



## Les erreurs dans le positionnement par satellite en environnement urbain

par Adrien Schmitt et Gaëlle Tournié  
Société de Calcul Mathématique SA

Mars 2017

### Résumé

On utilise couramment un smartphone ou un GPS pour obtenir des coordonnées  $(x, y)$  à partir d'une adresse postale, par exemple 111 Faubourg Saint Honoré. Cette conversion est indispensable au fonctionnement du système de navigation.

Notre travail montre que cette conversion peut être entachée d'erreur. Pour le système Google, l'erreur moyenne est de 40 m pour Paris, 50 m pour la banlieue, 65 m pour la grande banlieue. Pour l'Île-de-France dans son ensemble, une fois sur deux l'erreur est supérieure à 50 m, ce qui est tout de même significatif, aussi bien pour les usagers que pour l'organisation des secours : pompiers, ambulances, dépanneuses, etc.

Cette erreur peut avoir deux origines :

- Mauvais fonctionnement du GPS, surtout en environnement urbain ;
- Imprécision de l'adresse postale, surtout s'agissant de lieux vastes et mal définis : gares, complexes industriels, stades, etc.

Pour corriger, nous proposons :

- La création d'une base de données, alimentée par les usagers et administrée par le responsable de la géolocalisation. À chaque fois qu'une anomalie est constatée, l'utilisateur envoie une description de sa position réelle.
- La "sectorisation" des adresses complexes, comme les gares, en précisant par exemple le nom de la sortie et/ou les points cardinaux.

## I. Présentation

Il est de plus en plus courant, de nos jours, d'utiliser son smartphone ou bien son GPS pour se déplacer d'un point à un autre. Il existe un réel besoin de convertir une adresse postale (i.e. 111 rue du Faubourg Saint Honoré) en coordonnées  $(x, y)$  sur une carte plane. On veut ainsi déterminer la position "vraie" d'une adresse postale, calculer l'itinéraire pour la rejoindre et, dans certains cas, le temps de trajet.

Les usagers ont ce besoin couramment, mais il est particulièrement évident pour les services de secours : lutte contre l'incendie, ambulances, dépanneurs, etc. Lorsqu'une personne appelle une ambulance, elle précise son adresse postale et non ses coordonnées. L'intervention doit être rapide et toute erreur sur l'identification des coordonnées réelles est préjudiciable à l'efficacité du système de secours.

Or, comme nous allons le voir, les coordonnées associées à l'adresse peuvent être erronées. Cette erreur peut être plus ou moins importante, de quelques mètres à plusieurs centaines de mètres.

## II. L'usage du GPS

Chaque GPS dispose d'un système de géolocalisation qui associe à une adresse des coordonnées  $x$  et  $y$  (degrés décimaux). Par exemple le "111 rue du Faubourg Saint Honoré" a comme coordonnées  $(x, y)$  : (48.8726167 ; 2.31071972) [Google]. Parmi les systèmes les plus connus, on trouve Google, Tomtom, Mapping, etc.

Tous disposent d'une base de données qui liste l'ensemble des adresses et leurs coordonnées GPS. Selon le système de géolocalisation, les coordonnées associées à la même adresse ne sont pas nécessairement les mêmes.

De nombreux articles portent sur la mauvaise localisation GPS ; ils listent les sources d'erreurs causées par la mauvaise réception du signal entre le satellite et le système GPS (erreurs atmosphériques, trajets multiples, erreurs de chronométrage et d'orbite des satellites, etc.) [Galileo]. Par exemple, un conducteur est au 111 rue du Faubourg Saint Honoré mais le signal GPS le situe au 9 de la même rue.

Dans notre article, nous ne nous intéressons pas aux erreurs de "signal", mais aux erreurs concernant l'association de coordonnées  $(x, y)$  à une adresse.

Nous avons étudié les performances du système Google pour la région Île-de-France en nous appuyant sur un historique de trajets. Chaque conducteur fournit deux informations : les coordonnées reçues par le GPS pour l'adresse sélectionnée et ses coordonnées réelles, lorsqu'il est arrivé à destination. La distance entre elles mesure l'erreur de référencement pour cette adresse.

Nous avons analysé l'historique pour éliminer les données aberrantes et manquantes. Des données aberrantes peuvent être la conséquence d'un dysfonctionnement du système GPS. Si la base de données regroupe des trajets réalisés sur la zone Ile-de-France, il n'est pas acceptable de retrouver un utilisateur en Chine. Connaissant le secteur étudié, nous avons retiré les trajets où le conducteur déclare être en dehors du secteur.

Le système de géolocalisation Google n'est pas parfait. Sur l'ensemble de la zone Ile-de-France, l'erreur médiane de localisation est de 45 mètres. Si maintenant nous découpons la zone Ile-de-France en Paris, banlieue et grande banlieue, les résultats diffèrent. L'erreur médiane est de 40 mètres pour Paris, de 50 mètres pour la banlieue, de 65 mètres pour la grande banlieue.

