Dependability modelling and evaluation of vehicular applications based on mobile ad-hoc networks. Modélisation et évaluation de la sûreté de fonctionnement d'applications véhiculaires basées sur des réseaux ad-hoc mobiles

Ossama Hamouda

▶ **To cite this version:**

Ossama Hamouda. Dependability modelling and evaluation of vehicular applications based on mobile ad-hoc networks. Modélisation et évaluation de la sûreté de fonctionnement d'applications véhiculaires basées sur des réseaux ad-hoc mobiles. Computer Science [cs]. Université Paul Sabatier - Toulouse III, 2010. English. NNT : . tel-00546260

HAL Id: tel-00546260

https://theses.hal.science/tel-00546260

Submitted on 14 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

DÉLIVRÉ PAR *Université Paul Sabatier*

Discipline ou spécialité :

Systèmes Informatiques Critiques

Présentée et soutenue par *Ossama Mohamed Fawzi HAMOUDA*

Titre : *Modélisation et évaluation de la sûreté de fonctionnement d'applications véhiculaires basées sur des réseaux ad-hoc mobiles*

Dependability modelling and evaluation of vehicular applications based on mobile ad-hoc networks

Le *19 Juillet 2010*

JURY

M. Bruno CICIANI	*Professeur de l'Université de Rome*	*Italie*
M. Guy JUANOLE	*Professeur de l'Université Paul Sabatier, Toulouse*	*France*
M. Mohamed KAÂNICHE	*Directeur de Recherche du LAAS-CNRS, Toulouse*	*France*
Mme. Karama KANOUN	*Directeur de Recherche du LAAS-CNRS, Toulouse*	*France*
M. Raymond MARIE	*Professeur de IFSIC/IRISA, Rennes*	*France*
M. Hans-Peter SCHWEFEL	*Professeur de l'Université d'Aalborg*	*Danemark*

Ecole doctorale :	*Systèmes*
Unité de recherche :	*LAAS-CNRS*
Directeurs de Thèse :	*M. Mohamed KAÂNICHE, Mme. Karama KANOUN*
Rapporteurs :	*M. Bruno CICIANI, M. Raymond MARIE*
Président :	*M. Guy JUANOLE*

THÈSE

en vue de l'obtention du

DOCTORAT DE L'UNIVERSITE DE TOULOUSE

délivré par l'**Université Toulouse III – Paul Sabatier**

Ecole Doctorale : **Systèmes**

Discipline : **Systèmes Informatiques Critiques**

présentée et soutenue

par

Ossama HAMOUDA

le 19/07/2010

Titre :

Modélisation et évaluation de la sûreté de fonctionnement d'applications véhiculaires basées sur des réseaux ad-hoc mobiles

Dependability modelling and evaluation of vehicular applications based on mobile ad-hoc networks

Directeur de thèse : **M. Mohamed KAÂNICHE**

Co-directeur de thèse : **Mme. Karama KANOUN**

JURY

M. Bruno CICIANI	Rapporteur
M. Raymond MARIE	Rapporteur
M. Guy JUANOLE	Examinateur
M. Mohamed KAÂNICHE	Examinateur
Mme. Karama KANOUN	Examinateur
M. Hans-Peter SCHWEFEL	Examinateur

Dedicated to my parents, Mohamed Fawzi Hamouda and Aziza Hanbouta (late)…

Acknowledgements

This dissertation presents my PhD thesis work carried out at the *Laboratoire d'Analyse et d'Architecture des Systèmes* of the French National Center for Scientific Research (LAAS-CNRS), a research laboratory located in Toulouse, France. I would like to thank Malik Ghallab and Raja Chatila, Directors of LAAS-CNRS since my entry, for welcoming me in this laboratory. This work was done in the Dependable Computing and Fault Tolerance research group (TSF). I would like to thank Jean Arlat and Karama Kanoun, the successive leaders of this research group for having given me the chance to join the group and for their support.

I would like to thank Bruno Ciciani (Professor, University of Roma, Italy) and Raymond Marie (Professor, IFSIC/IRISA, Rennes France) for kindly accepting to review my dissertation and for being part of the jury and I would also like to express my gratitude to Hans-Peter Schwefel (Associate Professor, Aalborg University, Denmark) and Guy Juanole (Professor Emeritus, University of Toulouse, France) for being part of my jury. I value the insightful and comprehensive feedback of the reviewers as well as the jury members on my dissertation.

