles) à l'Ecole Nationale Supérieure des Télécommunications (Telecom ParisTech) en 2007. En septembre 2007, il rejoint NXP, fondé par Philips, qui deviendra plus tard ST-Ericsson. Chez ST-Ericsson, il travaille en tant qu'ingénieur en développement d'algorithmes sur le projet Long Term Evolution Advanced (LTE-A). En parallèle de sa position chez ST-Ericsson, il poursuit une thèse à Supélec. Ses sujets de recherche incluent les communications sans fil, les techniques de détection multi-utilisateurs et multi-antennes, la radio cognitive et la théorie des matrices aléatoires. Romain Couillet a reçu la récompense de la meilleure publication par un étudiant à la conférence Valuetools (Athènes, 2008).

Mérouane Debbah est né à Madrid, en Espagne. Ancien élève de l'École Normale Supérieure de Cachan, il a reçu son doctorat en 2002. De 1999 à 2002, il a travaillé pour Motorola Labs sur les réseaux locaux sans fil. En 2002, il a rejoint le Centre de recherches en télécommunications de Vienne (ftw. À Vienne en Autriche) pour travailler sur la modélisation de canaux multi-antennes. De 2003 à 2007, il fut membre du département de communications mobiles de l'institut Eurecom, Sophia-Antipolis (France) en tant que maître de conférence. Il est à présent professeur à Supélec, titulaire de la chaire Alcatel-Lucent en radio flexible. Ses thèmes de recherche concernent la théorie de l'information, le traitement du signal et les télécommunications sans fil. Mérouane Debbah a reçu le "Mario Boella" prize award en 2005, le 2007 General Symposium IEEE GlobeCom best paper award, le Wi-Opt 2009 best paper award ainsi que le Valuetools 2007, Valuetools 2008 et CrownCom2009 best student papers award. Il a été nommé Fellow du WWRF en 2009.