



Projections algorithmiques et villes ubiquitaires

Thierry Berthier

Chaire de cybersécurité & cyberdéfense Saint-Cyr Thales Sogeti

Juillet 2015 – Article I.3

Abstract

Le formalisme des projections algorithmiques fournit une représentation pratique des données produites et stockées lors des interactions entre un usager et un système numérique. Il permet entre autres de distinguer la part d'information créée de manière volontaire par cet usager de celle qui relève d'une production strictement systémique. Le ratio des volumes de données volontaires / systémiques constitue alors un bon indicateur de l'évolution d'un environnement "intelligent" et de la densité de ses infrastructures connectées. Le développement de villes ubiquitaires comme U-Songdo en Corée du Sud induit une exploitation massive des projections algorithmiques des habitants. On propose alors d'utiliser le ratio des volumes de données volontaires / systémiques pour définir le niveau d'ubiquité d'une ville "intelligente". On observe ensuite l'effet des algorithmes prédictifs déployés au sein de la ville ubiquitaire pour décrire l'émergence de boucles systémiques qui modifient en profondeur les pratiques urbaines et sociales des usagers.

Mots-clés

Projections algorithmiques, données numériques, information ubiquitaire, ville ubiquitaire, niveau d'ubiquité, Smart City, Songdo, boucle systémique, aléatoire.

I - Les projections algorithmiques d'un individu, définitions et formalisme

Le concept de projection algorithmique a été introduit en 2013 [1] pour décrire la production de données et de métadonnées résultant des interactions d'un opérateur humain avec les systèmes qui l'entourent. Par son formalisme élémentaire, il permet de généraliser la notion floue de traces numériques produites volontairement ou non par un individu. Prenons l'exemple d'un message twitter rédigé par un utilisateur du réseau social. Son auteur croit n'émettre que 140 signes et être donc en mesure de les contrôler. Toutefois, plus de 4000 signes sont réellement envoyés. La différence tient à l'ensemble des métadonnées qui accompagne chaque message (compte de l'émetteur, adresse IP, date et lieu d'émission, etc.). Encore s'agit-il là d'une démarche volontaire de l'individu, d'une émission consciente. Mais chaque fois que l'individu traverse la ville, il passe sous l'œil intrusif de caméras de surveillance, franchit le tourniquet du métro, met en œuvre de multiples senseurs qui comptabilisent son activité. Une part toujours plus grande de ses actes est ainsi captée par un système d'observation

qui agrandit, sans qu'il en ait conscience, sa projection algorithmique. Ainsi, la projection algorithmique d'un individu donné est bien plus importante que ce qu'il croit émettre par ses usages conscients...

Définition d'une projection algorithmique : Lorsqu'un individu H décide ou provoque l'exécution d'un algorithme A sur un système de calcul et de stockage S, une partie de l'information résultant de cette interaction est archivée sur S. Nous appelons alors projection algorithmique de H sur S selon A, et notons $P_S(H/A)$ l'ensemble de mots binaires finis (des mots formés d'une suite finie de 0 et de 1) archivés sur S et résultant de l'exécution de A sur S décidée ou provoquée par H.

On peut alors décomposer la projection algorithmique d'un individu selon deux partitions distinctes : la partition d'accessibilité qui sépare la projection en projection ouverte et projection fermée et la partition de libre arbitre qui la sépare en projection volontaire et projection systémique.

La partition d'accessibilité : La projection algorithmique est scindée en deux sous-ensembles disjoints appelés projection ouverte et projection fermée. La projection ouverte notée $PO_S(H/A)$ contient l'information archivée sur S accessible à tout utilisateur ou tout système extérieur. Elle constitue la composante ouverte et publique de la projection algorithmique. La projection fermée notée $PF_S(H/A)$ réunit les mots binaires archivés sur les unités de stockage de S, maintenus privés, et réservés au seul groupe supervisant le système S (ses administrateurs par exemple, dans le cas d'une supervision de nature humaine). La projection algorithmique s'exprime alors comme une réunion disjointe :