This thesis was partially financed by the Arab Academy for Science and Technology and Maritime Transport (AASTMT), Egypt. Many thanks to Gamal El Din Mokhtar and Mohamed Farghaly, Presidents of the AASTMT, Ossama Badawy, Ossama Ismaïl, Ahmed Fahmy, Gamal Sleem, Lubna Sherif, and Mohamed Abu El-Dahab, professors at the AASTMT, since my start, for giving me the approval of the scholarship to do my PhD in France.

I state my sincere thanks to my supervisors Mohamed Kaâniche (*Directeur de recherche*, LAAS-CNRS) and Karama Kanoun (*Directeur de recherche*, LAAS-CNRS) for supporting me during the three years and half of my PhD, of having been present when necessary, for reviewing my work in perfect manner, and for the trust they gave me. In fact, their advice and criticism have undoubtedly contributed greatly to this work.

Mohamed Kaâniche, most probably, I am the most stubborn and hard to tutor student you have ever worked with. I will not thank you only for your help, but also for your friendship. You have always been available, understanding and easy to talk to.

Thanks to all the members of the Dependable Computing and Fault Tolerance research group: permanents, PhD students, and interns. Amongst them, I felt to be at home and I have much appreciated the family-like atmosphere. I am very pleased to mention David Powell and Marc-Olivier Killijian for their time and for their very interesting remarks. Yves Crouzet, the angel of our group, deserves special thanks for his great services and for his availability even during the most difficult moments. I would also like to take the chance to thank Géraldine Vache. You have been always there for me whenever I needed any help. Even when it is hard for me to express myself in French, you have always had the time to wait and listen. My thanks go also to my colleagues Rim Akrout, Kossi Tiassou, Eric Alata and Ana-Elena Rugina, with whom I shared the office during my stay at LAAS and Gina Briand for her kindness and her help on administrative issues. I wish also to thank the staff members of the postgraduate school EDSYS (*École Doctorale Systèmes*): Sophie Achte, Caroline Bérard, Agnan Bonneval, Hélène Thirion, Véronique Villemur.

My warm thanks go to my Egyptian friends: Mohamed Gad-El-Rab, Ahmed Akl, Ahmed Ali, Heba Shaarawy, Mahmoud Mostafa, Ussama Zaghlol, Hesham Kotb, El Awaday Attia, and Hany Gamal. Their presence around me (especially Mohamed Gad-El-Rab and A. Akl) gave me a lot of enthusiasm and the desire to work more.

I never forget to express my sincere thanks to Ludovic Courtès (*Ingénieur de Recherche*, INRIA) and Erling Matthiesen Møller (PhD student, Aalborg Universitet) for their great and enormous support and efforts during my PhD.

To Mohamed Fawzi Hamouda (Dad), David Hamouda (uncle), Jeanne Hamouda (uncle's wife), Ahmed Hamouda (brother), for all my life, I have wanted you to be proud of me, not because you are difficult to please, but because you have always been my role models. You have been my great supporters; I wish I can be as half as good as you are with me.

To my sisters and brothers, you have always been my best family. When I was young, I hated it, but now, I understand that you have always cared. You will always be my best friend.

Last but not the least; I would like to express my gratitude to my friends Wala'a El-Sawah, Mohamed El Masri, Sandy Rahme and Layale Saab. They have been a constant and tireless source of support for me and their support played a key role in enabling me to achieve this significant milestone in my life. I am really lucky to have you all.

I thank Tarek Elbaumyi (father-in-law), and Houda Ismaïl (mother-in-law) for their advices and support they gave me all the time.

I cannot thank my wife enough as she is the invisible heroine of the history. We have lived together the last days before my defense moment by moment: those of hope as well as the times of disappointments. She has tolerated all my defects that were multiplied in the last days before defending the thesis.

Finally, I would like to say that I came here thanks to Allah (God) Almighty who provided me the health and the power to pursue research in the exciting domain of dependable computing systems in this strange and beautiful country. It is only with His help and His guidance that I have accomplished this milestone.

Again, thank you all.