Une erreur médiane de plus de 50 mètres peut s'avérer problématique pour les utilisateurs réguliers et les services de secours : elle peut signifier par exemple que l'on se retrouve de l'autre côté d'une grande artère.

Pour comprendre l'origine de ces erreurs, nous avons recherché les adresses les plus affectées. On retrouve entre autres la gare ROISSY TGV pour la zone grande banlieue avec une erreur médiane de 175 mètres. Ceci est compréhensible : il n'existe pas de coordonnées précises pour un bâtiment aussi grand qu'une gare. On retrouve le même problème pour des rues avec peu de bâtiments, des aéroports, des parcs d'activité commerciale, les zones en travaux ou des complexes industriels (i.e. rue des Longs Rideaux avec une erreur médiane de 340 mètres). Pour la zone banlieue, on retrouve par exemple l'avenue Jules Rimet (Saint-Denis), proche du stade de France, avec une erreur médiane de 45 mètres.

D'autres sources d'erreur proviennent des multitrajets en zone urbaine : le signal GPS est réfléchi par des bâtiments (voir [Berton]) ;

### III. Approche mathématique

Nous utilisons la méthode probabiliste développée par la SCM, décrite dans le livre [NMP]. Elle consiste à regarder la répartition des erreurs de localisation.

L'analyse se fait en plusieurs étapes :

- Calculer la distance entre les coordonnées GPS fournies par le système de géolocalisation et celles fournies par l'utilisateur lorsqu'il déclare être à l'adresse : c'est l'erreur que nous cherchons à analyser.

La distance retenue est la distance euclidienne :

$$d = \sqrt{(x_1 - x_u)^2 + (y_1 - y_u)^2} \quad (1)$$

où  $(x_1, y_1)$  sont les coordonnées fournies par le système de géolocalisation et  $(x_u, y_u)$  les coordonnées fournies par l'utilisateur sur place.

- Créer des intervalles réguliers pour les distances ;

Nous déterminons les distances minimales et maximales rencontrées dans la base de données et découpons cet intervalle en petits intervalles réguliers. Le choix du nombre et de la taille dépend de la finesse voulue pour l'analyse.

Dans notre étude nous avons trouvé une erreur minimale de 0 mètre et une erreur maximale de 22 845 mètres (gare de Roissy TGV).

Les intervalles utilisés sont du type (en mètres) :  $[0 ; 10[$ ,  $[10 ; 20[$ , etc. Chaque intervalle s'appelle une "classe" et  $K$  sera le nombre total de classes.

- Compter le nombre de fois dans la base de données où l'erreur tombe dans l'intervalle considéré ;

Par exemple, dans la base de données il y a 10 trajets qui ont une erreur comprise entre 10 et 20 mètres (intervalle  $[10 ; 20[$ ).

- Calculer la probabilité pour chaque intervalle de distance ;

Nous déterminons la probabilité que la distance soit dans l'un des intervalles définis précédemment par la formule (voir [NMP])

$$p_k = \frac{n_k + 1}{N + K} \quad (2)$$

où :

$p_k$  est la probabilité d'avoir une distance dans le  $k^{\text{ème}}$  intervalle

$n_k$  est le nombre de trajets dans le  $k^{\text{ème}}$  intervalle

$N$  est le nombre total de trajets

$K$  est le nombre de classes.

- Calculer la probabilité de dépassement de seuil (PDS) ;

Fixons un seuil d'erreur  $x$  arbitraire (par exemple 30 mètres). Il s'agit d'évaluer la probabilité qu'une adresse ait une erreur au moins égale à  $x$ . C'est par définition :

$$PDS(x) = P(X \geq x) \quad (3)$$

$$PDS(k) = P\{X \geq k\} = \frac{n_k + n_{k+1} + \dots + n_K + 1}{N + 2} \quad (4)$$

Les figures ci-dessous présentent les deux probabilités présentées précédemment.

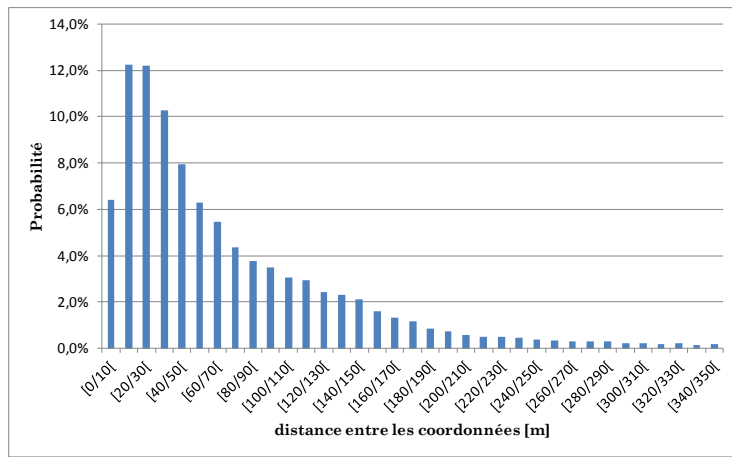


Figure 1 – Loi de probabilité pour la zone Île-de-France

Les erreurs les plus probables pour la source Google sont de 20 à 30 mètres. La probabilité décroît rapidement pour des erreurs plus élevées.

