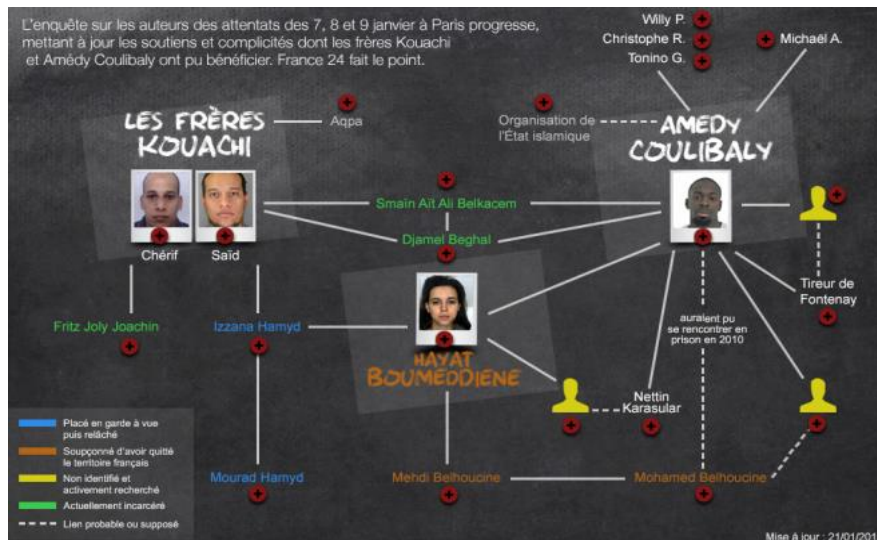


L'information, en n dimensions

«We're still waiting for a major breakthrough in understanding [non-linear systems], and it could be a long time in coming. I think we may be missing the conceptual equivalent of calculus, a way of seeing the consequences of the myriad interactions that define a complex system. It could even be that this ultracalculus, if it were handed to us, would be forever beyond human comprehension. **We just don't know**», S. Strogatz, *Sync - The Emerging Science of Spontaneous Order*, 2003.

Comme beaucoup, je suis resté rivé à mon poste de télévision en ce mois de janvier 2015. Je n'ai pas pu m'en détacher lors de l'attentat contre *Charlie Hebdo* et puis lors de la prise d'otages de *l'Hypercashier*, sidéré par les proportions que prenaient ces événements sur les médias. A vrai dire j'étais plutôt *englouti* par les flots d'informations qui écrivaient l'histoire devant moi, en direct. Ce n'est que petit à petit que je comprenais les implications de ces événements à plus longs termes, qu'elles soient culturelles, politiques, légales ou...techniques. Passée la période d'émotion, se profilaient les questions éthiques et politiques de la surveillance des réseaux d'activistes islamistes, de l'accès pour les services de renseignement à des données inédites, à des sources multiples et de leurs croisements, sans oublier les aspects juridiques qui allaient émerger. Pour le moment, j'étais pris dans un flot d'images et de commentaires d'où émergea subitement l'envie de noter sur un plan les faits et les personnages, un instrument de compréhension globale où se relieraient de façon logique (ou causale) les événements et que l'on pourrait inscrire à mesure qu'ils apparaissaient. Vous savez ce genre de grand tableau synthétique qui équipe probablement les rédactions de nos médias ou les *war-room* et sur lequel s'accumulent les informations en temps réel. Quelques médias, comme *France24* (ci-dessous), en diffusèrent d'ailleurs quelques-uns où se trouvent épinglés les frères Kouachi et Amedy Coulibaly, les acteurs principaux autour desquels s'agrègent d'autres éléments qui contribuent, justement, à *faire tableau*. Mon envie se transforma rapidement en projet. Après tout, j'avais de quoi bricoler mon propre schéma car j'avais tout sous la main et autour de moi: l'équipement qui m'accompagnait dans l'exploration du web. Et ce petit détour du côté des graphes de réseaux terroristes ne fut pas sans conséquence théorique et méthodologique pour la quête qui m'animait depuis déjà quelques années autour de la forme du web...



Je n'étais pas spécialiste du renseignement ou des questions de sécurité mais je savais, cependant, que pour rôder une méthodologie, affiner des instruments, produire des hypothèses ou même un modèle général (ici d'un type particulier, celui d'une d'organisation sociale) il faut d'abord disposer d'une bibliothèque de cas réels. C'est alors que je suis revenu sur le travail de Valdis Krebs ⁽¹⁾, l'un des premiers à avoir démontré tout l'intérêt du *social network analysis* (SNA) pour les questions de sécurité et de description des organisations terroristes.

Dans les jours qui ont immédiatement suivi les attentats du 11 septembre 2001, Valdis Krebs a commencé à reconstituer sous forme de graphe le puzzle des cellules qui ont opéré au nom d'*AlQaeda* aux Etats-Unis. Comme à chaque fois, tout commence avec une table de données (type fichier *Excel*) dans laquelle il va accumuler, rivé aux chaînes de télévision et au web, différentes informations en agrégeant les sources comme le ferait un détective. Dans ce travail remarquable de précision et de patience, sa compréhension du phénomène se construit par élargissement progressif de son horizon d'observabilité, au fur et à mesure des versions successives de sa "carte". Chaque nouvelle information vient nourrir la perspective: V. Krebs ajoute, en particulier, de *nouveaux liens* entre les 19 membres des quatre cellules qui concernent les points communs de leurs différentes biographies (certains sont de la même famille, ont fréquenté les mêmes lieux, vécu dans les mêmes camps d'entraînement en Afghanistan, sont des amis d'enfance ou ont côtoyé les mêmes imams). Il ajoute aussi différentes informations juridiques, commerciales ou administratives qui lui permettent alors de comprendre le lent travail préparatoire des attentats (location commune de voiture, logements communs, voyage à deux ou trois membres...sur une période de plusieurs mois ou années). Une partie significative de son travail a consisté à intégrer aussi de nouveaux noeuds dans la carte, tous ces acteurs qui ont assuré la logistique des opérations et/ou qui ont croisé les "kamikazes" au cours de leurs vie. Le principe cumulatif adopté par V. Krebs le conduit donc à intégrer de nouveaux noeuds dans le système en fonction du *type de liens* auquel il s'attache.