Résumé en Français

"An expert is a man who has made all the mistakes, which can be made, in a very narrow field."
Niels Bohr, Danish physicist, 1885-1962

Introduction

L'évolution des technologies de communication sans fil a engendré de nouvelles applications dans l'informatique mobile, qu'il s'agisse par exemple de la domotique, la santé, la protection de l'environnement ou encore de l'automobile ou des transports en général. Ces applications sont basées sur la présence de processeurs embarqués dans des objets physiques : téléphones, ordinateurs, appareils photo..., pouvant dialoguer entre eux via des réseaux *ad-hoc* utilisant des communications sans fil ou accéder à des services via des infrastructures fixes. Pour que les services fournis par ces applications soient adoptés par le grand public, la question de leur sûreté de fonctionnement et de la confiance qu'on peut leur porter est essentielle. En effet, une période d'indisponibilité de service peut conduire à une situation catastrophique en fonction de la durée d'interruption de service et de la criticité. Pour y répondre, plusieurs aspects sont à prendre en compte dans la conception et la validation de ces applications : mobilité des objets et des utilisateurs, défaillances d'origine accidentelle ou malveillante, déconnexions fréquentes, autonomie limitée, restrictions de bande passante et de traitement

L'objectif de cette thèse est de développer des méthodes et des modèles permettant d'évaluer des mesures quantitatives caractérisant la sûreté de fonctionnement de services mobiles telle qu'elle est perçue par les utilisateurs. Ces modèles et mesures permettront de fournir aux concepteurs des éléments de décision objectifs pour l'analyse et la sélection des architectures les plus adaptées pour satisfaire les exigences de sûreté de fonctionnement.

Cette problématique est nouvelle dans le contexte des applications et services mis en œuvre sur des dispositifs mobiles utilisant des réseaux ad-hoc, et peu de résultats existent dans ce domaine. La plupart des travaux menés autour de l'évaluation de services mobiles concerne l'évaluation des performances et cible particulièrement les protocoles de communication. Notre objectif est de nous focaliser sur les couches applicatives en ciblant les mécanismes de tolérance aux fautes.

Pour atteindre ces objectifs, deux aspects essentiels doivent être abordés : 1) décrire les caractéristiques liées à la mobilité et analyser leur impact sur la sûreté de fonctionnement des applications envisagées en tenant compte différents environnements et scénarios de mobilité, et 2) trouver des moyens efficaces pour maîtriser la complexité des modèles à la fois pour leur construction et leur traitement.

Les concepteurs d'applications et de systèmes sont généralement intéressés par la compréhension et la quantification de l'impact des défaillances à différents niveaux de décomposition des systèmes considérés. Une seule technique n'est généralement pas suffisante pour effectuer ces évaluations, et il est généralement nécessaire de combiner différentes techniques et plusieurs outils. Cela est dû en particulier à la complexité de l'analyse et à la nécessité de prendre en compte les différentes caractéristiques des systèmes cibles qui ne peuvent être caractérisés de manière efficace par une seule méthode (la description des scénarios de mobilité, la modélisation

de défaillances de composants et de stratégies de restauration de service, etc.). La composition de modèles utilisant conjointement des techniques analytiques (méthodes combinatoires, chaînes de Markov, réseaux de Petri stochastiques, etc.) et des techniques de simulation est une approche intéressante pour maîtriser la complexité des modèles. Ce type de méthodes a été privilégié dans l'approche de modélisation et les cas d'études étudiés dans de cette thèse. En particulier, l'évaluation quantitative des mesures de sûreté de fonctionnement est basée sur l'utilisation combinée : i) de modèles analytiques basés sur les réseaux d'activités stochastiques (SAN) [Meyer *et al*. 1985, Sanders & Meyer 2001], ii) de techniques de simulation et iii) du traitement statistique de traces de mobilité. Les deux dernières techniques sont utilisées pour déterminer les caractéristiques de connectivité dans différents scénarios de mobilité. Les paramètres issus de ces analyses et les distributions de probabilité associées sont utilisés comme entrées dans des modèles basés sur des réseaux d'activités stochastiques permettant l'évaluation de mesures de sûreté de fonctionnement à un niveau supérieur d'abstraction des applications et systèmes cibles. Les réseaux d'activités stochastiques (SAN) sont bien adaptés pour maîtriser la complexité des modèles grâce à la composition et la réplication de sous-modèles.

