

Thème prioritaire :

Données des transports, sphère privée et avalanches de données

Traité par: EPFL

Rapport : 9 juin 2016 (2016-1)

Les données de trafic et de mobilité sont essentielles pour le déploiement des services ITS. L'information sur l'état du trafic prend de plus en plus d'importance dans l'organisation de nos déplacements. Pour un usage efficace, il faut que l'information soit adaptée aux besoins des utilisateurs, qu'elle soit facilement interprétable et surtout que sa qualité soit contrôlée. Ce rapport présente le concept d'évaluation de la qualité de l'information trafic en définissant les principes généraux et certaines méthodes de calculs. Quelques perspectives sont données quant au développement de directives qui décrivent les processus d'évaluation basés sur des indicateurs « standards ».

Responsable:

Name: Pierre-Yves Gilliéron

Organisation: Ecole polytechnique fédérale de Lausanne

Adresse: EPFL - Laboratoire de Topométrie, Station 18, 1015 Lausanne, pierre-yves.gillieron@epfl.ch

Evaluer la qualité de l'information trafic

1. Contexte

1.1 Rappel du contexte (Rapport 2)

Le premier rapport¹ sur les données des transports a défini des thèmes à approfondir en fixant des priorités et des actions à entreprendre. Le thème sur le **monitoring de la qualité** est en première priorité et quelques activités liées à ce thème ont fait l'objet de propositions. Le développement du thème du monitoring de la qualité des données et des services ITS se compose des parties suivantes avec quelques questions clés:

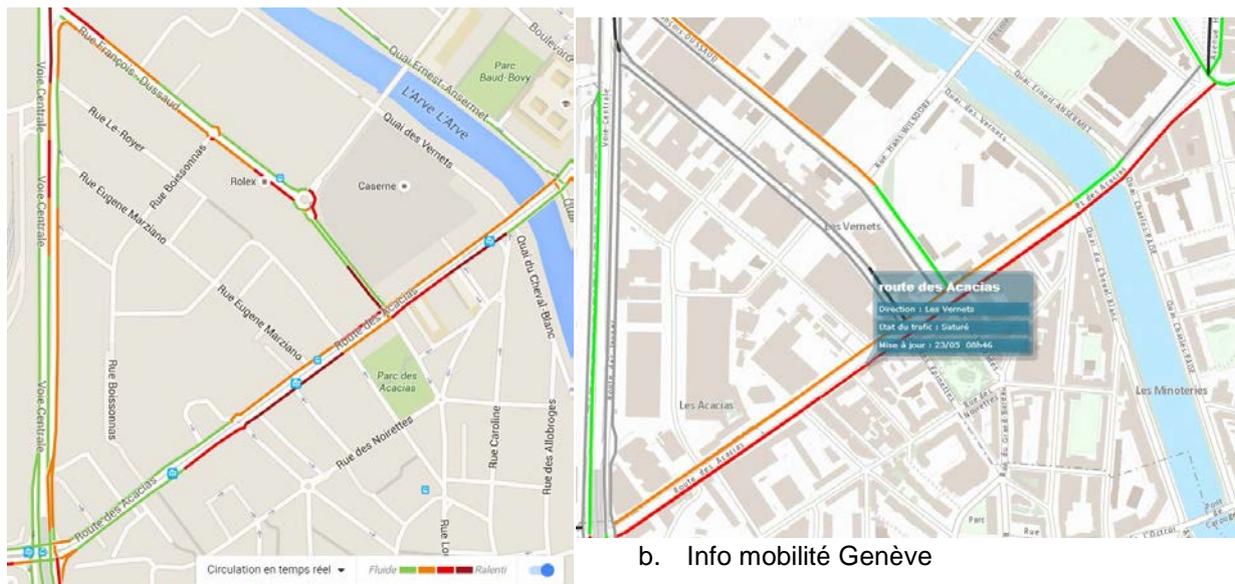
- **Problématique** et situation actuelle de la qualité des données ITS en Suisse et en Europe
 - o Quels sont les problèmes/enjeux majeurs dus à la qualité des données ?
 - o Quels sont les risques pour les fournisseurs de services ITS ?
 - o Quelles sont les difficultés (techniques, organisationnelles) pour gérer la qualité des données/services ITS ?
- **Moyens d'évaluation**, normes et bonnes pratiques, et organisation
 - o Quels sont les concepts de base et quelle méthodologie sont envisageables pour évaluer et gérer la qualité des données
 - o Quels sont les exemples d'évaluation de la qualité des données brutes ? (FCD, Bluetooth,...)
 - o Quels sont les normes, directives et bonnes pratiques développées en Suisse et en Europe ?
 - o Quel mode d'organisation faut-il mettre en place entre fournisseurs de données, prestataires de services (public, privé), autorités ?
- **Projets clés** pour la Suisse
 - o Qualité des données/service ITS et différentes échelles de gestion (routes nationales, routes cantonales, agglomérations)
 - o Réflexion sur un concept d'évaluation de qualité : modèle, méthodologie, organisation, responsabilité, coûts
 - o Impacts économique et sociaux avec des services ITS de qualité