¹ <http://www.orgnet.com/hijackers.html>, http://insna.org/PDF/Connections/v24/2001_1-3-7.pdf

Il s'agit là d'un principe d'expansion contrôlée du réseau sur des critères très factuels et qualitatifs, un travail d'orfèvre.

La carte finale que livre V. Krebs comporte ainsi plus de 60 membres (dont certains sont seulement soupçonnés d'avoir participé de façon indirecte aux quatre attentats, d'autres ont été condamnés puis relâchés aujourd'hui et d'autres encore à l'identité invérifiable). La plupart des quarante nouveaux noeuds de la carte ont été ajoutés à partir d'un double critère: la nature de leurs liens avec les 18 membres des 4 cellules opérationnelles et le fait qu'ils ont été en contact (pour la plupart) avec au moins deux terroristes (presque sur le même principe que la co-citation en bibliométrie).

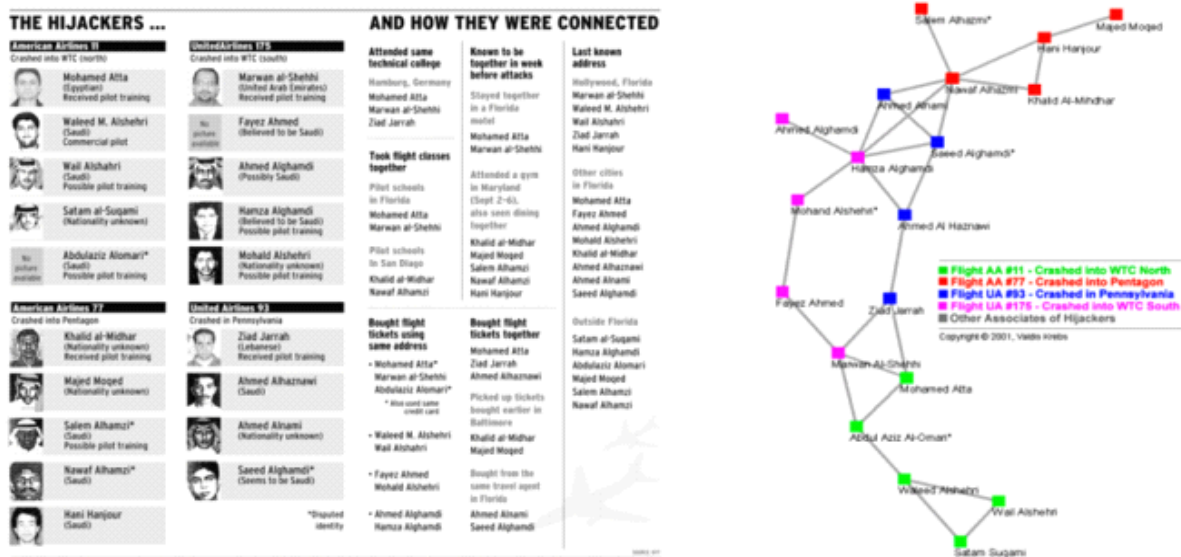


Figure 1. Early Hijacker Matrix

Je me suis donc mis à construire mon propre jeu de données sur les événements du 11 septembre pour rejouer l'expérimentation, comme on le ferait en sciences expérimentales pour valider une «manip» et ses résultats. Les informations ont été ajoutées petit à petit dans la table de données disponible sous Gephi (une sorte de feuille Excel directement visualisable sous forme de graphe).