Nous avons examiné trois cas d'études dans nos travaux, nommément un service de réplication dans le domaine *ad-hoc*, une application boîte noire virtuelle, et une application de conduite automatisée utilisant le principe de *platooning*. Ces cas d'études présentent des caractéristiques de sûreté de fonctionnement qui sont typiques des applications du domaine de l'automobile utilisant des communications de Véhicules-à-Véhicules (*Car-to-Car communications*–C2C) et des communications pour connecter des véhicules à une infrastructure fixe (*Car-to-Infrastructure communications*–C2I).

Ces travaux ont été effectués en partie dans le cadre du projet européen IST FP6 HIDENETS *HIghly DEpendable ip-based NETworks and Services* [HIDENETS 2006a]. Ce projet porte sur le développement et la validation d'architectures et de méthodes permettant d'assurer la sûreté de fonctionnement d'applications et de services, mis en œuvre dans un contexte mobile avec des réseaux ad-hoc, en considérant comme domaine privilégié l'automobile.

Ce mémoire de thèse est structuré comme suit. Le premier chapitre présente les motivations et le contexte du travail de thèse. Nous y présentons une discussion sur la modélisation de la sûreté de fonctionnement en commençant par les concepts principaux de la sûreté de fonctionnement. Nous présentons également un bref aperçu de l'évaluation de la sûreté de fonctionnement et les méthodes qui peuvent être utilisées pour évaluer les applications mobiles. Une approche de modélisation hiérarchique élaborée dans le cadre du projet HIDENETS pour évaluer la sûreté de fonctionnement de bout-en-bout de ces applications est également présentée. La dernière partie de ce chapitre présente des informations générales sur les applications du domaine de l'automobile utilisant des technologies sans fil et examine les défis qu'ils représentent du point de vue d'évaluation de la sûreté de fonctionnement. Une brève description des trois cas d'études présentées dans les chapitres suivants de cette thèse est également incluse. Nous finissons ce chapitre en donnant un bref aperçu des sujets abordés dans le mémoire. L'état de l'art et les travaux connexes sont discutés dans le contexte de chaque chapitre.

Le deuxième chapitre est consacré à l'estimation de certaines caractéristiques reflétant la connectivité entre véhicules dans les systèmes véhiculaires qui utilisent les réseaux sans fil (*ad-hoc*). Nos analyses concernent plus particulièrement la distribution du temps entre deux rencontres, en utilisant des preuves analytiques, des simulations, et du traitement statistique de

traces de mobilité issues de l'observation de trafic réel. Ces distributions sont utilisées dans la modélisation de la sûreté de fonctionnement des trois cas d'études qui sont présentés dans la thèse. Les trois chapitres suivants illustrent respectivement la modélisation et l'évaluation de sûreté de fonctionnement de nos trois cas d'études. Enfin, le chapitre six conclut sur les contributions de cette thèse et identifie les pistes de recherche futures.

Chapitre 1 : Contexte et problématique

Dans ce chapitre nous présentons le contexte et les objectifs des travaux développés dans ce mémoire de thèse. Nous introduisons dans un premier temps quelques définitions liées à la sûreté de fonctionnement, en nous concentrant en particulier sur l'évaluation quantitative des propriétés de sûreté de fonctionnement. Ensuite, nous décrivons les caractéristiques principales des applications véhiculaires qui utilisent des communications C2C et C2I, et les défis soulevés par ces applications du point de vue de l'évaluation de la sûreté de fonctionnement. Enfin, nous présentons les thèmes principaux et les contributions de la thèse.

1.1 Sûreté de fonctionnement et évaluation

La sûreté de fonctionnement est un domaine de recherche actif, avec ses propres concepts et terminologie [Avizienis *et al.* 2004, Laprie *et al.* 1995]. Elle est définie comme la propriété d'un système permettant à ses utilisateurs de placer une confiance justifiée dans le service qu'il leur délivre. Nos travaux portent sur l'évaluation de mesures permettant de quantifier des propriétés de sûreté de fonctionnement telles que la fiabilité, la disponibilité, ou la sécurité-innocuité pour des applications de l'automobile utilisant des communications de type C2C et C2I. Nous nous intéressons plus particulièrement aux méthodes d'évaluation basées sur des modèles qui sont bien adaptées pour aider les concepteurs dans la sélection d'architectures qui répondent au mieux aux exigences de sûreté de fonctionnement dès la phase de conception.