$$P_S(H/A) = PO_S(H/A) \cup PF_S(H/A)$$

La partition de libre arbitre : La projection algorithmique est scindée en deux sous-ensembles disjoints appelés projection volontaire et projection systémique. La composante volontaire notée $P_{VOL-S}(H/A)$ contient l'information volontairement transmise au système par l'individu lors de l'interaction. La composante systémique notée $P_{SYST-S}(H/A)$ contient quant à elle l'information "non volontaire" ou purement systémique engendrée lors de l'exécution de l'algorithme A sur le système. Les métadonnées créées durant l'interaction figurent dans la composante systémique. La projection algorithmique s'exprime alors comme réunion disjointe des composantes volontaires et purement systémiques :

$$P_S(H/A) = P_{VOL-S}(H/A) \cup P_{SYST-S}(H/A)$$

Définition de la S-projection algorithmique : Nous appelons S-projection algorithmique d'un opérateur H sur le système S, la réunion des projections algorithmiques $P_S(H/A)$ prise sur tous les algorithmes exécutables sur S. La S-projection est notée $P_S(H)$ et vérifie :

$$P_S(H) = \bigcup_A P_S(H/A)$$

On retrouve dans la S-projection une partition d'accessibilité : $P_S(H) = PO_S(H) \cup PF_S(H)$ dans laquelle $PO_S(H)$ désigne la S-projection ouverte de H et $PF_S(H)$ sa S-projection fermée et une partition de libre arbitre : $P_S(H) = P_{VOL-S}(H) \cup P_{SYST-S}(H)$, $P_{VOL-S}(H)$ désignant la S-projection volontaire de H et $P_{SYST-S}(H)$ sa S-projection systémique.

Définition de la projection globale : La projection globale notée $P(H)$ est obtenue en considérant la réunion sur tous les systèmes S des S-projections algorithmiques. Cette projection algorithmique globale s'écrit : $P(H) = \bigcup_S P_S(H)$ et doit être vue comme notre reflet numérique global. Elle se décompose d'une part en composantes ouverte et fermée : $P(H) = PO(H) \cup PF(H)$ et d'autre part en composantes volontaire et systémique $P(H) = P_{VOL}(H) \cup P_{SYST}(H)$. En tant qu'ensemble de données, la projection globale débute sa croissance à la naissance de l'individu et se stabilise à son décès.

II - La ville ubiquitaire, génératrice de projections algorithmiques

La ville connectée participe très directement au développement des projections algorithmiques de ses habitants. Les interactions avec les réseaux de capteurs et les objets connectés [2] présents dans le paysage urbain contribuent à l'enrichissement de la projection globale et en particulier de sa composante systémique. Le volume de la projection algorithmique systémique tend d'ailleurs à dépasser celui de la projection volontaire. Il s'agit là d'une tendance forte qui ne peut que se trouver renforcée par l'émergence d'une ville ubiquitaire (U-city) ultra connectée.

Le concept de ville ubiquitaire (U-city) se rapporte à une ville hyper connectée dans laquelle l'information est présente, tout le temps, partout, accessible à tous et invisible [3]. La ville ubiquitaire a su intégrer l'information ubiquitaire dans son architecture. Elle est construite autour d'un système d'information centralisé (U-Media Center) qui fonctionne comme le cerveau de la ville. Le U-Media Center collecte l'ensemble des données issues des réseaux de capteurs implantés dans l'espace urbain, les interprète à l'aide d'algorithmes de traitement Big Data puis oriente le fonctionnement des automates connectés en fonction des résultats de calculs. Cet enchaînement systémique opère en temps réel, selon des échelles spatiales et temporelles très variées. L'information ubiquitaire est véhiculée par une informatique ubiquitaire qui s'est urbanisée (le concept d'Ubiquitous computing date du début des années 1990).