Nodes	Id	Label	TYPES	BIO	Eccentricity	Closeness Cent...	Betweenne...	NAT
● Mohamed Atta	1	Mohamed Atta	1	http://fr.wikipedia.org/wiki/Mohammed_Atta	3	1,806	162,203	Egypte
● Marwan Al-Sheh	2	Marwan Al-Shehhi	2	http://fr.wikipedia.org/wiki/Marwan_al-Shehhi	3	1,75	88,906	EAU
● Hani Hanjour	3	Hani Hanjour	4	http://en.wikipedia.org/wiki/Hani_Hanjour	3	1,889	138,169	Arabie Saoudite
● Essid Sami Ben	4	Essid Sami Ben Khemais	0	http://en.wikipedia.org/wiki/Sami_Essid	4	2,778	0	Tunisie
● Nawaf Alhazmi	5	Nawaf Alhazmi	4	http://fr.wikipedia.org/wiki/Nawaf_al-Hazmi	4	2,111	69,543	Arabie Saoudite
● Ramzi Bin al-Sh	6	Ramzi Bin al-Shibh	0	http://fr.wikipedia.org/wiki/Ramzi_Bin_al-Shibh	4	2,417	40,417	Yemen
● Ziad Jarrah	7	Ziad Jarrah	3	http://fr.wikipedia.org/wiki/Ziad_Jarrah	3	1,806	108,622	Liban
● Abdul Aziz Al-O	8	Abdul Aziz Al-Omari*	1	http://en.wikipedia.org/wiki/Abdulaziz_al-Omari	3	1,722	87,3	Arabie Saoudite
● Fayed Ahmed	10	Fayed Ahmed	2	http://fr.wikipedia.org/wiki/Fayed_Benhammed	4	2,056	17,642	EAU
● Salem Alhazmi*	11	Salem Alhazmi*	4	http://fr.wikipedia.org/wiki/Salem_al-Hazmi	4	2,389	8,986	Arabie Saoudite
● Satam Suqami	12	Satam Suqami	1	http://fr.wikipedia.org/wiki/Satam_al-Suqami	4	2,167	24,933	Arabie Saoudite
● Zacarias Moussaoui	13	Zacarias Moussaoui	0	http://fr.wikipedia.org/wiki/Zacarias_Moussaoui	5	3,389	0	France
● Hamza Alghamdi	14	Hamza Alghamdi	2	http://fr.wikipedia.org/wiki/Hamza_al-Ghamdi	4	2,083	17,472	Arabie Saoudite
● Said Bahaj	15	Said Bahaj	0	http://en.wikipedia.org/wiki/Said_Bahaj	4	2,056	54,854	Allemagne
● Khalid Al-Mhdhar	16	Khalid Al-Mhdhar	4	http://fr.wikipedia.org/wiki/Khalid_al-Mhdhar	4	2,639	0	Arabie Saoudite
● Saeed Alghamdi*	17	Saeed Alghamdi*	3	http://fr.wikipedia.org/wiki/Saeed_al-Ghamdi	3	2,167	29,662	Arabie Saoudite
● Wail Alshehri	19	Wail Alshehri	1	http://fr.wikipedia.org/wiki/Wail_al-Shehri	4	2,25	0	Arabie Saoudite
● Waleed Alshehri	20	Waleed Alshehri	1	http://fr.wikipedia.org/wiki/Waleed_al-Shehri	4	2,222	1,672	Arabie Saoudite
● Agus Budiman	22	Agus Budiman	0	http://911review.org/Sept11Wiki/Budiman,Agus.shtml	4	2,194	7,028	Indonesie
● Ahmed Alghamdi	23	Ahmed Alghamdi	2	http://fr.wikipedia.org/wiki/Ahmed_al-Ghamdi	4	2,028	24,244	Arabie Saoudite
● Zakariya Essab	25	Zakariya Essabar	0	http://en.wikipedia.org/wiki/Zakariya_Essabar	4	2,528	0,783	Maroc
● Ahmed Al Hazni	26	Ahmed Al Haznawi	3	http://en.wikipedia.org/wiki/Ahmed_al-Haznawi	3	2,056	14,51	Arabie Saoudite
● Imad Eddin Bar	27	Imad Eddin Barakat Yarkas	0	http://en.wikipedia.org/wiki/Abu_Dahdah	4	2,528	3,672	Espagne

J'ai ainsi produit un graphe de 37 acteurs dont 19 ont constitué les 4 groupes kamikaze. J'ai repris une partie des informations à l'étude de V. Krebs mais j'en ai rajouté d'autres empruntées aux biographies, parfois très détaillées, des terroristes sur wikipedia (2). Quelques-uns parmi eux ont une identité supposée, d'autres sont toujours en prison aujourd'hui. C'est pour cela que je n'ai retenu que 18 acteurs associés aux équipes opérationnelles des attentats, selon différents aspects dont le rôle joué par certains dans le soutien logistique, d'autres parce que sont des compagnons sur les différents terrains d'action du terrorisme islamiste (Afghanistan, Pakistan, Irak). Il est important de noter que, quand on étend le graphe des relations au-delà de ces 18 acteurs associés, on arrive rapidement à des réseaux issus d'*Al Qaeda* ou de l'*Etat Islamique* dans lesquels ont gravité les frères Kouachi et A. Koulibaly. En deux « clics » ou en deux accointances sociales de distance, on peut donc construire une vaste carte où seraient présents une grande partie des réseaux islamistes actifs aujourd'hui. A une condition cependant: faire porter sur le processus d'expansion du graphe ou de la carte des critères qualitatifs rigoureux (séjours communs dans un camp d'entraînement en Afghanistan, fréquentation de la même mosquée radicale...). On imagine aisément à quelle situation absurde mènerait une expansion générale où pourrait figurer au final une bonne partie de l'humanité! Après tout, nous sommes tous liés les uns aux autres, directement ou indirectement, d'une façon ou d'une autre, ne serait-ce que sur *Facebook* ou *Gmail*. Le danger du *6 degrees* et des proximités structurales (3).

Je dois avouer que j'entrais là dans un univers qui s'est vite révélé différent de celui de l'exploration du web et de ses masses de données. Dans le travail que j'entamais et où une grande attention est portée à la vérification des informations pour un graphe comportant seulement 34 acteurs, on prend rapidement conscience en le construisant de l'amoncellement des difficultés méthodologiques. Le travail qualitatif de l'enquête (la mienne étant fictive) repose en grande partie sur l'analyse de la richesse et de la diversité des connexions de type social. Un autre aspect problématique de la démarche est d'assumer, pas à pas, un horizon d'observabilité du phénomène de plus en plus élargi et complexe. La prise de décision de faire entre ou pas un acteur ou un lien dans le graphe relève d'un exercice difficile entre le trop grand nombre d'informations potentiellement pertinentes et la nécessité de les qualifier précisément. L'une des grandes différences avec les univers du *data mining* et de son ingénierie parfois massive est d'assumer une forme d'incomplétude générée par la nature des relations causales des univers sociaux qui sont constitués d'événements, de faits ou d'aléas historiques. Contrairement au lien hypertexte qui indique de façon univoque et universelle l'existence d'une connexion ou d'une relation mais ne contient pas, en quelques sortes, son explication. Il faut de surcroît ajouter à l'exercice cette difficulté qui consiste à opérer dans un univers temporaire et aux frontières floues: certains éléments (nœuds ou liens) intégrés peuvent se révéler par la suite inutiles et/ou injustifiés ou d'autres encore peuvent être utiles à un moment et inutiles à un autre. Quand il s'agit de d'acteurs ou de groupes sociaux, on comprend les implications éthiques ou morales de la cartographie ou, dans des techniques plus quantitatives, la production de *false positive* (individus innocents répondant cependant aux critères d'identification des algorithmes en *big data*). Pour couronner le tout, il apparaissait que les éléments cartographiés étaient reliés entre eux par des liens de différentes

