

Towards the Ranking of Important Smart Vehicles in VANETs - An Information-centric Approach

Junaid Ahmed Khan*[†] Yacine Ghamri-Doudane[†] and Ali El Masri*[†]

*University Paris-Est, LIGM Lab, Marne-la-Vallée, France

[†]L3i Lab, University of La Rochelle, France

khan@univ-mlv.fr, yacine-ghamri@univ-lr.fr, ali.elmasri@u-pem.fr

Abstract—Vehicles today are equipped with various sensors and cameras to gather and share tremendous amount of heterogeneous data from urban streets. The challenge is to identify the appropriate vehicles from the fleet of vehicles, important for the collection, distribution and storage of such massive data. This paper addresses the autonomous identification of such "Smart Vehicles" without relying on the infrastructure network. Therefore, we propose "InfoRank" as an Information-centric algorithm for a vehicle to first rank different location-aware information associated to it. It then uses the information importance to analytically find its relative importance in the network. InfoRank is the first step towards identifying the best information hubs to be used in the network for the efficient collection, storage and distribution of urban sensory information. Results from scalable simulations using realistic vehicular mobility traces show that InfoRank is an efficient ranking algorithm to find popular information facilitator vehicles in comparison to other ranking metrics in the literature.

Keywords—Réseau ad hoc de véhicules ; réseau centré information ; centralité

I. INTRODUCTION

Les véhicules modernes sont munis de composants électroniques, capteurs, caméras et équipements de communications qui servent à assister et améliorer la sécurité et le confort de nos voyages. Dans un environnement urbain, un véhicule intelligent peut désormais être considéré comme une instance de l'Internet des objets (IoT, Internet of Things) qui vise à collecter et partager des données provenant des différents endroits de la ville. Ce collecte/partage de données est nécessaire dans les différentes applications des systèmes de transport intelligents (ITS, Intelligent Transportation System) telles que la gestion du trafic routier et la surveillance urbaine [1][2].

Ces applications exigent que les véhicules génèrent, d'une façon continue, une énorme quantité de données. En outre, la plupart des données générées sont de "pertinence locale" (les utilisateurs intéressés se trouvent dans le réseau véhiculaire). Par conséquent, l'utilisation d'un réseau d'infrastructure pour collecter et partager ces données de type *Big-Data* est coûteux et inadéquate. Une solution pour éviter le passage à travers un réseau d'infrastructure est d'utiliser les communications véhicule-à-véhicule (V2V, Vehicle-to-Vehicle). D'autre part, la diffusion de toutes les informations captées par les véhicules produit un surcoût de communication massive (*overhead*) et une énorme redondance d'information dans le réseau. Par conséquent, le principal défi pour le succès et le passage à l'échelle des applications ITS consiste à mettre en place un système

efficace pour le collecte/partage des informations captées par les véhicules (la forte mobilité et la connectivité intermittente d'un réseau ad hoc de véhicules (VANET, Vehicular Ad hoc NETWORK) doivent être prises en considération).

Dans ce contexte, nous nous intéressons à l'identification des véhicules capables de se comporter comme centres d'informations dont le rôle est de collecter les informations du réseau, partager les informations collectées, et les distribuer aux véhicules intéressés. Pour identifier des tels nœuds, une solution type est d'utiliser les méthodes de mesure de centralité telles que *Degree*, *Closeness*, *Betweenness* et *Eigenvector*. Cependant, ces méthodes ne sont pas appropriées à la nature sporadique de la topologie d'un réseau véhiculaire. Pour résoudre ce problème, dans cet article, nous proposons, pour la première fois dans les réseaux véhiculaires, un algorithme de classement, appelé InfoRank, qui permet à un véhicule de mesurer sa centralité (se classer), et donc, évaluer son importance dans le réseau [3].

InfoRank évalue l'importance d'un véhicule en se basant sur l'importance des informations provenant des différents lieux visités par le véhicule. Pour cette raison, InfoRank utilise une approche centrée information (ICN, Information-Centric Networking [4]). ICN est un paradigme de communication qui vise à remplacer les réseaux IP actuels (ICN est récemment utilisé pour le routage dans les VANETs [5][6]). Dans ICN, un nœud diffuse son intérêt pour un contenu/information en utilisant le nom de ce contenu. Chaque nœud du réseau peut répondre, à cet intérêt, s'il possède le contenu demandé, ce qui rend le contenu indépendant d'un nœud (adresse) spécifique dans le réseau. Ceci est réalisé en permettant aux nœuds intermédiaires de stocker les informations lors du transfert des intérêts/réponses entre une source et une destination.

Les principales contributions de cet article peuvent être résumées comme suit :

- Une première métrique qui permet de calculer l'importance d'une information en fonction de (1) la pertinence de l'information vis-à-vis des intérêts des véhicules du réseau, (2) la validité de l'information à l'instant actuel et (3) la distance entre le véhicule et le lieu où l'information est captée ;
- Une deuxième métrique qui permet de calculer la centralité (importance) d'un véhicule en fonction de (1) l'importance des informations provenant des différents lieux visités par le véhicule et (2) la couverture du véhicule (l'ensemble des lieux fréquemment visités par le véhicule) ;

- Un algorithme de classement (InfoRank) qui utilise les deux métriques proposées pour permettre à un véhicule de se classer (évaluer son importance) d'une façon autonome.