Ce cinquième rapport complète le thème développé dans (Gilliéron, 2014) en se focalisant plus spécifiquement sur l'évaluation de la qualité de l'information trafic en temps réel. Il présente la problématique, quelques définitions de base, les principaux modes de calcul et les modes d'évaluation de la qualité.

1.2 Problématique

Les moyens de diffusion de l'information trafic en temps réel se sont multipliés ces dernières années, notamment avec l'avènement des systèmes de navigation qui intègrent ces services et le développement d'applications disponibles sur smartphone. Cette diversité des services d'information génère des résultats de qualité variable ainsi que des modes de représentation différents. Pour l'utilisateur, cette variabilité va impacter sa perception de l'état du trafic et l'interprétation de l'information, et par conséquent le choix d'options d'itinéraires alternatifs.

¹ http://www.its-ch.ch/fileadmin/redacteur/pdf/2_its-ch_bericht_16.05.2013_Verkehrsdaten_20130503_def_01.pdf



a. Carte Google

b. Info mobilité Genève



c. TomTom Live Traffic

Figure 1 : Exemples de représentation de la charge de trafic de plusieurs fournisseurs

Pour fournir les informations aux utilisateurs, les fournisseurs de services s'appuient sur différentes sources de données, sur des historiques et sur des éléments transmis par la communauté d'utilisateurs (Gilliéron, 2015). De nombreux paramètres entrent en jeu dans l'élaboration de l'information trafic, comme par exemple :

- L'emplacement des compteurs ou boucles de trafic
- Le type de mesures enregistrées (flux, vitesses, ...)
- Le recours à des données de véhicules traceurs (floating car data (FCD))
- Le niveau d'agrégation des données

Pour les données de base, les sources d'erreurs sont multiples et doivent être prises en compte dans l'ensemble du processus, voir dès la conception du système d'information. Il peut s'agir par exemple, d'une boucle de trafic dont l'emplacement n'est pas prévu initialement pour le comptage du trafic ou du pourcentage de véhicules traceurs présents dans le réseau au moment de l'estimation.

Le type d'information sur l'état du trafic joue également un rôle essentiel. Sur l'exemple de la figure 1, la carte Google (a) donne une indication sur la fluidité du trafic avec un code de couleurs, alors que l'information fournie par TomTom (c) est basée sur un temps de parcours additionnel le long d'un trajet. On constate également que le niveau d'agrégation de l'information est différent entre la carte Google (a) et la carte du service d'info-mobilité de l'Etat de Genève (c). De plus les codes de couleurs et la terminologie diffèrent (Info-mobilité : saturé, rouge ; Google : ralenti, rouge foncé ; TomTom : orange, trafic avec des à-coups). Cet exemple illustre bien la problématique de la qualité de l'information et de sa représentation.

2. Approche de la plateforme européenne pour les ITS (EIP)

La problématique de l'évaluation de la qualité des services d'information fait partie intégrante du plan européen d'Action sur les ITS, notamment dans le domaine consacré à l'usage optimal des données de la route, de la circulation et des déplacements. De nombreuses études ont été consacrées à ce domaine avec entre autre la publication de directives pour le déploiement de services ITS dans le cadre du projet Easyway, comme par exemple celle sur l'information des conditions de circulation et des temps de déplacement (EIP, 2015). Afin de soutenir cette démarche, une plateforme européenne pour les ITS (EIP) a été créée et ses membres ont publié une recommandation sur les exigences de qualité pour les services d'information sur la circulation. Ce travail se poursuit avec la démarche d'évaluation de la qualité en testant les recommandations dans la pratique avec les partenaires de l'EIP et les membres de TISA (Traveller Information Services Association).

