Algorithmic projections and ubiquitous cities

Thierry Berthier

Saint-Cyr Thales Sogeti Cyber-security & Cyber-defense Chair

July 2015 - Article I.3

Abstract

The formalism of algorithmic projections provides a practical representation of the data generated and stored during the interactions between a user and a digital system. Among others, it makes it possible to distinguish the proportion of data voluntarily created by the user compared to strictly systemically-generated data. The ratio of voluntary/systemic data volumes provides a good indicator of the evolution of a "smart" environment and the density of its connected infrastructures. The development of ubiquitous cities such as U-Songdo in South Korea induces a massive exploitation of the algorithmic projections of its inhabitants. It is proposed to use the ratio of voluntary/systemic data volumes to define the level of ubiquity of a "smart" city. The effects of the predictive algorithms deployed in the city can then be observed to describe the emergence of systemic loops which deeply modify the urban and social behaviour of the users.

Keywords

Algorithmic projections, digital data, ubiquitous information, ubiquitous city, ubiquity level, Smart City, Songdo, systemic loop, random.

I - Algorithmic projections of an individual, definitions and formalism

The concept of algorithmic projection was introduced in 2013 [1] to describe the production of data and metadata resulting from the interactions between a human operator and the systems in his environment. By its elementary formalism, it makes it possible to generalize the vague notion of digital trails produced voluntarily or not by an individual. Let us use the example of a twitter message written by a user of this social network. The author believes he is only emitting 140 characters and that he is therefore able to control these. However, over 4000 characters are actually emitted. The difference lies in the amount of metadata accompanying each message (user account, IP address, posting date and location, etc.) . And yet this is a voluntary action on the part of the user, a conscious emission. But every time the user crosses the city, he crosses the field of intrusive surveillance cameras, goes through the subway gates, activates multiple sensors which record his activity. An ever-increasing proportion of his actions are recorded by an observation system which, without his knowledge, increases his algorithmic projection. As a result, the algorithmic projection of an individual is much more important than what he believes he is emitting with his conscious actions.

Definition of an algorithmic projection: When an individual H decides or triggers the execution of an algorithm A on a calculation and storage system S, part of the information resulting from this interaction is stored in S. We then define as the algorithmic projection of H onto S according to A, noted $P_S(H/A)$, the set of finite binary words (words composed of a finite sequence of 0 and 1) archived on S and resulting from the execution A onto S, decided or triggered by H.

The algorithmic projection of an individual can then be decomposed into two distinct partitions: the accessibility partition separating the projection in open and closed projection, and the “free will” partition separating it in voluntary and systemic projection.

The accessibility partition: The algorithmic projection is divided in two disjoint subsets called open projection and closed projection. The open projection, noted $PO_S(H/A)$, contains the data archived on S which is accessible to any user or external system. It constitutes the open and public component of the algorithmic projection. The closed projection, noted $PF_S(H/A)$, regroups the binary words stored in the storage units of S, kept private and only accessible to the group supervising the S system (its administrators, for example, in the case of human supervision). The algorithmic projection can then be expressed as a disjoint reunion:

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

The free will partition: The algorithmic projection is divided in two disjoint subsets called voluntary projection and systemic projection. The voluntary component, noted $P_{VOL\text{-}S}(H/A)$, contains the data voluntarily transmitted by the individual to the system during the interaction. The systemic component, noted $P_{SYST\text{-}S}(H/A)$, contains the “involuntary” or purely systemic data generated during the execution of the algorithm on the system. Metadata created during the interaction is part of the systemic component. The algorithmic projection can then be expressed as a disjoint reunion of the voluntary and purely systemic components:

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

Definition of the S-algorithmic projection: We call S-algorithmic projection of an operator H onto the system S the reunion of algorithmic projections $P_S(H/A)$ taken from all algorithms executable on S. The S-projection is noted $P_S(H)$ and verifies the following :

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

The accessibility partition can be found in the S-projection: $P_S(H) = PO_S(H) \cup PF_S(H)$ in which $PO_S(H)$ designates the open S-projection and $PF_S(H)$ the closed S-projection, as well as the free will projection: $P_S(H) = P_{VOL-S}(H) \cup P_{SYST-S}(H)$, $P_{VOL-S}(H)$ designating the voluntary S-projection of H and $P_{SYST-S}(H)$ its systemic S-projection.

Definition of the global projection : The global projection, noted P(H), is obtained by considering the reunion on the S systems of the S-algorithmic projections. This global algorithmic projection is noted $P(H) = \cup_S P_S(H)$ and should be considered as our global digital reflection. It can be decomposed in either open and closed components : P(H) = PO(H) U PF(H) or in voluntary and systemic components $P(H) = P_{VOL}(H) \cup P_{SYST}(H)$. As a data set, the global projection begins its growth at the time of birth of the individual and stabilizes at the time of his death.