Les performances d'InfoRank sont évaluées par simulation en utilisant des traces réelles de trois milles véhicules. Les résultats obtenus montrent la capacité d'InfoRank à identifier les véhicules importants dans le réseau ainsi que son aptitude à passer à l'échelle et sa conformité aux principes d'ICN.

La suite de l'article est organisée comme suit. La section II mentionne les travaux connexes à nos contributions. Nous présentons l'algorithme InfoRank et les deux métriques proposées dans la section III. Dans la section IV, nous évaluons par simulation les performances d'InfoRank. Une conclusion et des perspectives sont données dans la section V.

II. TRAVAUX EXISTANTS

L'utilisation des méthodes de mesure de centralité pour identifier les nœuds importants a été considérée dans plusieurs applications telles que les réseaux sociaux et les sciences médicales [7]. Par exemple, l'algorithme PageRank de Google [8] utilise une méthode de mesure de centralité pour évaluer l'importance d'une page web lors d'une recherche sur Internet. Les principales méthodes utilisées sont *Degree*, *Closeness*, *Betweenness* et *Eigenvector*.

Dans la méthode *Degree*, la centralité d'un nœud N est mesurée en se basant sur le nombre de voisins directs de N (voisins à un saut de N). La méthode *Closeness* mesure la centralité du nœud N en considérant la longueur du chemin le plus court de N vers chaque nœud du réseau ; la centralité du N est l'inverse de la somme de toutes les longueurs. La méthode *Betweenness* considère le chemin le plus court entre chaque source-destination dans le réseau ; la centralité de N est fonction du nombre de chemins courts qui passent par N . Dans la méthode *Eigenvector*, la centralité de N est fonction de son influence dans le réseau [9]. Récemment, en combinant plusieurs de ces méthodes, ML-SOR [10] sélectionne les nœuds importants pour la diffusion de données et le routage dans les réseaux mobiles opportunistes.

Cependant, les méthodes mentionnées ci-dessus ne peuvent pas être utilisées dans les réseaux véhiculaires. Ceci est dû à plusieurs raisons : (1) ces méthodes considèrent des topologies statiques (la forte mobilité des véhicules rend la topologie VANET très dynamique) ; (2) le calcul de centralité dans des méthodes telles que *Closeness*, *Betweenness* et *Eigenvector* utilisent des paramètres globaux du réseau (de telles informations peuvent ne pas être disponibles à bord d'un véhicule dans un VANET) ; et (3) ces méthodes utilisent des métriques telles que le chemin le plus court entre deux nœuds du réseau (des tels chemins sont instables dans les VANETs). Par conséquent, il est nécessaire de développer des méthodes de mesure de centralité qui seront spécialement conçues aux caractéristiques spécifiques des réseaux véhiculaires.

III. L'ALGORITHME INFORANK

Dans cette section, (1) nous modélisons le réseau véhiculaire en utilisant trois graphes, (2) nous proposons une première métrique qui permet de calculer l'importance d'une

information, (3) nous proposons une deuxième métrique qui permet de calculer la centralité (l'importance) d'un véhicule et (4) nous présentons l'algorithme InfoRank qui utilise les deux métriques proposées pour permettre à un véhicule de se classer d'une façon autonome.

A. Modèle du réseau

Dans cette sous-section, (1) nous modélisons le réseau véhiculaire en utilisant trois graphes (topologie du VANET, plan de la ville et *véhicule-information*), (2) nous présentons la nomination ICN utilisée pour référer à une information et (3) nous illustrons la structure temporelle utilisée dans InfoRank.

Les graphes: nous modélisons la topologie d'un VANET par un graphe non-orienté $G(\mathbb{V}(t), \mathbb{E}^v(t))$, où $\mathbb{V}(t) = \{v\}$ est l'ensemble des sommets v représentant les véhicules à un instant t et $\mathbb{E}^v(t) = \{e_{jk}(t) \mid v_j, v_k \in \mathbb{V}, j \neq k\}$ est l'ensemble des arêtes $e_{jk}(t)$ représentant les liens de communication entre les véhicules ($e_{jk}(t)$ représente le lien de communication entre les véhicules j et k à l'instant t). Le nombre des arêtes dans l'ensemble $\mathbb{E}^v(t)$ dépend de la portée de communication de chaque véhicule. D'autre part, nous modélisons le plan de la ville par un graphe non-orienté $G(\mathbb{X}, \mathbb{E}^x)$, où $\mathbb{X} = \{x\}$ est l'ensemble des sommets x représentant les lieux (quartiers) de la ville et $\mathbb{E}^x = \{e_{pq} \mid x_p, x_q \in \mathbb{X}, p \neq q\}$ est l'ensemble des arêtes représentant les routes (ou ensemble de routes) reliant les lieux de la ville. Enfin, nous définissons le graphe *véhicule-information* comme un graphe biparti $G(\mathbb{V}, \mathbb{X}, \mathbb{E})$, où \mathbb{V} est l'ensemble des sommets v (véhicules) du graphe modélisant la topologie du VANET $G(\mathbb{V}(t), \mathbb{E}^v(t))$ et \mathbb{X} est l'ensemble des sommets x (lieux) du graphe modélisant le plan de la ville $G(\mathbb{X}, \mathbb{E}^x)$. L'ensemble des arêtes $\mathbb{E} = \{e_{ij} \mid v_i \in \mathbb{V}, x_j \in \mathbb{X}\}$ associe un ensemble de lieux $X_v \subset \mathbb{X}$ à chaque véhicule v (ceci signifie que le véhicule v possède des informations provenant de l'ensemble de lieux X_v). Dans cet article, nous utilisons les termes *information* et *lieu* d'une façon interchangeable. Cette utilisation est dû au fait qu'une information est toujours captée dans un lieu bien spécifié.