1.2 Contexte : applications véhiculaires basées sur des réseaux ad-hoc mobiles

Les applications étudiées dans le cadre de la thèse sont destinées à être mises en œuvre dans des véhicules équipés d'interfaces réseau sans fil qui peuvent communiquer avec d'autres véhicules via des communications de véhicule-à-véhicule notées C2C ou avec un ensemble de serveurs sur les infrastructures fixes en utilisant des communications de type C2I, via des points d'accès disponibles répartis géographiquement sur le réseau routier. Les scénarios considérés peuvent être basés sur des communications C2C uniquement, des communications C2I, ou bien sur une combinaison hybride des deux types de communications.

Plusieurs applications utilisant des communications de type C2C et C2I peuvent être envisagées. En effet, différents types de service peuvent être offerts aux utilisateurs incluant des services de loisir (messagerie électronique, audio et vidéo à la demande, consultation de guides touristiques et de serveurs d'information en ligne, etc.) ainsi que des services plus critiques (supervision de trafic, détection de dangers et avertissement des usagers, etc.).

Nous avons considéré trois cas d'études le cadre de cette thèse : 1) un service de réplication de données sur des réseaux ad-hoc, 2) un système d'autoroute automatisée, et 3) une boîte noire virtuelle basée sur la réplication et la sauvegarde coopérative des données. Dans les chapitres

suivants, nous décrivons ces cas d'études et les modèles associés permettant d'évaluer la sûreté de fonctionnement. Un aspect délicat de la modélisation concerne la maîtrise de la complexité au niveau de la construction et du traitement des modèles. La complexité peut résulter de : 1) la représentation fine des interactions entre composants, tant matériels que logiciels, 2) le nombre élevé de composants et d'états à considérer explicitement dans le modèle, et 3) la prise en compte de plusieurs niveaux de dégradation de service. Les méthodes de modélisation hiérarchiques et modulaires constituent une piste intéressante pour la maîtrise de la complexité des modèles. Ce type de méthodes a été privilégié dans le cadre de la thèse.

1.3 Contributions

Deux contributions principales sont présentées dans la thèse.

La première concerne la caractérisation des scénarios de mobilité et l'estimation des paramètres de connectivité pour les communications C2C et C2I. Différentes techniques sont utilisées, notamment la simulation, des modèles de connectivité analytiques, et le traitement statistique des traces de connectivité disponibles sur Internet Les paramètres de connectivité d'intérêt correspondent par exemple au taux de rencontre des véhicules, à la durée pendant laquelle ces véhicules restent à portée de la connectivité et au taux auquel un véhicule est à portée d'un point d'accès à l'infrastructure fixe.

La deuxième contribution concerne le développement de modèles qui permettent la description et l'évaluation quantitative de l'impact des défaillances des composants et des restaurations sur la sûreté de fonctionnement au niveau du système, en tenant compte des informations issues des analyses de connectivité pour différents scénarios de mobilité. La construction et le traitement des modèles de sûreté de fonctionnement sont basés sur le formalisme SAN, et l'outil associé Möbius [Sanders & Meyer 2001, Dali *et al.* 2000].

Chapitre 2 : Estimation de paramètres de connectivité dans les réseaux ad-hoc sans fils véhiculaires

Ce chapitre a pour objectif d'étudier la distribution de certains paramètres caractérisant la connectivité entre les véhicules dans les réseaux véhiculaires sans fil *ad-hoc,* en utilisant des preuves analytiques, des simulations et des traces issues de mesures dans un contexte opérationnel. Les résultats de ce chapitre sont utilisés pour l'évaluation de la sûreté de fonctionnement des applications présentées dans les chapitres suivants.