Pour ce rapport, on prend comme référence le cadre défini par la plateforme européenne sur les ITS avec une publication récente sur le paquet qualité pour les services d'information sur la circulation en temps réel et en relation avec la sécurité (EIP, 2016). On tient compte également des directives émises par l'association TISA pour le calcul d'estimation de la qualité de l'information sur les flux de trafic (TISA, 2016 a, b).

2.1 Terminologie et définitions

La chaîne de valeur, qui va de la détection d'évènements à l'information à l'utilisateur, est décrite dans les documents de l'EIP et de TISA. La figure 2 montre la relation entre la partie « contenu » et la partie « service » où la notion de gestion de la qualité est primordiale pour que les résultats fournis aux usagers soient interprétés correctement.



Figure 2 : chaîne de valeur pour l'information sur la circulation (TISA, 2016)

Parmi les éléments de base proposés dans (EIP, 2016), on relève les définitions principales des critères pour :

- Les niveaux de service
 - o Couverture géographique : pourcentage du réseau routier couvert par le service
 - o Disponibilité : pourcentage du temps pendant lequel le service est disponible
- Les niveaux de qualité
 - o Rapidité (début) : délai entre l'apparition d'un évènement et la première détection validée de l'évènement (ex. information provenant d'une source fiable).
 - o Période de déclaration : intervalle de temps pour mettre à jour le statut des déclarations

- Rapidité (mise à jour) : délai entre la fin ou le changement (en relation avec la sécurité) de condition et la détection de cette fin ou de ce changement
- Latence (contenu): délai entre la première détection (si possible validée) d'un évènement et le moment où l'information est fournie par la partie « contenu » à la partie « service »
- Exactitude de position : précision relative de l'emplacement référencé pour l'évènement publié par rapport à l'emplacement effectif de l'évènement
- Exactitude de la déclaration : exactitude relative de la valeur déclarée (ex. vitesse, temps de parcours) par rapport à la valeur effective
- Taux d'erreur : pourcentage d'évènements publiés qui sont répertoriés comme incorrects
- Couverture d'évènement : pourcentage des évènements apparus et effectifs qui sont répertoriés comme ayant été correctement détectés et publiés
- Couverture de déclaration : pourcentage des emplacements de déclaration pour lesquels un statut de déclaration a été reçu pour toute période donnée de déclaration

Pour chacun de ces paramètres, l'EIP propose 3 niveaux (de base, amélioré, avancé) avec des éléments quantitatifs.

2.2 Méthodes d'évaluation de la qualité

L'EIP propose une dizaine de méthodes pour l'évaluation de la qualité que l'on décrit succinctement ci-dessous. En considérant l'ensemble des processus contenus dans la chaîne de valeur (Fig. 2), on comprend qu'il est difficile de développer une méthode universelle d'évaluation et que c'est par un découpage en sous-ensembles que l'EIP a proposé sa série de méthodes d'évaluation.

- Monitoring continu de la performance des équipements et leur disponibilité : cette méthode permet un monitoring en continu du fonctionnement des capteurs sur un réseau. Le but est d'obtenir des alertes sur les pannes ou mauvais fonctionnement d'équipements. Elle se situe en amont de la chaîne de valeur et devrait permettre d'écarter des mesures aberrantes.
- Vérification manuelle des évènements ou des conditions : elle permet une vérification manuelle de l'exactitude des évènements ou conditions reportés par rapport à la réalité effective. Le recours à une telle méthode demande du personnel ou des moyens techniques (ex. caméra de surveillance). Cette vérification s'applique à toute la chaîne de valeur.
- Tests de référence des données collectées : ces tests de référence sont utilisés pour vérifier la conformité des conditions de circulation, des temps de parcours et des informations sur les évènements. Les données ou informations à analyser sont comparées à des sources connues et fiables (« ground truth »). Cette comparaison est faite pour une période limitée et avec un échantillon de données restreint. Parmi les méthodes, on trouve les analyses quantitatives en fonction du temps et de l'espace (voir ci-dessous).
- Méthodes de tests de références orientées espace/temps : ce groupe de méthodes consiste à comparer des valeurs mesurées en fonction du temps et de l'espace à des valeurs de références (« ground truth »).
 - QKZ : « Qualitätskennzahl », approche macroscopique qui permet d'estimer les différences entre les informations de congestion fournies et celles effectivement détectées (ex. boucles) sur le réseau (BASt, 2011). La méthode QKZ fournit 2 indicateurs : QKZ1 qui est le taux (%) de détection et QKZ2 qui est le taux de fausses alarmes. La méthode est utilisée dans la chaîne de valeur depuis la détection de contenu jusqu'à la fourniture de service. Elle est utilisée par des sociétés commerciales comme TomTom.