La nomination ICN: afin d'identifier une information, nous proposons la nomination ICN suivante "*région/section-route/type-information*". La partie *section-route* de la nomination est généralement définie par la municipalité de la ville alors que la partie *type-information* est spécifiée par les différentes applications ITS (sécurité routière, congestion de la route, *infotainment*, etc.). Enfin, la partie *région* de la nomination ICN est spécifiée en utilisant la *partitionnement de Voronoï* [11], où les véhicules qui se trouvent dans une même zone sont considérés dans la même régions de Voronoï $x \in G(\mathbb{X}, \mathbb{E}^x)$

La structure temporelle: le temps est divisé en intervalles de longueur constante $\bar{t} = t_{k+1} - t_k$. L'intervalle débute à l'instant t_k et se termine à l'instant t_{k+1} . Chaque véhicule calcule sa centralité à l'instant t_{k+1} .

B. Importance des informations

Dans cette sous-section, nous proposons une nouvelle métrique qui permet de calculer l'importance d'une information. L'importance d'une information est fonction de trois paramètres : (1) la pertinence de l'information vis-à-vis des intérêts des véhicules du réseau, (2) la validité de l'information

à l'instant actuel et (3) la distance entre le véhicule et le lieu où l'information est captée.

Définition 1 (Fréquence de satisfaction des intérêts): afin de mesurer la pertinence d'une information vis-à-vis des intérêts des véhicules du réseau, nous définissons $I_x^v(\bar{t}) = \frac{r_x(\bar{t})}{R_x}$ comme étant la fréquence des intérêts qui ont été satisfaits (ont eu des réponses positives) durant l'intervalle précédent \bar{t} , où $r_x(\bar{t})$ est le nombre des réponses positives aux intérêts demandant l'information x durant cet intervalle et R_x est le nombre total des réponses positives aux intérêts demandant l'information x .

Définition 2 (Validité d'une information): la validité d'une information est fonction du temps (par exemple, les informations de la congestion de la route sont valables uniquement durant la congestion). Par conséquent, afin de mesurer la validité d'une information à l'instant actuel, nous définissons $\tau(t_{k+1}) = \begin{cases} 1 & t_{k+1} \leq t_x^f + \bar{t}_d \\ e^{-\delta \bar{t}_d} & t_{k+1} > t_x^f + \bar{t}_d \end{cases}$, où δ est un paramètre d'ajustement dépendant des besoins temporels de l'application concerné, t_x^f est le temps de la dernière réponse positive à un intérêt demandant l'information x et \bar{t}_d est la moyenne des durées de validité spécifiées par les intérêts (nous supposons que chaque intérêt spécifie une durée dans laquelle l'information est considérée valable). $\bar{t}_d = \frac{1}{n} \sum_n t_x$, où n est le nombre des intérêts reçus durant l'intervalle précédent et t_x est la durée de validité spécifiée par chaque intérêt.

Si à un certain instant, (1) il n'y a aucun intérêt actif et (2) la moyenne des durées de validité spécifiées par les intérêts est expirée, $\tau(t_{k+1})$ suit une diminution exponentielle. Pour les informations dont leur nature exige qu'elles seront toujours disponibles dans le réseau, $\tau(t_{k+1})$ est fixée à 1.

Distance d'une information: Nous définissons la distance d'une information $d(x, x_k)$ comme étant la norme Euclidienne entre le lieu $x \in X_v$ où l'information est captée et le lieu $x_k \in X_v$ du véhicule à l'instant t_{k+1} . Nous supposons que chaque véhicule possède tout le graphe $G(\mathbb{X}, \mathbb{E}^x)$ vu que le plan de la ville est statique et disponible auprès de la municipalité. Cependant, un véhicule possède juste une partie du graphe $G(\mathbb{V}, \mathbb{X}, \mathbb{E})$ car il est assez difficile pour un véhicule de traverser toutes les routes de la ville. Les informations d'un véhicule se divisent en deux parties : (1) les informations captées par le véhicule lui-même (i.e., il est le véhicule source de ces informations) et (2) les informations transférées par ce véhicule (i.e., il était un véhicule intermédiaire lors de la transmission de ces informations entre une source et une destination).