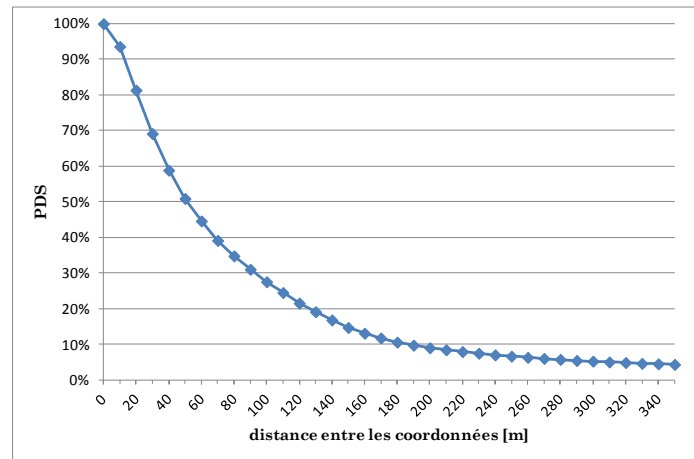


Figure 2 – Probabilité de dépassement de seuil pour la zone Île-de-France

Pour l'Île-de-France tout entière, la probabilité que l'erreur soit supérieure à 50 mètres est de 50%. Dans 95% des cas, l'erreur est inférieure à 180 mètres.

#### IV. Conclusion

Nous avons mis en évidence des erreurs sur le positionnement GPS de certaines adresses. La question maintenant est de savoir comment corriger les adresses erronées.

Une première méthode consistera à regarder la distribution spatiale du positionnement des conducteurs pour une même adresse et à en déduire son véritable emplacement. Par exemple, si vingt conducteurs déclarent tous être à l'adresse aux coordonnées  $(x_c, y_c)$ , alors que l'adresse est géolocalisée aux coordonnées  $(x_g, y_g)$ , il est préférable de changer les coordonnées de l'adresse en  $(x_c, y_c)$ . De manière générale, si l'on parvient à constituer une base de données

d'utilisateurs correspondant à une même adresse de géolocalisation, la correction sera la moyenne des déclarations d'erreurs faites par chacun des usagers.

Cette base de données d'utilisateurs devrait être mise à jour régulièrement, à partir des déclarations des conducteurs. Celui qui constate que l'adresse où il se trouve n'est pas correctement géolocalisée envoie un message d'erreur, qui sera conservé dans une base de données et analysé par les administrateurs du système, en l'occurrence Google. C'est le principe des communautés d'utilisateurs, que l'on rencontre en particulier avec Waze (déclaration de points dangereux sur les routes, etc.) ou bien Google Map Maker.

Pour les grands bâtiments (gares, aéroports, complexes industriels, etc.), il faudrait parvenir à une "sectorisation" : chaque sortie, ou chaque point cardinal, aurait sa propre adresse. Prenons l'exemple de la gare Montparnasse : celle-ci dispose de cinq sorties auxquelles il est possible d'associer des coordonnées  $(x, y)$ . Si un service d'urgence doit agir rapidement à cette adresse, elle peut alors demander la sortie la plus proche comme complément d'adresse.

## Références

[Google] Coordonnées GPS et Google Map.

Disponible sur : <<http://www.coordonnees-gps.fr/>> [consulté le 23/02/2017]

[Galileo] Jean-Marc Piéplu : GPS et Galileo, Systèmes de navigation par satellites.

Disponible sur : <<http://www.wearealgerians.com/up/uploads/139534969626851.pdf>> [consulté le 23/02/2017]

[NMP] Bernard Beauzamy : Nouvelles méthodes probabilistes pour l'évaluation des risques. Ouvrage édité et commercialisé par la Société de Calcul Mathématique SA. ISBN 978-2-9521458-4-8. ISSN 1767-1175, avril 2010.

[Berton] Gottfried Berton : Amélioration du positionnement par satellite en environnement urbain, janvier 2016. Disponible sur :

[http://www.scmsa.eu/archives/SCM\\_amelioration\\_positionnement\\_GPS\\_2016\\_01\\_25.pdf](http://www.scmsa.eu/archives/SCM_amelioration_positionnement_GPS_2016_01_25.pdf)