- Monitoring de la complétude et de la latence des données : l'objectif de ce monitoring de la latence est d'évaluer le temps de traitement de l'information dans les centres de gestion ou d'information du trafic.
- Echantillonnage régulier de la complétude et de l'exactitude des contenus de messages ou données : vérification basée sur des échantillons et pour une durée donnée des contenus et des messages sur la circulation.
- Vérification et calibrage des pronostics des conditions de trafic et météo : cette méthode permet une vérification constante des pronostics des conditions de circulation et des conditions météo sur les routes. Ces pronostics sont systématiquement comparés aux conditions observées pendant les périodes considérées.
- Enquête de la perception de la qualité par les utilisateurs : le but d'une telle enquête est de mesurer la perception par les utilisateurs et leur expérience par rapport aux services d'information sur le trafic. Ces enquêtes sont réalisées avec une période de l'ordre d'une année.
- Récolte de feedback direct des usagers : cette récolte en direct peut se faire par plusieurs canaux par le fournisseur de service.
- Monitoring de statistiques d'usage opérationnel : cette méthode indirecte se base principalement sur le comptage du nombre d'accès au service.

Dans le document (EIP, 2016), les méthodes décrites ci-dessus sont croisées avec les différents critères de qualité afin de donner une vue d'ensemble du potentiel d'évaluation. En conclusion, on constate qu'aucune méthode ne permet de couvrir l'ensemble des critères et que la description des processus d'évaluation est très hétérogène et qu'il est difficile de valider les exigences de qualité en se basant sur ces approches.

Face à cette complexité, le recours à des indicateurs de qualité basés sur une approche quantitative s'avère pertinent et permet avec objectivité d'évaluer certains critères. C'est l'approche préconisée par TISA et qui est résumée dans le chapitre suivant.

3. Calcul de l'indicateur QBench de TISA

L'association TISA a publié récemment une directives (TISA, 2016 a) et des spécifications de calculs (TISA, 2016 b). Ces directives décrivent les métriques pour déterminer la qualité de l'information des flux de trafic.

3.1 Principe du QBench

La méthode QBench est principalement basée sur la mesure de temps de parcours. Elle permet d'évaluer l'exactitude des retards déclarés par rapport aux retards effectifs vécus par les usagers. La méthode peut s'appliquer aux différentes étapes de la chaîne de valeur. Une série de métriques sont définies pour ces calculs d'estimation, comme le temps de parcours en flux libre (free flow), le temps de parcours déclaré. Ces métriques, principalement basée sur les vitesses et les temps de parcours, permettent de dériver les valeurs du QBench. On ne tient pas compte d'autres informations comme la déclaration d'évènements ou des valeurs de densité du trafic.

Le calcul du QBench estime le bénéfice « réel » pour l'utilisateur. C'est le rapport entre la somme des bénéfices de temps effectifs par rapport à la somme des bénéfices de temps idéaux. Le bénéfice de temps idéal est le temps de parcours « ground truth » moins le temps de parcours en flux libre (free flow).

3.2 Comparaison à des valeurs de référence

Pour déterminer les indices de QBench, il faut récolter des données de référence (ground truth) de manière indépendante et fiable afin de calculer les métriques nécessaires (vitesse, temps de parcours). Cette récolte de données peut se faire avec des mesures tests de trajectoires de véhicules.

Il suffit d'utiliser un *logger* GPS qui permet d'enregistrer des positions à intervalles réguliers et de les dater. Le calcul du temps de parcours le long d'un segment s'obtient par la différence entre le temps à l'entrée du segment et celui à la sortie. Le niveau d'agrégation doit être compatible avec la définition des segments routiers utilisés pour la génération de l'information sur la circulation (par exemple les links TMC). Pour que ces tests soient représentatifs, il faut choisir les routes à parcourir suivant les conditions de trafic dans la journée. Il faut évidemment que le nombre de tests soit suffisant pour que les résultats soit représentatifs d'un point de vue statistique (Brouwer, 2014).

Avec le déploiement des applications mobiles et des périphériques équipés de GPS (smartphone, navigateur), on peut envisager une récolte automatisée sur le principe des véhicules traceurs (FCD). Pour que les résultats soient représentatifs, il faut que le taux de véhicules équipés soit adapté et que ces données ne soient pas utilisées pour générer les informations sur la circulation. Si la flotte répond à ces critères, elle permet une estimation fiable à l'échelle détaillée d'un réseau routier. Toutefois, il ne faut pas sous-estimer les opérations nécessaires pour éliminer les données aberrantes issues des véhicules traceurs.