Ainsi, l'information se dissout à la fois dans les objets de la vie quotidienne et dans les comportements [4]. Selon Mark Weiser (Xerox Park), l'information ubiquitaire correspond à la troisième ère de l'informatique : la première ère est celle du modèle Mainframe : un ordinateur et de nombreuses personnes connectées. La seconde ère correspond à celle du PC : un ordinateur et une personne connectée. La troisième ère est celle du U-computing : une personne et de nombreux ordinateurs accessibles. Les comportements humains intègrent l'information ubiquitaire d'autant plus facilement que l'informatique ubiquitaire devient transparente pour l'utilisateur. L'informatique ubiquitaire se situe exactement à l'opposé de la réalité virtuelle qui place l'homme au milieu d'un monde virtuel issu du calcul numérique alors qu'il s'agit ici de faire disparaître l'informatique et de la dissoudre dans l'espace réel.

La ville ubiquitaire « redocumentarise » ses composantes à toute échelle d'espace et de temps. L'objet ordinaire produit de l'information, devient un support et un document. L'interaction de l'homme avec cet objet engendre des projections algorithmiques qui à leur tour deviennent information pour le système central de gestion des données. La recherche d'information et la médiation algorithmique produisent de l'information. Les contenus se fragmentent et subissent une hybridation. Tous les secteurs d'activités de la ville sont concernés par la métamorphose de son espace informationnel. Le U-shopping intègre les comportements en exploitant les projections algorithmiques du consommateur. Le mobilier urbain connecté facilite l'accès nomade instantané à une information adaptée et personnalisée [5]. Il participe directement au ciblage intelligent de la bonne offre commerciale à la bonne personne au bon moment.

Niveau d'ubiquité d'une ville connectée

Le volume d'une projection algorithmique d'un individu H créée sur un système S selon un algorithme A est noté $v(P_S(H/A))$. Il peut s'exprimer en bits ou en octets.

Considérons à présent un individu évoluant dans une ville connectée durant l'intervalle de temps $[0, T]$. Au cours de cette période, il va produire des projections algorithmiques volontaires et systémiques (involontaires) dont on mesure le volume total,

de la contribution volontaire :

$$V_{\text{vol}}(H, [0, T]) = \sum_{A, S, [0, T]} v(P_{\text{vol-S}}(H/A))$$

et de la contribution systémique :

$$V_{\text{Syst}}(H, [0,T]) = \sum_{A, S, [0,T]} v(P_{\text{Syst-S}}(H/A))$$

Les sommes sont prises sur tous les algorithmes exécutés par H sur tous les systèmes durant la période [0,T].

On s'intéresse ensuite au ratio des volumes "volontaire / systémique" durant la période considérée :

$$R(H, [0,T]) = V_{\text{Vol}}(H, [0,T]) / V_{\text{Syst}}(H, [0,T])$$

puis à la valeur moyenne $\mu(R(H, [0,T]))$ de ce ratio prise sur tous les individus fréquentant la ville durant la période [0,T]. On peut alors définir le niveau d'ubiquité d'une ville connectée en fonction de cette valeur moyenne.

Définition du niveau d'ubiquité d'une ville : Une ville est dite ubiquitaire de niveau N sur la période [0,T] si $\mu(R(H, [0,T])) < 10^{-N}$

Plus N est grand et plus la partie systémique des projections est prépondérante sur la partie volontaire. Cela signifie que durant la période considérée, la densité des objets, systèmes de surveillance vidéo et infrastructures connectées de la ville provoque cette dissymétrie. Le niveau N est globalement croissant dans une ville "intelligente" connectée. Le modèle de ville ubiquitaire est entré dans sa phase d'application en 2003 avec le projet U-Songdo.