En se basant sur les définitions précédentes, l'importance d'une information (première métrique proposée) est calculée comme suit :

$$C_x^v(t_{k+1}) = C_x^v(t_k) + \tau(t_{k+1}) I_x^v(\bar{t}) (1 + d(x, x_k))^{-\lambda} + s_x^v(t_{k+1}) \quad (1)$$

Où $C_x^v(t_{k+1})$ est l'importance de l'information x à l'instant t_{k+1} , λ est un paramètre d'ajustement dépendant de l'application ITS, et $0 \leq s_x^v(t_{k+1}) \leq 1$ est le pourcentage de temps durant lequel le véhicule se comporte comme source

de l'information x ($s_x^v(t_{k+1})$ est un indice de la fiabilité du véhicule par rapport à l'information, le véhicule v est considérée plus fiable s'il est la source de l'information x). Nous pouvons observer de l'équation 1 que : (1) l'importance de l'information x à l'instant t_{k+1} ($C_x^v(t_{k+1})$) dépend de son importance à l'instant t_k ($C_x^v(t_k)$) et (2) si aucun intérêt demandant l'information x n'est reçu durant l'intervalle précédent, $I_x^v(\bar{t})$ sera égale à 0, ce qui évite toute augmentation inutile de l'importance de l'information.

C. Centralité des véhicules

Dans cette sous-section, nous proposons une nouvelle métrique qui permet de calculer la centralité (l'importance) d'un véhicule. La centralité d'un véhicule est fonction de deux paramètres : (1) l'importance des informations provenant des différents lieux visités par le véhicule (cette importance est mesurée en utilisant l'équation 1) et (2) la couverture du véhicule (cette couverture représente l'ensemble des lieux fréquemment visités par un véhicule et elle est mesurée en utilisant l'entropie de couverture définie ci-dessous).

Définition 4 (Entropie de couverture): Nous définissons $H^v = - \sum_{x \in \mathbb{X}} p(x) \log p(x)$ comme étant l'entropie de couverture d'un véhicule par rapport au plan de la ville (sous-graphe $X_v \in G(\mathbb{X}, \mathbb{E}^x)$), où $p(x)$ est la probabilité qu'un lieu $x \in \mathbb{X}$ sera visité par le véhicule v avant le prochain instant de calcul t_{k+1} .

Afin de mieux comprendre cette définition, nous considérons que la couverture d'un véhicule est représentée par un ensemble de lieux. Nous supposons donc que $M^A = \{x_3, x_2, x_2, x_3, x_5, x_2\}$ et $M^B = \{x_1, x_2, x_3, x_5, x_4, x_1\}$ sont les couvertures des véhicules A et B , respectivement. M^A indique que le véhicule A visite les lieux x_2, x_3 et x_5 avec les probabilités $\frac{2}{6}, \frac{3}{6}$ et $\frac{1}{6}$, respectivement, alors que M^B indique que le véhicule B visite les lieux x_1, x_2, x_3, x_4 et x_5 avec les probabilités $\frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}$ et $\frac{1}{6}$, respectivement. L'entropie de couverture de chaque véhicule est alors calculée comme suit :

$$\begin{aligned} H^A &= -\frac{2}{6} \log \frac{2}{6} - \frac{3}{6} \log \frac{3}{6} - \frac{1}{6} \log \frac{1}{6} = 0.439, \\ H^B &= -\frac{2}{6} \log \frac{2}{6} - \left(\frac{1}{6} \log \frac{1}{6} * 4\right) = 0.639, \end{aligned}$$

Les résultats montrent que le véhicule A a une couverture plus étroite (A se déplace dans une zone géographique limitée), alors que le véhicule B a une couverture plus large (B se déplace dans une zone géographique plus large).

En se basant sur les définitions précédentes, la fonction de centralité d'un véhicule est calculée comme suit :

$$f_I^v(t_{k+1}) = \frac{(1 + I_m^v(t_{k+1}))^{-\epsilon}}{|X_v|} \sum_{x \in X_v} C_x^v(t_{k+1}) \cdot w_x + H^v(t_{k+1}) \quad (2)$$

Où $f_I^v(t_{k+1})$ est la centralité du véhicule v à l'instant t_{k+1} , $C_x^v(t_{k+1})$ est l'importance de l'information x à l'instant t_{k+1} , $I_m^v(t_{k+1})$ est le taux des intérêts manqués par le véhicule (le véhicule n'a pas pu répondre positivement à ces intérêts, le taux des intérêts manqués mesure la capacité d'un véhicule à répondre positivement aux intérêts reçus), $|X_v|$ est la cardinalité du sous-graphe $X_v \subset \mathbb{X}$, ϵ est le paramètre d'ajustement, et $w_x = \frac{R_x}{R_T}$ (où R_x est le nombre des réponses

positives aux intérêts demandant l'information x et R_T est le nombre total des réponses positives (à tous les intérêts)).