² L'un des sites les plus utiles aura été sans conteste le site du programme *S.T.A.R.T.* (National Consortium for the Study of Terrorism and Responses to Terrorism), une référence incomparable à ma connaissance. Parmi les projets du programme START, figure par exemple un projet de base de données de profils du radicalisme et issues de l'analyse de milliers de sources d'information (Projet PIRUS - Profiles in Individual Radicalization in the United States).

³ Cf. Chapitre 2, *Questions de distance*.

durées (longue, courte, occasionnelle) et de différentes intensités (connexion occasionnelle dans un meeting politique ou religieux, connexion logistique quand deux acteurs habitent au même endroit près des lieux des attentats, connexion forte et ancienne quand on appartient à la même famille ou quand on a passé des mois dans un camp d'entraînement en Afghanistan).

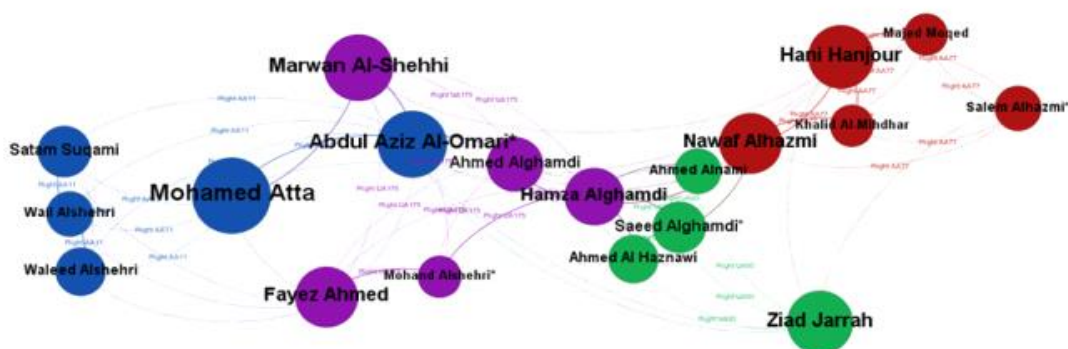
Pour toutes ces raisons, il n'est guère envisageable de carte finale ou définitive en matière de sécurité, sauf à la construire de façon rétrospective et/ou en ayant connaissance de toutes les dimensions d'une activité criminelle. L'intérêt des cartes de réseaux criminels réside plutôt dans la dynamique de leur construction, à mesure qu'adviennent les éléments de connaissance qui permettent de comprendre un contexte. Elles sont proches des instruments de *mind-mapping* ou de cartographie sémantique et sont très contrôlées qualitativement. Ce sont de véritables instruments de *dynamic knowledge discovery* mais d'une nature différente de celles qui accompagnaient l'exploration des structures du web et de ses *patterns* statistiques massifs. La construction de ma carte générale a débuté par l'intégration des 19 kamikazes du 11 septembre 2001, au moment où les membres des 4 cellules ont été identifiés (V. Krebs a commencé sa carte à partir des deux premiers suspects identifiés).



A ce stade ce sont quatre clusters indépendants (plus haut), un pour chacun des avions (vol AA11 en rouge s'écrase sur la tour 1 du *World Trade Center*, vol AA77 en violet s'écrase sur la *Pentagone*, vol UA175 s'écrase sur la tour 2 du *World Trade Center*, vol UA93 s'écrase en Pennsylvanie alors qu'il se dirigeait vers la *Capitole*). Dans ce type de travail, les attributs associés aux liens sont d'une grande importance: ils permettent de faire apparaître des labels sur les liens mais, surtout, de qualifier et de pondérer de nombreuses familles potentielles de connexions. La question d'un travail tant qualitatif que quantitatif sur les connexions est un aspect central de la cartographie des réseaux criminels. Pour commencer, je n'ai donc que les liens "logistiques" assurant la présence simultanée des membres du réseau dans les quatre vols différents. Ces quatre clusters sont isolés et ne constituent pas, à proprement parler, un réseau cartographiable: l'adjonction de liens supplémentaires permet, dans un second temps, de faire apparaître un graphe connexe où chaque acteur se trouve lié aux autres via la découverte de différents types de relation. Il aura fallu des mois aux enquêteurs pour les identifier à partir de très nombreuses sources d'information que V. Krebs résume sous forme d'un tableau.