4. Conclusion

Ce rapport a présenté une partie du développement des méthodes d'évaluation de la qualité des informations sur le trafic qui s'insère dans la démarche d'harmonisation des services ITS en Europe. On constate qu'il n'y a pas de méthode universelle, mais qu'il est indispensable de considérer l'ensemble de la chaîne de valeur et d'y associer des méthodes d'évaluation fiables et basées sur une approche quantitative. L'avènement des véhicules traceurs ou des applications basées sur la localisation, offre un potentiel important pour mieux estimer les conditions de trafic et la détection d'incidents. Toutefois, il faut rester attentif, car ces données sont, pour la plupart, gérées par des acteurs privés qui cherchent à en tirer des bénéfices et qui ne vont pas forcément fournir tous les détails sur les sources.

5. Bibliographie

Bast (2011), Qualität von on-trip Verkehrsinformationen im Strassenverkehr, Bast Kolloquium 23.-24.03.2011, Bericht der Bundesanstalt für Strassenwesen, Heft F 82

Brouwer Jeroen (2014), Measuring real-time traffic data quality based on floating car data, TomTom

EIP-European ITS Platform (2015), Traveller Information Services, TRAFFIC CONDITION AND TRAVEL TIME INFORMATION, Deployment Guideline, <https://dg.its-platform.eu/DGs2012>

EIP-European ITS Platform (2016), Quality package for safety related and real-time traffic information services, EIP+ 3.1 Testing and Validating of the Quality Recommendations and Results for Traffic Information from EIP, www.easyway-its.eu

Gilliéron, P.-Y. (2014), Qualité et avalanche de données : est-ce compatible ?, Route et Trafic, pp 33-35, Vol. 03-2014

Gilliéron, P.-Y. (2015), Crowdsourcing pour la mobilité et les transports, Route et Trafic, pp 40-43, Vol 07-08-2015

TISA (2016, a), Guidelines for TISA QBench Calculations

TISA (2016, b), Specification for TISA QBench Calculations

Texas Transportation Institute (2011), Guidelines for evaluating the accuracy of travel time and speed data

Rakha H. et al (2011), Assessment of Data Quality Need for use in Transportation Applications, Virginia Polytechnic Institute and State University, University of Maryland

Liens Internet

Portland State University, portail d'information trafic
<http://portal.its.pdx.edu/home> , consulté le 31.03.2016

TomTom traffic Index
http://www.tomtom.com/fr_ch/trafficindex/ , consulté le 31.03.2016

Houston, Transtar
<http://traffic.houstontranstar.org/layers/> , consulté le 31.03.2016

ITS Action Plan
http://ec.europa.eu/transport/its/road/action_plan/action_plan_en.htm

6. Annexe

Chronologie des activités

Cette partie reprend les thèmes identifiés dans le premier rapport et définit les priorités et les actions à entreprendre. Le thème sur le **monitoring de la qualité** est en première priorité et quelques activités liées à ce thème sont proposées.

Priorités

Cette partie propose quelques actions possibles selon les thématiques identifiées et fixe des priorités pour leurs études

[Tableau 1] Liste des propositions de thèmes et actions

Thème	Types d'actions	Priorité
<i>Monitoring de la qualité</i>	- Concepts, normalisation - Développement d'outils	1
<i>Certification des données</i>	- Aspects organisationnels et juridiques - Développement d'outils et procédures	2
<i>Anonymisation des données</i>	- Développement d'outils - Développement de moyens de contrôle	3
<i>Lutte contre la fraude</i>	- Aspects organisationnels et juridiques - Développement d'outils de diagnostic - Développement de concepts	3

Activités sur le monitoring de la qualité (**rapport 1 du 16.5.13**)

Parmi les activités envisagées sur ce premier thème, on peut citer les actions à court et moyen termes :

- Organisation d'une journée (*its-ch* - Tagung) en automne 2014 avec le développement d'une thématique autour des données du trafic et de la mobilité (programme en cours de préparation)
- Participation à des rencontres nationales et internationales sur le thème de l'acquisition et de la gestion des données du trafic (par exemple : séminaire IFSTTAR, SETRA, CERTU du 23.05.13 sur le véhicule traceur comme capteur de données trafic).
- Etablissement d'un groupe de travail entre la VSS-FK9 et l'OFROU pour la définition de projets de recherche sur la qualité de données liée à l'information trafic
- Inventaire des principaux projets, directives et normes sur la qualité des données et leur évaluation
- Identification des principaux acteurs concernés et représentatifs des enjeux de ce thème
- Coordination du thème « données des transports » avec les autres thèmes prioritaires