La centralité d'un véhicule (deuxième métrique proposée) est mise à jour à l'instant t_{k+1} en utilisant une moyenne glissante comme suit :

$$C_v(t_{k+1}) = (1 - \theta)C_v(t_k) + \theta f_I^v(t_{k+1})$$

où θ est un paramètre d'ajustement.

D. InfoRank

L'algorithme 1 (InfoRank) représente les étapes permettant à un véhicule de se classer d'une façon autonome. Les étapes de cet algorithme représentent une application des métriques proposées ci-dessus.

Algorithm 1 InfoRank

INPUT : $G(\mathbb{V}, \mathbb{X}, \mathbb{E})$: graphe *véhicule-information*
OUTPUT : la centralité du véhicule à l'instant t_{k+1}
for chaque véhicule $v \in \mathbb{V}$ **do**
 for chaque information $x \in X_v$ **do**
 Calculer $d(x, x_k), \tau(t_{k+1}), s_x^v(t_{k+1}), w_x$
 Calculer $I_x^v(\bar{t}) \leftarrow \frac{r_x(\bar{t})}{R_x}$
 if $I_x^v(\bar{t}) \neq 0$ **then**
 Mettre à jour $C_x^v(t_{k+1})$ en utilisant l'équation 1
 else
 $C_x^v(t_{k+1}) = C_x^v(t_k) + s_x^v(t_{k+1}),$
 end if
 end for
 Calculer le taux des intérêts manquées $I_m^v(t_{k+1})$
 Calculer l'entropie de couverture $H^v(t_{k+1})$
 Calculer $f_I^v(t_{k+1})$ en utilisant l'équation 2
end for
return $C_v(t_{k+1})$

IV. ÉVALUATION DE L'ALGORITHME INFORANK

Pour évaluer InfoRank, nous l'avons comparé avec les méthodes *Degree*, *Closeness*, *Betweenness* et *Eigenvector*. Les simulations ont été réalisées en utilisant *Network Simulator-3* (NS-3) [12]. Nous avons utilisé des traces de mobilité réelles de la ville de Cologne en Allemagne. À notre connaissance, ces traces sont considérées les plus précises parmi les traces de mobilité dans les réseaux véhiculaires [13]. Les paramètres de simulation sont présentés dans la table I. Chaque simulation est répétée cinq fois en modifiant l'ensemble des couples source-destination (appelée par la suite producteur-consommateur).

Table I: Paramètres de simulation

Paramètre	Valeur
Simulateur	NS-3
Nombre de nœuds	2986
Origine des traces de mobilité	Cologne, Allemagne
Zone	6 x 6 km ² centre-ville
Durée de simulation	1 heure
Portée radio	100 m
Taille d'un paquet	1024 bytes
Précision de temps	1 sec
Nombre de simulations	5

Table II: Les 10 véhicules les mieux classés en utilisant InfoRank dans cinq différentes simulations.

Simulation	1		2		3		4		5		Moyenne
	ID	Score	ID	Score	ID	Score	ID	Score	ID	Score	
1	34	1	1013	1	59	1	1210	1	10	1	1
2	2414	0.9719	1051	0.9966	63	0.9984	113	0.9930	56	0.9957	0.9911
3	1239	0.9639	543	0.9958	108	0.9969	14	0.9923	1195	0.9838	0.9865
4	108	0.9612	442	0.9919	348	0.9967	2372	0.9883	127	0.9833	0.9843
5	517	0.9567	2366	0.9808	392	0.9958	184	0.9881	577	0.9801	0.9803
6	414	0.9557	318	0.9788	152	0.9917	157	0.9875	658	0.9754	0.9778
7	31	0.9516	2728	0.9770	48	0.9890	189	0.9834	5	0.9749	0.9752
8	571	0.9506	268	0.9738	26	0.9839	197	0.9808	54	0.9742	0.9727
9	502	0.9501	54	0.9735	37	0.9821	196	0.9803	1888	0.9718	0.9716
10	2174	0.9489	150	0.9729	137	0.9819	30	0.9789	192	0.9714	0.9708

A. Scénario de simulation

Nous avons utilisé le module ndnSIM [14] de NS-3 pour intégrer l'architecture *Named Data Networking* (NDN). Le scénario suivant a été implémenté :

Producteur: un véhicule producteur est une source d'informations dans le réseau, où les informations fournies sont les lieux visités par le véhicule avant l'instant de calcul de centralité.

Consommateur: un véhicule consommateur est un utilisateur souhaitant visiter un lieu. Ce véhicule diffuse un intérêt pour les informations du lieu à visiter. L'intérêt est acheminé vers le véhicule producteur en utilisant l'architecture ICN.