III - U-Songdo, la ville ubiquitaire

Sortie de terre en 2003 sur 610 hectares de terrains gagnés sur la mer Jaune à proximité d'Incheon (Corée du Sud) et à 65 km à l'ouest de Séoul, Songdo offre certainement le tout premier exemple de ville ubiquitaire hyper-connectée à développement durable. Cette ville intelligente accueille aujourd'hui plus de 76 000 habitants et concentre environ 300 000 emplois sur place. Le coût de développement de Songdo estimé à plus de 35 milliards de dollars est supporté par un consortium privé composé du groupe Gale International (61%) concepteur du projet [8], du groupe Posco, producteur d'acier (30%) et de la banque d'investissement Morgan Stanley (9%). Songdo est avant tout un concept de ville intelligente ultra-connectée exportable et applicable à d'autres sites géographiques. Cette ville souhaite devenir un centre d'affaire de tout premier ordre, un pôle de recherche universitaire technologique majeur et une cité-laboratoire expérimentant le principe ubiquitaire en vraie grandeur. C'est à Songdo que l'on trouve les plus hautes tours de Corée du Sud dépassant les 480 mètres. La ville a été pensée et construite autour d'un système de calcul centralisé (le U-Média Center) collectant l'ensemble des données transmises par la multitude de capteurs, senseurs, caméras de surveillance implantés dans les édifices et le mobilier urbain.

La conception de Songdo est inédite dans le sens où l'architecture de la ville s'accorde parfaitement avec sa projection algorithmique globale. Tous les bâtiments ont été « informatisés » dans leurs structures. Les caméras de sécurité, omniprésentes, filment l'intégralité de l'activité urbaine en temps réel et permettent par exemple la reconnaissance optique des plaques d'immatriculations comme la reconnaissance biométrique des habitants [6]. Les données collectées sont transmises en continu au système de calcul, le U-Média Center. Elles sont ensuite analysées par des algorithmes sophistiqués qui produisent et émettent des avis, des préconisations et des prévisions destinés à réguler les trafics et les flux. L'information ubiquitaire, « accessible tout le temps, partout et pour tout le monde » sert également à optimiser la dépense énergétique globale de la ville et à minimiser son empreinte carbone. Les données collectées et analysées impactent et influencent le « comportement écologique de

Songdo ». Construite autour d'un parc central de 41 hectares, la ville est suréquipée en pistes cyclables, terrains de golf, taxis fluviaux. Elle est dotée d'un métro ne produisant aucun rejet de CO₂, les places de parking sont souterraines, l'eau de pluie est collectée et filtrée de manière systématique et les déchets ménagers sont acheminés vers la centrale de traitement par un aspirateur centralisé puis transformés en électricité. Les bâtiments quant à eux sont dotés de toitures végétalisées et de panneaux solaires à haute performance. Chacun d'entre eux a été construit selon le standard LEED (Leadership in Energy and Environmental Design), c'est-à-dire à haute qualité environnementale. L'optimisation énergétique est calculée en temps réel par le U-Media Center qui règle et adapte en permanence les infrastructures urbaines selon un objectif de minimisation des consommations. La ville ubiquitaire s'inspire de la biomimétique en cherchant à reproduire des écosystèmes qui intègrent les technologies humaines grâce aux énergies vertes.

Le géant américain des réseaux Cisco a fait de Songdo sa ville laboratoire et procède sur place à de nombreuses expérimentations. Parmi ces expériences de services ubiquitaires, on trouve la consultation médicale et psychologique en haute définition depuis chez soi avec les meilleurs spécialistes hospitaliers ou encore l'enseignement de l'anglais à domicile avec des enseignants américains. Ces services sont particulièrement appréciés par les mères de familles coréennes qui peuvent surveiller leurs enfants et leur faire bénéficier des enseignements les plus compétitifs sur le marché, sans sortir de chez elles. Les infrastructures numériques de Songdo favorisent aussi le télétravail. Cisco supervise le système U-Life et ses salles de commandes connectées en permanence à l'ensemble des appartements de la ville. Enfin, l'ONU a choisi d'installer le siège de son « Global Green Fund » à Songdo, ville qui devrait voir sa population croître et atteindre les 265 000 habitants en 2018, l'année de fin de chantier [7].