Développement dans le rapport 2 (20.11.2013)

Le 2^e rapport a permis de développer le thème du monitoring de la qualité des données ITS avec les parties suivantes :

Résumé

La multiplicité des sources de données sur les déplacements permet à de nombreux acteurs de la mobilité de proposer de nouveaux services qui sont diffusés sur l'Internet. Face à ces nouveaux canaux d'informations, les responsables de réseaux de transports et d'infrastructures doivent gérer le trafic, prendre des décisions et informer les usagers. Ces opérations ne peuvent être établies que sur la base d'informations fiables. Ce second article propose de développer le thème du monitoring de la qualité des données et services ITS en fonction de la situation des normes et projets en Suisse et en Europe. Il propose des initiatives ou idées de projets en abordant les questions méthodologiques, organisationnelles et de normalisation.

Situation et développement

- Problématique et situation actuelle de la qualité des données ITS en Suisse et en Europe
- Moyens d'évaluation et de contrôle, normes et bonnes pratiques, organisation
- Projets clés pour la Suisse
- Contexte du déploiement des services ITS
- Qualité des données et des services dans les ITS
- Etat du développement
 - Norme ISO, projet

Méthodologie

- Approche du projet Quantis (www.quantis-project.eu)

Bibliographie

Gilliéron, P.-Y. (2013), Big Data : un enjeu majeur pour gérer le trafic et la mobilité, Route et Trafic, pp 31-33, Vol. 09-2013

Gilliéron, P.-Y. (2014), Qualité et avalanche de données : est-ce compatible ?, Route et Trafic, pp 33-35, Vol. 03-2014

Développement dans le rapport 3 (14.05.2014)

Le 3^e rapport a permis de développer le thème de **l'intégrité du positionnement** dans les ITS et le transport routier. Ce rapport contient les parties suivantes :

Résumé

Les données de trafic et de mobilité sont essentielles pour le déploiement des services ITS. Parmi ces données, le rôle de la localisation n'est plus à démontrer dans les applications du transport terrestre. Avec la grande diffusion des systèmes de navigation basée sur le GPS, on a le sentiment que la question de la localisation est complètement réglée. Or, ce n'est de loin pas le cas pour les applications exigeantes qui nécessitent un haut niveau de confiance dans la localisation. Cet article fait le point sur les développements GNSS et les actions en cours pour normaliser et certifier les terminaux de localisation pour le transport terrestre.

Situation et état de l'art

- Développement des systèmes de navigation par satellites
- Etat de la normalisation
- Action COST SaPPART

Problématique de la qualité du positionnement

- Localisation par rapport aux exigences des ITS
- Caractéristiques des erreurs et notion d'intégrité
- Concept d'intégrité

Conclusion

Bibliographie

Gilliéron, P.-Y. (2014), Intégrité du positionnement dans le transport routier, Route et Trafic, pp 14-17, Vol. 09-2014

Développement dans le rapport 4 (7.05.2015)

Le 4^e rapport a permis de développer le thème du **crowdsourcing** pour la mobilité et les transports. Ce rapport contient les parties suivantes :

Résumé

Les données de trafic et de mobilité sont essentielles pour le déploiement des services ITS. L'avènement des nouvelles technologies de la communication et de l'informatique offre un fort potentiel pour la saisie et la transmission d'informations liées à la mobilité. L'usage de ces technologies permet à chacun de collecter des informations et de les transmettre dans une optique de production participative (crowdsourcing) pour des services de mobilités ou de gestion des transports. Ce quatrième rapport propose de préciser le concept de crowdsourcing et son usage dans les services d'information et de gestion des transports. Il présente la problématique du crowdsourcing avec ses éléments techniques, quelques exemples d'applications, quelques aspects juridiques et ainsi que les perspectives pour le domaine des transports.

Concept et principe

- Concept de la production participative
- Principe de fonctionnement

Application à la mobilité et aux transports

- Services basés sur la localisation
- Principales applications
- Questions juridiques et statut des données

Synthèse et perspectives

Bibliographie

Gilliéron, P.-Y. (2015), Crowdsourcing pour la mobilité et les transports, Strasse und Verkehr, vol. 07-08 2015, num. 07-08, p. 32-35, 2015.