Nous avons considéré l'application de la surveillance urbaine. Le plan de la ville est divisé en zones/lieux en utilisant la *tessellation de Voronoï* [11]. Les producteurs s'annoncent en tant que sources d'informations pour les lieux visités. Les consommateurs diffusent des intérêts aux informations des lieux à visiter. Les intérêts suivent une distribution Zipf où les intérêts aux informations populaires sont plus fréquents. Les paramètres d'ajustement λ , δ et ϵ sont fixés à 0.5. La moyenne des durées de validité des intérêts \bar{t}_d est fixée à 1 heure.

B. Résultats de simulation

L'objectif de notre simulation est de répondre à la question suivante : *Est-ce que l'algorithme InfoRank est capable de bien identifier les véhicules importants ?*

La table II présente les 10 véhicules les mieux classés en utilisant InfoRank dans cinq différentes simulations. Pour chaque véhicule, nous présentons l'ID et le score de centralité. La moyenne pour chaque niveau de classement, présentée dans la dernière colonne, se situe dans un intervalle de confiance de 0.01 pour un niveau de confiance de 95%. Les scores sont normalisés par rapport au meilleur score (le meilleur score est fixé à 1). La même normalisation est utilisée dans la suite de l'article. Nous observons que, dans la première simulation, le véhicule 34 a eu le meilleur score. Ce score signifie que, par rapport aux autres véhicules, le véhicule 34 répond plus aux intérêts reçus. Par conséquent, ce véhicule est classé comme le véhicule le plus important dans le réseau.

La comparaison entre InfoRank et les autres méthodes de mesure de centralité (*Degree*, *Closeness*, *Betweenness* et *Eigenvector*) est réalisée comme suit. Pour chaque simulation, nous sélectionnons les 5 véhicules les mieux classés en utilisant InfoRank et chacune des autres méthodes, et par la suite, nous considérons les paramètres suivants :

- Nombre des intérêts satisfaits par véhicule sélectionné;
- Stabilité de l'ensemble des véhicules sélectionnés ;

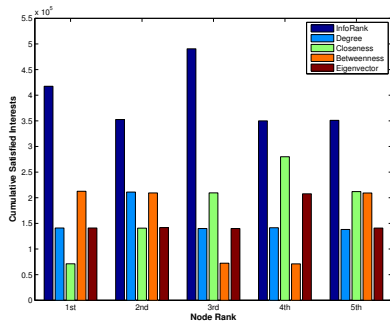


Figure 1: Nombre des intérêts satisfaits par les 5 véhicules les mieux classés en utilisant InfoRank et les autres méthodes de mesure de centralité.

- Débit par véhicule sélectionné ;
- Taux de succès de la mémoire cache dans l'ensemble des véhicules sélectionnés.

1) *Nombre des intérêts satisfaits*: La figure 1 présente le nombre des intérêts satisfaits, durant une heure, par les 5 véhicules les mieux classés, en utilisant InfoRank et les autres méthodes de mesure de centralité (les résultats affichés représentent la moyenne sur cinq simulations). Nous observons que, en utilisant InfoRank, plus d'intérêts sont satisfaits. Ceci est dû essentiellement au fait que les autres méthodes ne considèrent pas le taux de satisfaction des intérêts lors du calcul de la centralité d'un véhicule. Au contraire, dans InfoRank, ce taux de satisfaction est pris en compte par la *fréquence de satisfaction des intérêts* définie dans la section III. Cette fréquence est un facteur clé dans le calcul de l'importance d'une information, et par la suite, l'importance d'un véhicule.

2) *Stabilité de l'ensemble des véhicules les mieux classés*: En raison des changements très rapides dans la topologie d'un réseau véhiculaire, il est nécessaire d'analyser la stabilité, par rapport au temps, de l'ensemble des 5 véhicules les mieux classés. Plus précisément, nous allons exécuter le classement des véhicules à différents instants (une durée de 15 minutes sépare deux instants consécutifs), et nous allons étudier, si l'ensemble des 5 véhicules les mieux classés vont rester les mêmes après chaque classement ou vont être remplacés par d'autres véhicules. La figure 2 présente les 5 véhicules les mieux classés à quatre instants différents. La normalisation des scores de centralité est effectuée pour chaque méthode séparément. Une première observation des résultats montre la différence entre les classements des différentes méthodes.

Nous observons aussi qu'en utilisant InfoRank, l'ensemble des véhicules les mieux classés est presque stable (des petites remplacements ont eu lieu entre deux classement consécutifs). Au contraire, en utilisant les autres méthodes, la plupart de ces véhicules sont remplacés par d'autres véhicules après chaque classement. En effet, ces méthodes calculent la centralité d'un véhicule en fonction de la topologie (par exemple, les chemins les plus courts), ce qui nécessite à un véhicule d'avoir une vue globale du réseau. Une telle vue n'est pas possible dans un réseau très dynamique tel qu'un réseau véhiculaire. D'autre part, InfoRank n'est pas affecté par la nature dynamique du réseau véhiculaire vu que son calcul de centralité se base sur des métriques plus stables.

Enfin, nous observons que dans le premier classement de

la méthode *Eigenvector*, un seul véhicule atteint un score de centralité élevé. Ceci montre que la méthode *Eigenvector* n'est pas capable d'attribuer un score de centralité élevé à plusieurs véhicules dans un VANET, alors que l'algorithme InfoRank est capable d'éviter un tel comportement indésirable.