La Chine vient de commander au groupe américain Gale International sa première ville ubiquitaire en kit sur le modèle d'U-Songdo. Le marché chinois est logiquement considéré comme prioritaire chez les concepteurs de villes intelligentes. Songdo partage au niveau mondial le titre d'U-City avec deux autres villes ubiquitaires : Masdar construite en plein désert par les Émirats Arabes Unis avec un budget de 18 milliards de dollars, qui fonctionne à 100 % avec des énergies renouvelables et Fujisawa actuellement en cours de construction, développée par Panasonic au sud de Tokyo. Dans chacune de ces villes ubiquitaires, c'est le système de calcul central qui rend la ville « intelligente ». Plus ce système est algorithmiquement performant et plus la ville devient "écologique", à l'écoute de ses ressources propres et de sa consommation. Les interactions entre le cyberspace et l'espace physique interviennent alors pleinement dans le développement d'une intelligence de la ville qui sait exploiter l'information pour optimiser son fonctionnement. Un exemple simple concerne Fujisawa dont l'éclairage public ne s'allume que lorsque les capteurs détectent la présence d'un individu. C'est alors la projection algorithmique de cet individu qui va contribuer à l'économie des énergies et à la régulation des flux.



Fig 1- U-Songdo, vue d'ensemble

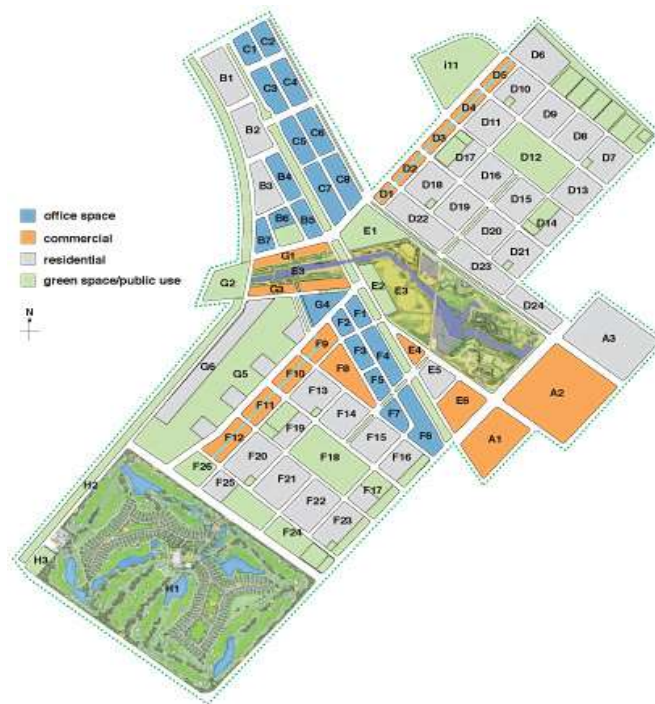


Fig 2 - Masterplan - Songdo

IV - Les algorithmes prédictifs

Ces algorithmes s'appuient en général sur l'exploitation d'un flux continu de données (données en streaming) pour construire des prévisions sous la forme d'événements probables associés à une probabilité de réalisation. Une fois analysé, le flux de données permet d'anticiper une tendance, une évolution dans le temps ou la valeur future d'une variable. L'objectif principal de l'algorithme prédictif est de maximiser la fiabilité des prévisions qu'il produit. Un algorithme qui se « tromperait » souvent n'aurait aucune valeur et ruinerait la réputation de ses concepteurs. Cet algorithme s'efforce donc de faire reculer l'aléatoire au profit d'un déterminisme facilitant les interactions de l'utilisateur avec son environnement. Les applications mobiles prédictives peuvent ainsi s'avérer très performantes pour prévoir l'affluence sur une ligne de métro en fonction de l'horaire, du lieu et de l'ensemble des données collectées en temps réel par les capteurs d'un réseau de transport.