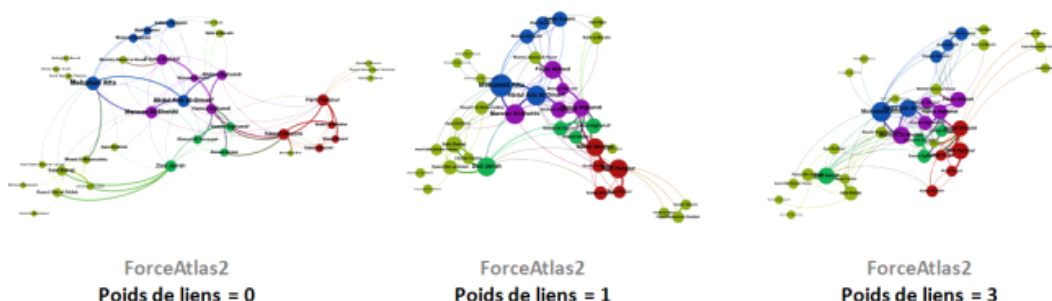
Relationship/Network	Data Sources
1. Trust	Prior contacts in family, neighborhood, school, military, club or organization. Public and court records. Data may only be available in suspect's native country.
2. Task	Logs and records of phone calls, electronic mail, chat rooms, instant messages, Web site visits. Travel records. Human intelligence: observation of meetings and attendance at common events.
3. Money & Resources	Bank account and money transfer records. Pattern and location of credit card use. Prior court records. Human intelligence: observation of visits to alternate banking resources such as Hawala.
4. Strategy & Goals	Web sites. Videos and encrypted disks delivered by courier. Travel records. Human intelligence: observation of meetings and attendance at common events.

C'est le principe des *layers* de liens, autrement dit des *couches* de liens en fonction de leur nature en termes de sources. Etant donné mes instruments et mes connaissances, j'ai réduit le principe à seulement deux familles (liens *forts* représentant des liens de famille, d'amitié, de scolarité commune ou de présence conjointe dans des camps d'entraînement, notamment en Afghanistan; liens *faibles* liés à des questions qui m'ont apparu plutôt logistique et opérationnelles comme les cours communs de pilotage, les locations de voitures et de logement, les flux bancaires). Comme le fait Valdis Krebs, j'ai attribué des forces différentes pour chacune des deux familles de liens (liens faibles = poids 1, liens forts = poids 5), selon une répartition assez arbitraire à priori. Les formules de traduction peuvent être nombreuses, tout comme la grammaire générale qui préside à la construction des liens. Par exemple, le poids des liens peuvent, ou non, être cumulés comme le fait V.Krebs (un lien faible entre A et B de poids 1 + un lien fort entre A et B de poids 5 = un lien de poids 6). Un tableau plus général se dessine alors où les 19 kamikaze sont effectivement reliés entre eux:



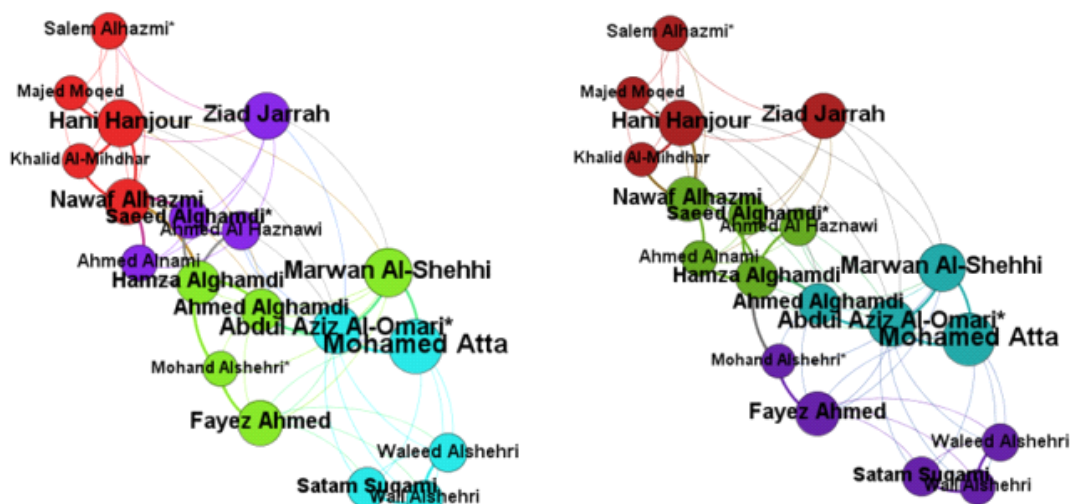
Les liens épais représentent les liens forts sous Gephi. Ils permettent de relier les nœuds du graphe en fonction de relations «profondes» d'amitié et d'appartenance familiale.

Quand on dispose ainsi d'un graphe connexe, on peut alors déployer les métriques de graphes qui ont guidé jusqu'à présent l'exploration des données web. Les attentats du 11 septembre 2001 peuvent *enfin* être considérés sous l'angle d'un réseau de relations sociales, historiques ou factuelles. Ce réseau possède donc un diamètre, une densité, des éléments hiérarchisables selon des scores de degrés ou regroupables en clusters. Spatialisée en *ForceAtlas2*, la structure de réseau permet par exemple d'apercevoir les rôles proéminents de Mohammed Atta mais aussi de Marwan Al-Shehhi ou de Hani Hanjour.