3) *Débit par véhicule*: La figure 3 présente le débit de chacun des 5 véhicules les mieux classés en utilisant InfoRank et les autres méthodes de mesure de centralité. Le débit présenté est la moyenne sur un ensemble de cinq simulations. Nous observons que l'algorithme InfoRank est plus performant que les autres méthodes ; le débit des 5 véhicules les mieux classés en utilisant InfoRank est supérieur à celui des véhicules classés en utilisant les autres méthodes. Encore une fois, l'impact de la considération de l'importance d'une information dans le calcul de l'importance d'un véhicule est observé. En effet, les véhicules classés en utilisant InfoRank possèdent des informations plus populaires.

D'autre part, nous pouvons observer qu'en utilisant InfoRank, le débit du véhicule classé en troisième position est légèrement supérieur au débit des autres véhicules. En effet, dans InfoRank, l'importance d'un véhicule n'est pas uniquement basée sur le nombre des intérêts reçus (d'autres paramètres sont considérés).

4) *Évaluation de la conformité aux principes d'ICN - taux de succès de la mémoire cache*: Nous évaluons la conformité aux principes d'ICN en considérant le taux de succès de la mémoire cache. En effet, ICN permet aux nœuds intermédiaires de stocker les informations lors du transfert des intérêts/réponses entre une source et une destination. Ceux-ci retournent les informations stockées dans le réseau (d'une façon distribuée) au lieu qu'elles soient stockées dans un seul nœud (d'une façon centralisée). Pour mesurer la capacité de l'algorithme InfoRank à stocker les informations d'une façon distribuée, nous considérons le taux de succès de la mémoire cache dans l'ensemble des 5 véhicules les mieux classés. Ce taux est présenté dans la figure 4. Nous considérons qu'une information est stockée dans la mémoire cache d'un véhicule si ce dernier répond positivement à deux intérêts demandant cette information. Les résultats montrent que les véhicules classés, en utilisant InfoRank, stockent plus d'informations en comparant avec les autres méthodes. Encore une fois, ces résultats expliquent l'intérêt de la considération de l'importance d'une information dans le calcul de l'importance d'un véhicule. Les véhicules possédant des informations importantes répondent, et par la suite, stockent plus par rapport aux autres véhicules. Ces résultats montrent que le stockage distribué d'ICN offre un accès plus efficace aux informations dans un réseau véhiculaire caractérisé par une forte mobilité et une connectivité intermittente.

5) *Discussion*: En revenant à la question qui représente le but de notre simulation : *Est-ce que l'algorithme InfoRank est capable de bien identifier les véhicules importants ?* Les résultats de simulation ont montré, qu'au contraire des autres méthodes de mesure de centralité, en utilisant InfoRank, l'ensemble des véhicules les mieux classés est presque stable (ces véhicules sont généralement des bus, taxi, etc.). En outre, les résultats ont montré que les véhicules les mieux classés en utilisant InfoRank sont capables de satisfaire plus les intérêts des autres véhicules dans le réseau. Ils sont capables aussi de stocker plus d'informations dans leurs mémoires. Par

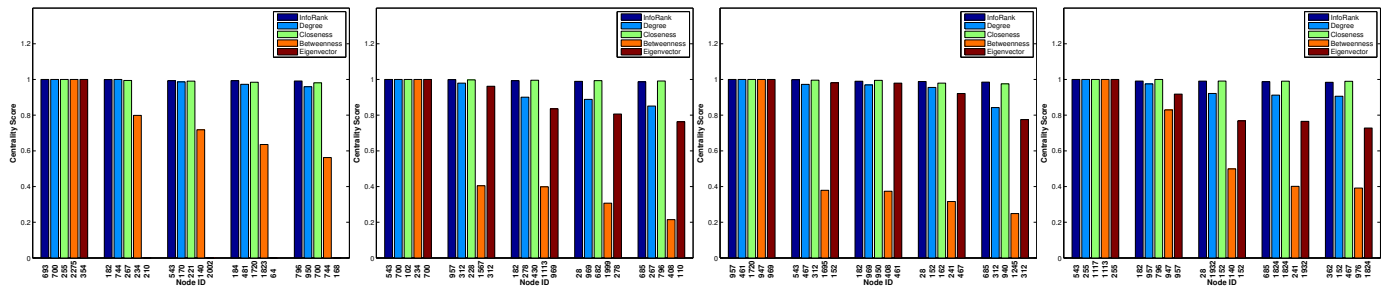


Figure 2: Les 5 véhicules les mieux classés à différents instants en utilisant InfoRank et les autres méthodes de mesure de centralité (une durée de 15 minutes sépare deux instants consécutifs).