Dans ce domaine, la société Snips a développé en 2012, en partenariat avec la SNCF, l'application Tranquilien qui prédit quelles seront les rames des trains du réseau Transilien les plus fréquentées et qui détermine les wagons qu'il faudra choisir afin de voyager plus sereinement. Les algorithmes de Tranquilien utilisent à la fois les données SNCF, les données ouvertes (Open Data) et les données de géolocalisation produites par les smartphones des usagers. Ces données sont ensuite croisées, interprétées, extrapolées de manière pertinente pour produire la prévision d'affluence qui est mise à jour en temps réel grâce aux informations fournies par les voyageurs. Le volet participatif-collaboratif est une composante importante du processus de construction de la prévision et il contribue largement à sa fiabilité. La prévision est ensuite diffusée auprès des usagers qui peuvent alors en tenir compte avant de monter dans un wagon. Le transport devient ainsi plus « intelligent » et plus interactif. L'exemple de Tranquilien illustre parfaitement le concept de boucle rétroactive qui se met en place lorsqu'une infrastructure prédictive s'active.

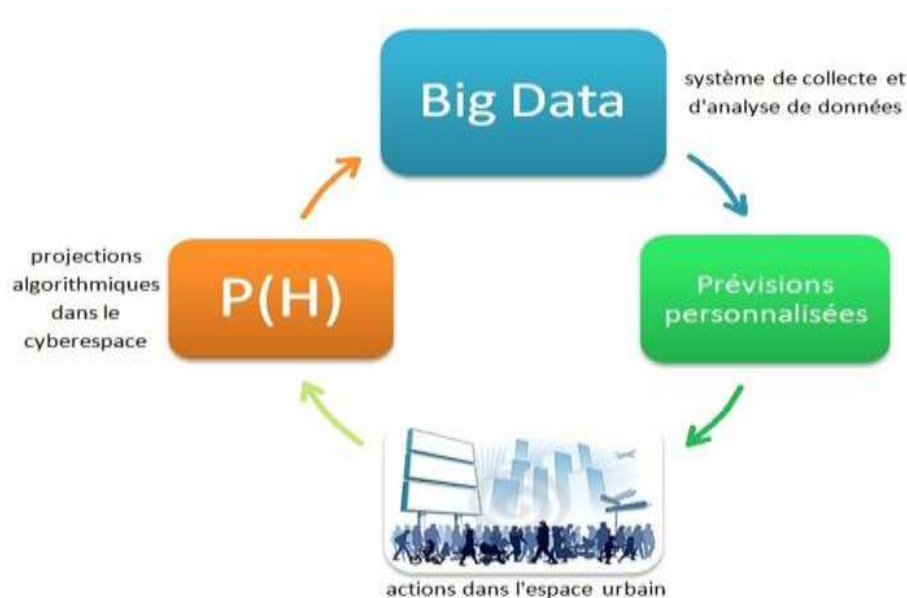
V - La boucle systémique « Data-prédicatif-action »

L'espace urbain ultra connecté devient ainsi l'acteur de la prévision. Il participe au mouvement global qui consiste à repousser l'aléatoire. Des boucles rétroactives apparaissent entre le citoyen usager de la ville connectée et les infrastructures urbaines. Elles s'installent selon la séquence suivante :

- 1 - La ville produit des données massives à partir des projections algorithmiques des usagers.
- 2 - Ces données sont analysées par des systèmes Big Data qui fournissent en temps réel prévisions et tendances.
- 3 - Les usagers de la ville tiennent compte de ces prévisions, adaptent leurs comportements et leurs actions puis produisent de nouvelles projections algorithmiques.

La séquence résumant la boucle systémique peut alors s'écrire sous la forme :

Données > Calculs > Prévisions > Adaptations > Données



Boucle systémique des algorithmes prédictifs

La boucle systémique provoque une circulation permanente des données résultant des projections algorithmiques des usagers ou produites par l'analyse prédictive à destination des usagers. Ce flux continu de données augmente en volume et renforce l'information ubiquitaire baignant l'espace physique. La boucle systémique agit alors comme un générateur de données à l'interface des espaces numérique et réel. Elle joue le rôle de passerelle active entre ces deux espaces tout en réduisant une partie de l'aléatoire qui impacte les événements du réel. Une partie seulement !

VI - Les limites des algorithmes prédictifs sur le hasard sauvage.