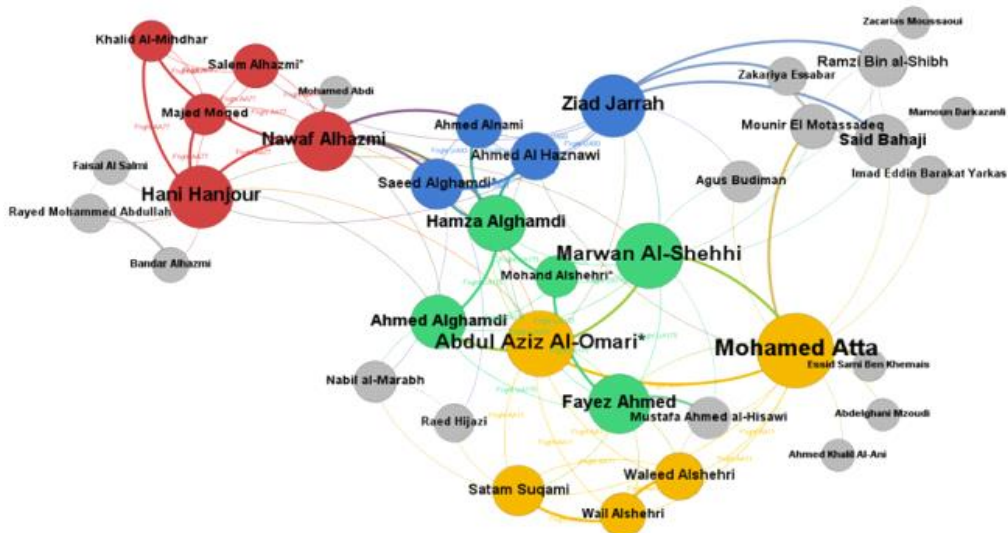
La pondération des liens dans la table de données permet alors de faire varier les vues sur les données avec l'algorithme ForceAtlas2 (fonction «poids des liens» activable sous Gephi). De gauche à droite, trois vues sur la même structure de graphe en augmentant petit à petit la prise en compte des scores de pondération. La vue de gauche est plutôt «fonctionnelle» tandis que celle de droite a tendance à regrouper les acteurs en fonction de leurs liens d'amitié et familiaux.

Toute la batterie de métriques réseau peuvent être déployées: on peut ainsi «pageranker» les acteurs ou encore (comme plus haut) faire varier les forces d'attraction entre les nœuds ou les acteurs du réseau en fonction du poids des liens. Parmi toutes ces métriques, une m'interrogeait: le diamètre du graphe (ramené ici à la longueur moyenne des chemins possibles entre les nœuds considérés par paire), autour de 4,5, ce qui paraît important pour un réseau de seulement 19 acteurs. Cette longueur apparente peut être considérée comme un principe de sécurité pour les membres du réseau (chaque élément étant éloigné des autres, il ne dispose, par exemple, que d'informations partielles sur l'ensemble de l'organisation). Mais cette longueur relative contredit un principe nécessaire de coordination (donc de distance moyenne *courte*) dans la préparation des attentats comme dans la synchronisation des équipes. Sous un autre aspect, des métriques comme les formes de centralité (notamment *betweenness*) indiquent le rôle prépondérant de certains acteurs, comme Mohamed Atta, qui ont pu jouer le rôle d'animateur de l'ensemble des 4 cellules. Sachant que ce dernier était aussi l'un des pilotes, ce djihadiste concentre une double compétence (*skill*) qui rend fragile l'organisation (son arrestation avant les attentats aurait peut-être permis de les éviter). L'étude de la distribution de *skills* nécessaires à l'échelle d'une organisation pour préparer et réaliser des attentats semble être un enjeu important de programmes comme S.T.A.R.T. et semble issue, pour partie, des méthodes d'analyse des réseaux sociaux. Elle nous incite à comprendre, notamment, comment la couche des "skills" permet de relier organiquement des individus à une organisation et comment elles sont distribuées depuis les équipes opérationnelles jusqu'aux cercles lointains du soutien logistique ou idéologique. Ainsi, différentes métriques permettent d'explorer le corpus sous différents angles. Par exemple on peut comparer la répartition des acteurs tels qu'ils se sont distribués en cellules opérationnelles dans les quatre avions et la façon dont un algorithme de détection de communautés comme *Modularity* les répartit. Dans la vue suivante, l'algorithme n'a pas été contraint en nombre de classes (il produit donc une distribution "optimale" en termes statistiques) mais certains acteurs forment désormais des groupes nouveaux, très liés apparemment à des questions de relations familiales, de nationalité et de coparticipation à des périodes d'entraînement en Afghanistan.



A gauche, la répartition par groupe telle que le révèle un algorithme de détection de communautés comme *Modularity* en prenant en compte le nombre, la variété et le poids des nouveaux liens qu'ont révélés les enquêtes. A droite, la même structure mais les couleurs et les groupes sont ceux des cellules opérationnelles. On peut ainsi considérer, à certains égards, que la mise en place de cellules opérationnelles s'appuie sur des connexions plus anciennes. En mixant les acteurs en fonction de leurs compétences (les cellules opérationnelles) et non pas de leurs accointances sociales ou de leurs connexions passées, on rend difficilement repérable une cellule active, composée à priori d'acteurs ayant des horizons différents.

Si l'on projette cette fois les 37 acteurs intégrés dans le jeu de données, une perspective plus globale s'ouvre pour la compréhension de qui fait réseau et qui imprime à une organisation sa morphologie particulière (peut-être comme une sorte d'empreinte de sa structure).



Le graphe laisse maintenant apparaître quelques propriétés nouvelles intéressantes. En termes de métrique, le diamètre se situe désormais autour de 2,5, bien que le nombre d'éléments ou d'acteurs cartographiés ait doublé. L'univers des membres d'*AlQaeda* semble plus réduit (effet *small-world*), univers dans lequel peuvent alors s'exercer une actions de synchronisation, de coordination ou d'animation de l'ensemble du réseau à partir de distances plus courtes. Si l'on poursuit à la lumière des quelques enseignements issus de l'exploration du web, ce type de propriétés annonce aussi un effet de hiérarchisation autour d'éléments qui semblent jouer un rôle central, soit parce qu'ils développent de nombreux liens avec tous les autres noeuds du système (les *hubs*) soit parce qu'ils occupent une "place centrale" dans le graphe (différents scores de centralité). Cette dernière propriété est intéressante à mobiliser pour identifier des faiblesses potentielles de ce type d'organisation: certains acteurs jouent un rôle déterminant en termes de structure non pas parce qu'ils sont les plus connectés (voire les plus "actifs" ou "en pointe") mais parce que c'est par eux que passe un grand nombre de connexions (chemins possibles) entre tous les éléments.