2.1 Introduction

Les dernières années, un effort important a porté sur le développement de protocoles et de systèmes nécessaires à la réalisation d'applications véhiculaires qui utilisent des communications C2C et C2I. La qualité du service et la sûreté de fonctionnement fournies par ces applications dépendent fortement des caractéristiques de connectivité entre les véhicules pour les communications C2C et entre un véhicule et une infrastructure fixe pour les communications C2I. Les paramètres de connectivité qui nous intéressent dans le cadre de nos travaux incluent en particulier le taux avec lequel les connexions peuvent avoir lieu et la durée des périodes de connectivité. Ces paramètres dépendent fortement de l'environnement de mobilité considéré

(autoroute, zone urbaine-ville, etc.). La plupart des travaux existants sur l'étude des paramètres de connectivité sont basés sur des modèles mathématiques et des simulations en raison des difficultés de mener des expériences dans un environnement réel sur route. Dans ce chapitre, nous nous sommes concentrés sur l'analyse des processus stochastiques décrivant les instants durant lesquels un véhicule entre dans la zone de transmission d'un autre véhicule (il peut donc communiquer avec lui) ou d'un point d'accès de l'infrastructure fixe.

Nous avons utilisé trois techniques complémentaires pour l'analyse des paramètres de connectivité : des modèles analytiques, la simulation, et le traitement de traces de mobilité disponibles sur Internet.

2.2 Paramètres de connectivité

Pour analyser la sûreté de fonctionnement des applications véhiculaires examinées dans nos travaux, nous nous sommes intéressés à l'étude de la distribution de probabilité relative aux trois variables aléatoires suivantes caractérisant les processus de rencontre de type C2C ou C2I.

- *Le temps entre deux rencontres C2C* : Il s'agit de l'intervalle de temps pendant lequel un véhicule est hors de portée de transmission d'un autre véhicule dans le domaine *ad-hoc*.

- *La durée de connectivité d'une rencontre C2C* : il s'agit de la durée durant laquelle un véhicule reste à portée de transmission d'un autre véhicule.

- *Le temps entre deux rencontres C2I* : il s'agit de la durée pendant laquelle un véhicule est hors de portée de transmission d'un point d'accès à l'infrastructure fixe.

2.3 Caractérisation des paramètres de connectivité

Dans ce mémoire de thèse, nous avons considéré trois méthodes complémentaires pour déterminer les caractéristiques des distributions associées aux variables aléatoires définies au paragraphe 2.2 : la modélisation analytique, la simulation, et le traitement de traces de mobilité disponibles sur Internet. L'objectif est aussi est d'étudier ces paramètres dans différents environnements et scénarios de mobilité (autoroute, zone urbaine) et d'analyser l'impact de différents facteurs tels que la densité de trafic, la portée de transmission, la vitesse des véhicules, etc.)

2.3.1 Modélisation analytique

Le développement de modèles théoriques permettant de déterminer les distributions des paramètres de connectivité que nous avons considérés dans des environnements de mobilité réalistes n'est pas toujours possible. Il est nécessaire dans ce contexte de considérer des hypothèses simplificatrices et d'étudier ensuite dans quelle mesures les conclusions obtenues restent valides si on assouplit certaines de ces hypothèses. Cette dernière analyse peut être effectuée en s'appuyant sur des techniques de simulation ou le traitement de traces réelles.

Dans les modèles que nous avons développés, nous avons focalisé sur la distribution des intervalles de temps entre les rencontres de véhicule-à-véhicule en considérant comme environnement une route rectiligne. Les hypothèses sont les suivantes :

- La route a une longueur infinie.
- Les véhicules se déplacent dans une seule direction
- Chaque véhicule est supposé avoir une portée de transmission de forme circulaire sur un rayon de longueur constante R_v.
- Chaque véhicule se déplace avec une vitesse constante échantillonnée à partir d'une variable aléatoire « V » avec une espérance E|V|. Les mouvements des véhicules sont stochastiquement indépendants.

Avec ces hypothèses, nous avons démontré que les intervalles de temps entre les rencontres véhicule-à-véhicule suivent une loi exponentielle. Le taux de rencontre α associé à cette distribution est égal à ρ.E|V| où ρ est la densité moyenne des véhicules correspondant au nombre de véhicules par mètre.