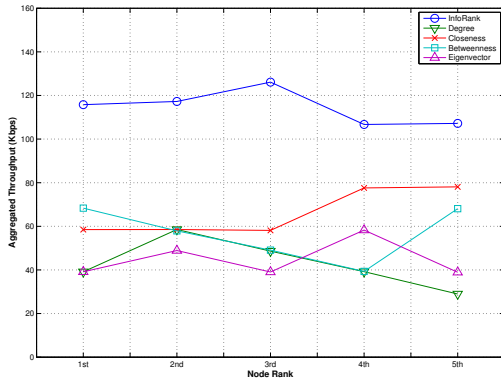


Figure 3: Le débit des 5 véhicules les mieux classés en utilisant InfoRank et les autres méthodes de mesure de centralité.

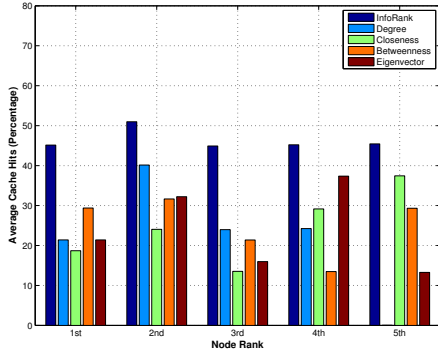


Figure 4: Le taux de succès de la mémoire cache dans les 5 véhicules les mieux classés en utilisant InfoRank et les autres méthodes de mesure de centralité.

conséquent, oui, InfoRank est capable de **bien** identifier les véhicules importants dans un réseau véhiculaire.

V. CONCLUSIONS ET PERSPECTIVES

Dans cet article, nous avons proposé InfoRank, le premier algorithme de classement des véhicules dans les réseaux véhiculaires. InfoRank est un algorithme distribué qui permet à un véhicule de se classer (mesurer son importance/centralité) d'une façon autonome. InfoRank est une approche centrée sur les informations qui se compose de deux métriques : la première métrique permet de calculer l'importance des informations alors que la deuxième permet de calculer l'importance d'un véhicule. Plusieurs paramètres sont pris en compte dans

le calcul des métriques proposées tels que la pertinence des informations aux intérêts des véhicules du réseau, la validité temporelle des informations et la couverture géographique des véhicules. Les résultats de simulation ont montré, qu'en comparant avec les méthodes de mesure de centralité traditionnelles, InfoRank est plus approprié aux réseaux véhiculaires ; il peut identifier un ensemble stable de véhicules qui sont capables de satisfaire plus les intérêts des véhicules du réseau.

L'algorithme InfoRank représente une première étape vers le succès et le passage à l'échelle des applications ITS. Les véhicules les mieux classés par InfoRank doivent se comporter comme des centres capables de collecter et partager les données dans un réseau véhiculaire. Le développement d'une approche de collecte/partage efficace de données est l'objectif de nos travaux futurs.

REFERENCES

- [1] U. Lee, B. Zhou, M. Gerla, E. Magistretti, P. Bellavista, and A. Corradi, "Mobeyes: smart mobs for urban monitoring with a vehicular sensor network," *Wireless Communications, IEEE*.
- [2] R. Bruno and M. Nurchis, "Efficient data collection in multimedia vehicular sensing platforms," *Pervasive and Mobile Computing*, 2014.
- [3] J. A. Khan, Y. Ghamri-Doudane, and D. Botvich, "Inforank: Information-centric autonomous identification of popular smart vehicles," in *IEEE Vehicular Technology Conference, VTC*, 2015.
- [4] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman, "A survey of information-centric networking," *Communications Magazine, IEEE*, vol. 50, no. 7, pp. 26–36, 2012.
- [5] M. Amadeo, C. Campolo, A. Molinaro, and G. Ruggeri, "Content-centric wireless networking: A survey," *Computer Networks*.
- [6] M. Gerla, E.-K. Lee, G. Pau, and U. Lee, "Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds," in *Internet of Things (WF-IoT), World Forum on*. IEEE, 2014, pp. 241–246.
- [7] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, "Identification of influential spreaders in complex networks," *Nature Physics*, vol. 6, no. 11, pp. 888–893, 2010.
- [8] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web." 1999.
- [9] S. P. Borgatti, "Centrality and network flow," *Social networks*, vol. 27, no. 1, pp. 55–71, 2005.
- [10] A. Socievole, E. Yoneki, F. De Rango, and J. Crowcroft, "MI-sor: Message routing using multi-layer social networks in opportunistic communications," *Computer Networks*, vol. 81, pp. 201–219, 2015.
- [11] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*. John Wiley & Sons, 2009, vol. 501.
- [12] "Network simulator 3 - <http://www.nsnam.org>."
- [13] S. Uppoor, O. Trullols-Cruces, M. Fiore, and J. M. Barcelo-Ordinas, "Generation and analysis of a large-scale urban vehicular mobility dataset," *Mobile Computing, IEEE Transactions on*, vol. 13, no. 5, pp. 1061–1075, 2014.
- [14] A. Afanasyev, I. Moiseenko, and L. Zhang, "ndnSIM: NDN simulator for NS-3," NDN, Technical Report NDN-0005, October 2012. [Online]. Available: <http://named-data.net/techreports.html>