Le principe de l'extension du corpus laisse apparaître de nombreuses difficultés, et quelques "surprises". Par exemple, l'apparition dans le graphe de Zacarias Moussaoui, de nationalité française et emprisonné à vie aux Etats-Unis. Son cas pose d'intéressants problèmes d'intégration au corpus: à priori, rien n'est encore venu prouver sa participation aux attentats du 11septembre (déjà emprisonné au moment des faits) mais il est lié de bien des façons à plusieurs acteurs de la carte (Afghanistan, séjours en Grande-Bretagne...). Autre "singularité" associée au principe de l'expansion du corpus: Valdis Krebs fait apparaître dans ses graphes la présence de Djamel Beghal, le franco-algérien, parfois présenté comme le "mentor" des frères Kouachi (Chérif avec qui il était en prison et que A. Coulibaly a visité en prison). Les trois se sont retrouvés par la suite dans le Cantal où Djamel Beghal a été assigné à résidence durant

plusieurs mois. Je ne sais pas si les attentats récents sur notre territoire (i.e. janvier 2015) représentent un "11 septembre français" mais, en terme de cartographie, nous ne sommes ainsi qu'en "distance 2" avec les événements du 11 septembre 2001 aux Etats-Unis. Une cartographie du *small world* du terrorisme est donc possible, à supposer que soit contrôlé précisément le processus d'expansion du graphe.

Durant les mois qui ont suivi les attentats du 11 septembre, les chercheurs (dont Valdis Krebs) et les spécialistes du renseignement aux Etats-Unis ont tiré quelques leçons essentielles dans la lutte qui s'engageait à grande échelle contre le terrorisme, particulièrement en matière de renseignements. Et ce sont des *hard lessons*. La première de ces leçons est qu'il est possible d'agrèger des informations largement publiques pour produire des instruments pertinents d'observation des organisations criminelles. Le web fournit aujourd'hui suffisamment de données accessibles pour isoler des frontières d'organisation et comprendre la nature de leurs activités. Un jour, peut-être, verra-t-on surgir un *djihad-leaks* issu d'un travail de fourmis réalisé par quelques journalistes motivés? Une autre leçon, complémentaire, est de type juridique, voire éthique. L'opération technique d'agrégation ouvre sur la question de l'accès à de nombreuses sources d'information ou du droit à les archiver et à les croiser. La nécessité d'un accès presque total à toutes les sources d'informations (auxquelles il faudrait ajouter aujourd'hui la géolocalisation des objets mobiles, tweeter, les bases biométriques...) fait figure d'évidence ne serait-ce qu'à la lumière de notre petite expérimentation cartographique en suivant les pas de Valdis Krebs. C'est toute la logique du *PatriotAct* et de la mise en place aux Etats-Unis du *Intelligence Reform and Terrorism Prevention Act* en 2004. Mais elle inaugure, aussi, une frénésie de l'archivage tous azimuts dont la N.S.A. semble être l'instrument principal. Comme beaucoup, j'ai été surpris de l'ampleur du système "d'écoutes" déployé durant des années par la NSA telle qu'elle apparaît avec les informations livrées par E. Snowden sans oublier, pourtant, cette impérieuse nécessité du rassemblement de l'information que produisent les individus ou les organisations sociales.

De là, une troisième leçon issue du *11/11/2001*, plus tardive mais tout aussi importante sur le plan méthodologique: la nécessité de développer des démarches qualitatives de contrôle des masses de données. En 2006, Jeff Jonas et Jim Harper reviennent sur les méthodes déployées en matière de renseignement pour en souligner les déficiences qualitatives:

"The details of the 9/11 story reveal that federal authorities had significant opportunities to unravel the 9/11 terrorist plot and potentially avert that day's tragedies. Two of the terrorists who ultimately hijacked and destroyed American Airlines flight 77 were already considered suspects by federal authorities and known to be in the United States. One of them was known to have associated with what a CIA official called a "major league killer." Finding them and connecting them to other September 11 hijackers would have been possible—indeed, quite feasible— using the legal authority and investigative systems that existed before the attacks (...) What the 9/11 story most clearly calls for is sharper focus on the part of our national security agencies and *the ability to efficiently locate, access, and aggregate information about specific suspects*. Investigators should use intelligence to identify subjects of interest and then follow specific leads to detect and preempt terrorism (...) They lacked this focus and capability—with tragic results".

Le focus qualitatif appliqué aux masses de données conduit ainsi naturellement à adopter un modèle *subject based*, en orientant «en profondeur» l'observation d'un sujet ou d'un groupe de sujets. C'est le principe d'une sorte de balance entre l'approche qualitative (celle des

enquêtes de terrain par exemple) et l'approche quantitative où le *big-data* est mobilisé à la recherche de patterns spécifiques associées aux organisations criminelles et/ou terroristes. La traçabilité du social suppose que l'on extrait des masses de données ces *smart data* qui viennent s'agréger en tableaux cohérents et précis. Le *small* ou le *smart data* pourraient permettre d'éviter la production de ces *false positives* (personnes innocentes dont le profil a été identifié statistiquement comme suspect possible) qui alimentent encore aujourd'hui le débat aux Etats-Unis. Comme l'écrit E. Dahl en 2013, "It's not Big Data, but *Little Data*, that prevents terrorist attacks", non pour revenir aux méthodes intégralement artisanales de l'enquête mais bien pour signaler qu'en matière d'intelligence des données appliquées à la sécurité c'est dans le croisement des deux méthodes que réside le potentiel prédictif des instruments ou des démarches (disons, de démarche *instrumentée* d'observation).