II - The ubiquitous city, an algorithmic projection generator

A connected city participates very directly in the development of the algorithmic projections of its inhabitants. The interactions with sensors and connected devices [2] present in the urban environment contribute to the growth of the global projection, and in particular to the growth of its systemic component. The volume of the systemic component tends to exceed that of the voluntary component. This is a strong tendency which can only be reinforced by the emergence of the highly-connected ubiquitous city (U-city).

The concept of ubiquitous city (U-city) refers to a hyper-connected city in which data is present, constantly and everywhere, available to all and invisible [3]. The ubiquitous city has integrated ubiquitous data in its architecture. It is built around a centralized data system (U-Media Center) which acts as the city's brain. The U-Media Center collects all of the data provided by the urban environment's networks of sensors and interprets it with Big Data treatment algorithms, and then adjusts the behavior of the connected devices based on the results of the interpretation. This systemic sequence operates in real-time, based on very varied spatial and temporal scales. The ubiquitous data is transmitted by urbanized ubiquitous computing (the concept of ubiquitous computing can be traced back to the early 90s).

In this manner, the information dissolves both into daily-life objects and behaviors [4]. According to Mark Weiser (Xerox Park), ubiquitous computing corresponds to the third era of computing: the first was that of the mainframe model - a computer with many users connected to it. The second was that of the PC: one computer and one connected user. The third era is that of the U-computing: one person and numerous accessible computers. The integration of ubiquitous information into human behavior is facilitated by the fact that ubiquitous computing becomes transparent to the user. Ubiquitous computing is the exact opposite of virtual reality, which places man in a virtual world born from digital computing, in that it attempts to make computing disappear and dissolve it in the real world.

The ubiquitous city "redocumentarises" its components to all spatial and temporal scales. The ordinary device generates information, becomes a medium and a document. The interaction of man with this device generates algorithmic projections which in turn become data for the centralized data management system. The search for information and algorithmic mediation generate data. Contents are fragmented and hybridized. All of the city's areas of activity are concerned by the metamorphosis of its information space. U-shopping integrates behaviors by exploiting the algorithmic projections of consumers. Connected street furniture facilitates instant nomadic access to adapted and personalized information [5]. It directly contributes to the intelligent supply of the appropriate commercial offer to the right person at the right time.

Level of ubiquity of a connected city

The volume of the algorithmic projection of an individual H onto a system S according to an algorithm A is noted v(PS(H/A)). It can be expressed in bits or octets.

Let us consider the case of an individual evolving in a connected city during a time interval noted [0, T]. During this period, he will generate voluntary and systemic (non voluntary) algorithmic projections for which we measure the total volume, of the voluntary contribution :

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

and systemic contribution :

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

The sums are taken from all algorithms executed by H on all systems during the period [0,T].

We then look at the ratio of the “voluntary/systemic” volumes during the period :

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

and then the average value $\mu (R (H , [0,T]))$ of this ratio for all individuals in the city during the same time period [0,T]. We can then define the ubiquity level of a connected city based on this average value.

Definition of the ubiquity level of a city: A city is said to be an N-level ubiquitous city during the [0,T] period if $\mu (R (H , [0,T])) < 10^{-N}$.

The greater N is, the more important the systemic component of the projections is, compared to the voluntary component. This means that, during the considered period, the density of devices, video surveillance systems and connected urban infrastructures are responsible for this asymmetry. The N-level is globally on the rise in an “intelligent” connected city. The ubiquitous city concept entered its application phase in 2003 with the U-Songdo project.

III - U-Songdo, the ubiquitous city

Created from scratch in 2003 on 610 hectares of ground reclaimed from the Yellow Sea near Incheon and 65 kilometers west of Seoul, Songdo is certainly the very first example of a hyper-connected ubiquitous city with sustainable development. This smart city now hosts over 76,000 inhabitants and concentrates 300,000 jobs on site. The cost of Songdo's development, evaluated at more than 35 billion dollars, is supported by a private consortium composed of the Gale International group (61%), which designed the project [8], the Posco group, an iron producer (30%), and the investment bank Morgan Stanley (9%). Songdo is first and foremost a ultra-connected smart city concept which can be exported and used in other geographical locations. This city wants to become a leading business centre, a major university technological research center and a laboratory experimenting at full-scale the ubiquitous principle. The highest towers in South Korea are to be found in Songdo, with a height of over 480 metres. The city was thought and designed around a central calculating system (the U-media center) collecting all of the data transmitted by the numerous sensors and surveillance cameras implanted in the city's buildings and street furniture.