Les infrastructures prédictives peuvent s'avérer très performantes sur un type d'aléatoire et beaucoup moins sur un autre. Ainsi, le système algorithmique PredPol (Predit Crime In Real Time - Predictive Policing) qui a été déployé dans la ville de Santa Cruz (Californie) en 2011 puis à Los Angeles, Memphis, Charleston et New York en 2012 a localement permis de faire chuter de 33 % les agressions et de 21 % les crimes violents. Le nombre de cambriolages a baissé de 20 % dans la ville de Santa Cruz en seulement six mois. À partir de données statistiques, PredPol parvient à prédire où et quand les prochains crimes et délits risquent de se produire. On notera que PredPol ne prédit pas qui va commettre ce futur crime...

Cependant, l'algorithme prédictif reste inefficace sur des événements aléatoires de type « cygnes noirs » étudiés par Nassim Nicholas Taleb [9]. Les cygnes noirs sont des événements à très faible probabilité de réalisation, à très fort impact sur le système et rétrospectivement prévisibles. Le concept de hasard sauvage décrit par Taleb s'appuie sur l'existence de cygnes noirs, par définition hors d'atteinte de tout système « Big Data » prédictif. Cette forme de hasard sauvage persiste donc au cœur des villes ubiquitaires. Le hasard « soft gaussien » répond quant à lui assez bien aux infrastructures prédictives et recule nettement sous l'effet des boucles de rétroactions systémiques. La cité ubiquitaire conserve pour l'heure cette part d'aléatoire sauvage.

Conclusion

La croissance exponentielle des projections algorithmiques des usagers de la ville "intelligente" transforme la texture même de l'espace urbain. L'information ubiquitaire recouvre désormais la ville en créant de puissantes boucles systémiques qui opèrent entre les espaces physique et numérique. Construit sur le ratio de volumes des projections volontaires et des projections systémiques, le niveau d'ubiquité d'une ville connectée permet de mesurer et de hiérarchiser sa capacité de production automatique de données numériques d'activités des usagers. D'autres mesures d'ubiquité peuvent être définies en s'appuyant cette fois sur l'efficacité des infrastructures algorithmiques prédictives implantées au cœur de l'architecture urbaine. Plus la ville est capable de limiter l'aléatoire en fournissant à ses habitants des prévisions fiables et plus l'ubiquité gagne en puissance. Le développement d'U-city comme Songdo livrées "clés en main" par de grands consortiums de bâtisseurs s'accélère aujourd'hui avec de nouveaux projets pour lesquels les défis d'optimisation énergétique et de gestion de l'aléatoire rencontrent plus que jamais ceux de l'intelligence artificielle.

Bibliographie

- [1] BERTHIER T. - " Projections algorithmiques et cyberspace " R2IE – revue internationale d'intelligence économique – Vol 5-2 2013 pp. 179-195.
- [2] "Big Data et objets connectés - Faire de la France un champion de la révolution numérique", rapport de l'Institut Montaigne, avril 2015.
- [3] EVENO E. & MESTRES J.M. Revue Urbanisme N°384, automne 2014, Dossier " villes numériques, villes intelligentes ?" pp 24 - 84
- [4] DANIELOU J., MENARD F. "L'art d'augmenter les villes, (pour) une enquête sur la ville intelligente", rapport PUCA, septembre 2013.
- [5] Actes du Forum "Villes & Numérique, nouveaux usages, nouveaux visages", Agence d'urbanisme pour le développement de l'agglomération lyonnaise, juillet 2009
- [6] Rapport Domotique News, "Ville numérique U-Songdo", N°248, septembre 2010.
- [7] LEE Yong Woo, "Ubiquitous (smart) City", EU Parliament Seminar, avril 2013.
- [8] Rapport GALE International, "City Builder - Songdo", 2010
- [9] TALEB N.N. "Le cygne noir; la puissance de l'imprévisible", Editions Les Belles Lettres, septembre 2008.

Chaire Cyberdéfense et cybersécurité

Fondation Saint-Cyr, Ecole militaire, 1 place Joffre, 75007 Paris

Téléphone : 01-45-55-43-56 – courriel : contact@chaire-cyber.fr – SIRET N° 497 802 645 000 18

La chaire remercie ses partenaires