L'équilibre de la balance quali-quantitative qu'appellent de leurs vœux les spécialistes du renseignement et de la sécurité n'est qu'un aspect local et particulier (et ô combien instructif!) de l'hybridation des deux dimensions du traitement de l'information dans un système où tout dépend de la façon dont sont distribués dans un système d'information des phases et des approches qualitatives et quantitatives, depuis les algorithmes jusqu'aux interfaces. C'est ce que j'appelle *l'alchimie quali-quantitative* où alternent d'un côté la production qualitative de profils d'entités qui doivent être nourris de données quantitatives pour valider certains de leurs traits. De façon complémentaire, une approche massive de type *big-data* en termes de recueil et d'archivage des données doit être nourrie de *patterns* (donc de profils) élaborés qualitativement pour optimiser les recherches (comme une sorte d'ontologie sociale où seraient archivées les entités nommées mais aussi le nombre et la nature de leurs connexions). Ces principes généraux résonnent de façon particulière dans le contexte des récents attentats en France, notamment sur les plateaux de télévision où nombre d'experts ont souligné tour à tour le "manque d'analyse" et l'impossibilité pour les enquêteurs d'avoir accès à de nombreuses sources d'information.

Les nouvelles lois sur la sécurité maintenant envisagées par les pouvoirs publics permettront, peut-être, d'apercevoir plus précisément les frontières et la morphologie des organisations criminelles. Ce type de recherche me paraît aujourd'hui constituer le terrain d'expérimentation le plus riche et le plus prometteur pour ses retombées pour les sciences humaines dans leurs différents questionnements du social, surtout en sociologie et en anthropologie. Mais ce terrain sécuritaire représente aussi un formidable espace d'innovation en termes d'ingénierie pour les difficultés méthodologiques qu'il suppose d'avoir levées. Et les enjeux sont importants puisque ces recherches sont supposées mener de l'observation à la *prédiction* et à la *prévention*.

L'espace occupé par cette recherche paraît maintenant assez balisée et portera massivement ces prochaines années sur le renouvellement des hypothèses, des concepts, des méthodes et des instruments dédiés à l'observation et à la modélisation du social et de ses entités interagissantes. Il m'apparaît clairement aujourd'hui que ces recherches sur les *data* liées aux questions de sécurité en constituent l'avant-garde, là où va se jouer notre avenir en termes de sécurité collective. Le chemin à parcourir pour les spécialistes du domaine est encore long: leur point de départ est analogue à celui des *web scientists* au début des années 2000, coincés entre deux façons d'envisager le terrorisme qui rappellent le *nuage* et les *abîmes* du *deep-web*.

D'un côté, notre petite expérience de cartographie qualitative montre que la théorie du «loup solitaire» ou du groupe isolé ne peut tenir longtemps, comme à chaque fois lorsque l'enquête se développe. Les frères Couachi et N. Koulibaly sont rapidement pris et associés dans un réseau de connexions ou de relations qui contextualisent (et parfois peuvent en partie expliquer) les opérations terroristes de ce mois de janvier 2015. Elles sont nombreuses et complexes, un long travail d'identification commençant. Cependant, d'un autre côté, la recherche de patterns sociaux, donc d'un principe d'organisation, permet de dépasser la notion de «nébuleuse djihadiste» dont parlent tant nos médias, comme si l'on avait à faire à une masse sociale indistincte. Ce qu'il y a de certain, et qui nourrit cette idée de la nébuleuse, c'est que l'on ne peut pas réduire, ni même appréhender, les réseaux djihadistes comme des organisations représentables sous forme d'arborescence (il suffirait alors dans ce cas d'en éliminer le leader ou les éléments des niveaux inférieurs de hiérarchie). C'est entre l'idée du *loup solitaire* et celle de la *nébuleuse* que se loge la recherche à venir, tout comme la recherche d'un modèle ou une carte du web se place entre le nuage et le *deep-web*. C'est là, dans cet espace, qu'il s'agit de placer quelques 5.000 suspects présents sur notre territoire (si l'on en croit le Ministre de l'intérieur à la fin du mois de janvier 2015), ainsi que le réseau de leurs relations directes ou indirectes. A-t-on commencé à répertorier toutes ces entités nommées et à recenser les différents types de connexion qui les unit probablement?

En regardant la tablette *Androïd* dans mon salon, je me dis que si nous laissons spontanément autant d'informations personnelles aux géants de l'information et à leurs services marketing, nous pourrions aussi le faire pour nos enquêteurs. Parmi toutes les approches possibles, celle par graphe serait diablement intéressante pour documenter ma vie personnelle. Il pourrait en être de même pour des suspects identifiés à condition de ne pas se laisser happer par la logique foisonnante des liens qui nous unissent tous via les technologies, ou non d'ailleurs. Dès décembre 2001 Thomas A. Stewart soulignait dans son article *Six Degrees of Mohamed Atta* ⁽⁴⁾ les implications sociales ou culturelles de ce que signifient les graphes, les calculs de diamètre ou la recherche de clusters avec des métriques de graphes. Les mesures de distance dans l'espace social, rappelons-nous, peuvent être considérées comme des indices de *proximité* mais aussi d'*éloignement*. Pour le cartographe que je suis, le plus frappant est de prendre conscience à quel point la cartographie du social englobe potentiellement aussi bien ceux qui ont tout de l'altérité radicale comme les djihadistes que sa propre famille ou ses amis. De ce point de vue, nous sommes tous dans la *même carte*.

⁴ <http://faculty.cbpp.uaa.alaska.edu/afgjp/PADM610/Six%20Degrees%20of%20Mohamed%20Atta.pdf>