Songdo's design is novel in that the city's architecture is perfectly adapted to its global algorithmic projection. All of the buildings' structures have been "computerized". Omnipresent security cameras film the entirety of the city's activity and, for example, allow optical reconnaissance of licence plates as well as the biometric recognition of the inhabitants [6]. The collected data is continuously fed to the calculating system, the U-Media Center. It is then analyzed by sophisticated algorithms which produce and emit opinions, recommendations and predictions meant to regulate traffic and flows. The ubiquitous information, "always available, everywhere and to all", is also used to optimize the city's energy consumption and minimize its carbon footprint. The collected and analyzed data impacts and influences "Songdo's ecological behaviour". Built around a 41-hectare central park, the city is over-equipped with cycling paths, golf courses, river taxis. It operates a subway which produces no CO_2, parking spaces are subterranean, rain water is systematically collected and filtered and household waste is routed by a centralized vacuuming system to a treatment plant and converted to electricity. The buildings are equipped with green roofs and high-performance solar panels. Each of them were built in accordance to the LEED standard (Leadership in Energy and Environmental Design), i.e. high environmental quality. Energetic optimization is calculated in real-time by the U-Media Center which constantly regulates and adapts urban infrastructures to minimize consumption. The ubiquitous city finds inspiration from biomimetics in trying to replicate ecosystems which integrate human technologies using green energy.

The American network giant Cisco has made Songdo its testing ground and runs many experiments on location. These ubiquitous services experiments include medical and psychological consultations in high definition from home with the best hospital-based specialists or English lessons from home with American teachers. These services are particularly appreciated by Korean mothers, who get to supervise their children while providing them with the market's most competitive education programs without having to leave their homes. Songdo's digital infrastructures also encourage telework. Cisco supervises the U-Life system and its command rooms which are continuously connected to all of the city's apartments. Finally, the UN elected to set up the headquarters of its "Global Green Fund" in Songdo, a town which should see its population increase and reach 265,000 inhabitants when its construction will end in 2018 [7].

China just ordered its first ubiquitous city kit from the American group Gale International, based on the U-Songdo model. The Chinese market is logically considered as a priority by the designers of smart cities. On a global scale, Songdo shares the title of U-City with two other ubiquitous cities: Masdar, built in the desert by the United Arab Emirates with an 18-billion dollar budget and operated solely with renewable energy, and Fujisawa, currently under construction south of Tokyo and designed by Panasonic. In each of these cases, it is the central calculating system which makes the city "smart". The more the system is algorithmically effective, the more the city becomes "ecological" and aware of its own resources and consumption. The interactions between cyberspace and physical space then fully intervene in the development of a city’s intelligence, capable of using the data to optimize its operation. A simple example can be found in Fujisawa where public lighting is only activated in the presence of individuals. In this case, the algorithmic projection of the individual contributes to energy savings and flow regulation.

Fig 1 - A general view of U-Songdo

Fig 2 - Masterplan - Songdo

IV - Predictive algorithms

These algorithms are generally based on the exploitation of a continuous data flow (streaming data) to build predictions of probable events associated to the probability of their occurrence. Once analyzed, the data flow makes it possible to anticipate a tendency, an evolution or the future value of a variable. The primary goal of the predictive algorithm is to maximize the reliability of the predictions it generates. An algorithm which would often “make mistakes” would be of no value whatsoever and would ruin the reputation of its creators. The algorithm therefore attempts to minimize randomness in favor of a determinism facilitating the interactions of the user and his environment. Predictive mobile applications can prove highly efficient in anticipating the congestion of a subway line based on the time, place and all of the data collected in real-time by the sensors of a transportation system.

In this field, the Snips company, partnering with the SNCF, developed in 2012 the Tranquilien application which predicts which trains of the Transilien network will be most used and determines the best wagons to board for best travel conditions. The Tranquilien’s algorithms use SNCF data, Open Data and geo-localization data generated by the users’ smartphones. These values are then combined, interpreted and extrapolated in a relevant manner to generate the congestion predictions, updated in real time using information provided by the travelers. The participative-collaborative component is an important part of the prediction generation process and largely contributes to its reliability. The predictions are then transmitted to the users who can take them into account before boarding a carriage. Thus, transportation becomes “smarter” and more interactive. The Tranquilien example is a perfect illustration of the retroactive loop concept that occurs when a predictive infrastructure activates itself.

