

Jeu de données pour l'évaluation d'algorithmes d'appariement de cartes

Matěj Kubička (L2S/IFP-EN), Arben Cela (ESIEE), Philippe Moulin (IFP-EN), Hugues Mounier (L2S) et S. I. Niculescu (L2S)

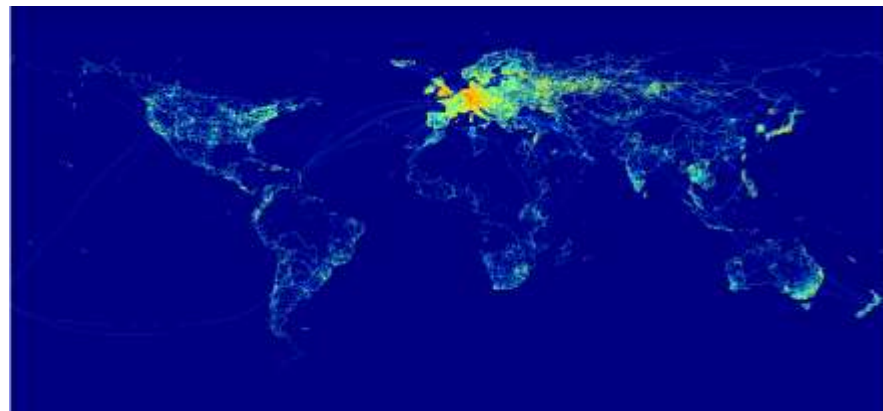
Les algorithmes d'appariement de cartes (map-matching) déterminent une position sur une carte à partir d'un système de positionnement (tel GNSS). Ils jouent un rôle important dans les SIG et les services utilisant le positionnement, tels la planification de route, la gestion et la surveillance de flotte et la location.

Des progrès considérables ont été réalisés au cours des deux dernières décennies, mais quelques problèmes persistent :

- Les Auteurs utilisent différentes méthodes pour évaluer la performance
- Le test en vraie grandeur est coûteux et chronophage

Objectifs

Le but est de résoudre les problèmes ci-dessus par de l'évaluation de performance de map-matching. La question principale est de collecter un échantillon de test suffisamment riche pour prouver les hypothèses d'appariement. Comme la carte et la précision du système de positionnement varient considérablement avec l'environnement, l'effort nécessaire excède en général les moyens mis oeuvre.



Résultats

Au lieu d'une coûteuse campagne de collecte de données, nous avons utilisé des traces GPS gratuites (pour leur distribution voir figure à gauche), sélectionné un sous-ensemble de traces de distribution uniforme, exécuté un algorithme de d'appariement et corrigé à la main les résultats. L'ensemble de données résultat contient 110 paires de traces GPS ainsi que leurs résultats de d'appariement corrigés à la main.

Il se compose de 2265,4 kilomètres de routes parcourues et de 150360 points GPS appariés.

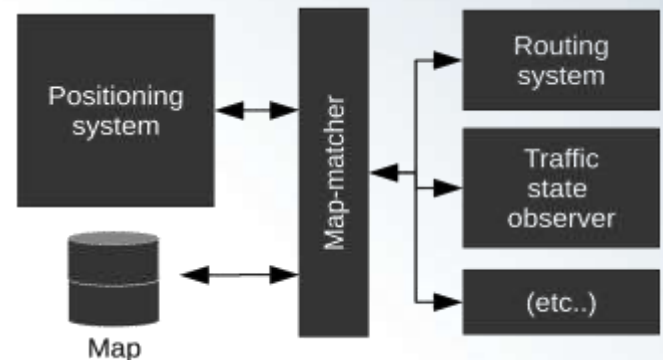
Grâce à la distribution presque uniforme des données mondiales, les données ne seront pas biaisées si l'on travaille dans une région spécifique du globe. Les données ayant été collectées sur une période de 9 ans, il n'y aura pas non plus de biais relatif aux propriétés d'un récepteur GPS particulier.

L'ensemble des données peut être utilisé de trois manières :

1. Comme un standard type pour l'analyse comparative et l'évaluation de divers algorithmes
2. Pour du test hors ligne d'hypothèses d'appariement
3. Pour de l'apprentissage hors ligne d'algorithmes adaptatifs

Conclusion

Pour la première fois un ensemble de données assez grand pour prouver ou réfuter des hypothèses d'appariement à l'échelle mondiale devient disponible. Jusqu'à présent, il y avait des obstacles considérables qui ont empêché chercheurs et ingénieurs de recueillir suffisamment de données pour des tests, étalonnage et apprentissage complets. Nous fournissons des données avec 2265,4km de routes correctement appariés dans le monde entier pour aider la communauté à surmonter cet obstacle.



↓ la suite d'outils développée pour synthétiser l'ensemble des données