V - The “Data-Predictive-Action” systemic loop

In this manner, the ultra-connected urban environment becomes an element of the prediction. It participates in the global movement to push back randomness. Retroactive loops appear between the connected city’s user and the urban infrastructures. They form in the following sequence:

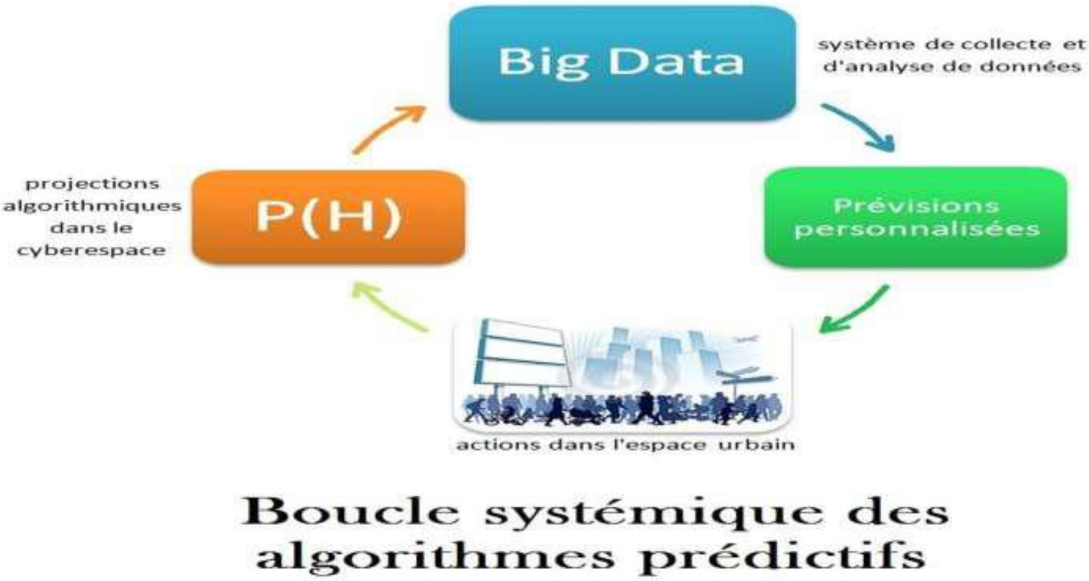
1 - The city generates massive data based on its users’ algorithmic projections.

2 - This data is analyzed by Big Data systems which elaborate real-time predictions and trends.

3 - The city’s users take these predictions into account, adapt their behavior and actions to them and generate new algorithmic projections.

The sequence resuming the systemic loop can then be written under this form:

Data > Calculation > Previsions > Adaptation > Data



Big Data
(Data collection and analysis system)
->
Individualized predictions
->
Urban environment actions
->
P(H)
Algorithmic projections in cyberspace
->
Systemic loop of predictive algorithms

The systemic loop provokes a constant traffic of data resulting from the users' algorithmic projections or generated by the predictive analysis destined to the users. This constant stream of data increases in volume and reinforces the ubiquitous data's presence in the physical space. The systemic loop then behaves as a data generator at the intersection of the digital and physical worlds. It acts as an active bridge between these two spaces while reducing part of the randomness of the events in the physical world. This is only a partial reduction!

VI - The limits of predictive algorithms on wild randomness

Predictive infrastructures can prove very effective on a given type of randomness and much less in another. Hence, the PredPol algorithmic system (Predict Crime in Real Time - Predictive Policing) which was deployed in the city of Santa Cruz, California, in 2011 and later in Los Angeles, Memphis, Charleston and New York in 2012 was able to locally decrease attacks by 33% and violent crimes by 21%. Burglaries in Santa Cruz went down by 20% in only six months. Using statistical data, PredPol can predict where and when the next crimes and misdemeanors will probably take place. One should note that PredPol cannot determine who the authors of these future crimes are…

However, predictive algorithms are inefficient in predicting random events of the "Black Swan" type studied by Nassim Nicholas Taleb [9]. Black swans are events with a very low probability of occurrence, high impact on the system and are retrospectively predictable. The wild randomness concept described by Taleb relies on the existence of the existence of black swans, which are by definition out of the reach of "Big Data" predictive systems. This form of wild randomness therefore persists in ubiquitous cities. On the other hand, "soft gaussian" randomness responds fairly well to predictive infrastructures and is largely minimized by retroactive systemic loops. At this time, the ubiquitous city is still subject to this wild randomness component.

Conclusion

The exponential growth of the algorithmic projections of users in "smart" cities is transforming the very texture of the urban environment. Ubiquitous data now covers the city by creating powerful systemic loops which operate between the physical and digital spaces. Defined by the ratio of voluntary and systemic data volumes, the ubiquity level of a connected city makes it possible to measure and compare its capacity to automatically generate digital user activity data. Other measures of ubiquity can be defined based on the efficiency of the predictive algorithmic infrastructures built in the urban architecture. The more capable the city is at limiting randomness by providing its users with reliable predictions, the greater the ubiquity. The development of Songdo-like U-cities sold as turnkey solutions by great building consortiums is on the rise with new projects in which the challenges of energy efficiency and randomness management meet those of artificial intelligence.

[1] BERTHIER T. - " Projections algorithmiques et cyberespace " 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.