2.3.2 La simulation

La simulation offre la possibilité de considérer des scénarios plus complexes. Notre objectif est d'étudier si le résultat obtenu par modélisation analytique concernant la distribution des temps de rencontre véhicule-à-véhicule reste valide en considérant des hypothèses plus générales. En particulier, nous avons considéré les hypothèses suivantes :

- La route a une longueur finie L et une largeur 2W.
- La route est constituée de deux voies avec des véhicules évoluant dans deux directions opposées, une voie étant destinée à chaque direction.
- Chaque véhicule a une portée de transmission de forme circulaire sur un rayon de longueur constante R_v.
- Chaque véhicule se déplace avec une vitesse constante, échantillonnée à partir d'une variable aléatoire « V » de distribution *Uniforme*. Les mouvements des véhicules sont stochastiquement indépendants..
- Les simulations sont effectuées en considérant une route de largeur W=15m, et de longueur L=5000m et différentes valeurs pour la densité des véhicules (ρ).

Les résultats de la simulation et les études de sensibilité que nous avons effectuées ont confirmé que la distribution des intervalles de temps entre des rencontres véhicules-à-véhicules peut être représentée dans ce cas aussi par une loi exponentielle comme dans le cas de l'étude analytique avec des hypothèses plus simples. La loi exponentielle offre également une bonne approximation : i) de la distribution des intervalles de temps entre les rencontres d'un véhicule avec un point d'accès de l'infrastructure fixe (rencontre C2I) et ii) de la distribution de la durée pendant laquelle un véhicule reste à portée de transmission d'un autre véhicule (durée de connectivité d'une rencontre C2C).

Nous avons réalisé plusieurs études de sensibilité afin d'analyser l'impact sur les résultats par exemple, de la densité des véhicules ρ et de la portée de transmission R_v. Pour le taux de rencontre de véhicule-à-véhicule, nous avons observé que pour une valeur donnée de la densité ρ, le taux de rencontre de véhicule-à-véhicule augmente progressivement avec R_v et atteint une asymptote quand la portée de transmission des véhicules couvre les deux voies de la route ($R_v \geq 2W$). Le taux de rencontre correspondant à l'asymptote est égal à $\rho.E|V|$, confirmant ainsi la même relation obtenue par modélisation analytique (voir §3.2.1). Ces conclusions ont été confirmées pour différentes valeurs de densité de véhicules ρ. En effet, les courbes donnant l'évolution du taux de rencontre de véhicule-à-véhicule en fonction de R_v pour différentes valeurs de ρ présentent la même allure.

2.3.3 Le traitement de traces de mobilité

L'une des hypothèses principales de la simulation de scénarios de mobilité considérés dans les paragraphes précédents est que les véhicules se déplacent indépendamment les uns des autres à une vitesse fixe (bien que différente entre les véhicules). Dans cette section, nous considérons des traces de mobilité issues de mesures en opération ou de simulations plus détaillées qui prennent en compte les dépendances entre les véhicules. Ces dépendances résultent par exemple des conditions du trafic (ralentir en cas de congestion, etc.), du comportement d'utilisateurs (changer de voies, sortir de l'autoroute, etc.).

Plusieurs traces de mobilité de véhicules issues de différents environnements (autoroute, trafic urbain, etc.) sont disponibles sur Internet. Nous avons considéré les deux traces suivantes qui sont pertinentes dans notre contexte et contiennent les informations qui sont nécessaires pour estimer les paramètres de connectivité étudiés dans nos travaux : i) les traces MMTS décrivant des scénarios de mobilité sur une autoroute pour différentes densités de trafic qui sont issues du simulateur *Multi-agent Microscopic Traffic Simulator* développé chez ETH Zurich [Traces_MMTS] ; ii) les traces CRAWDAD [Kotz & Henderson 2005] issues de mesures réelles dans un environnement urbain correspondant à des enregistrements de positions GPS de taxis dans la zone de San Francisco.

Les traces MMTS correspondent à des enregistrements de positions de véhicules évoluant sur sur une autoroute avec trois densités (ρ) différentes (faible, moyenne et forte). En faisant l'hypothèse que chaque véhicule a une portée de transmission entre 100 à 300m, l'analyse statistique de la distribution des intervalles de temps entre des rencontres véhicules-à-véhicules a montré que même si la loi exponentielle reste acceptable statistiquement pour représenter les données recueillies dans les traces, la loi de Pareto offre une meilleure adéquation. Pour les autres paramètres de connectivité (durée pendant laquelle un véhicule reste à portée de transmission d'un autre véhicule, intervalle de temps entre rencontres véhicule-à-infrastructure fixe), la loi exponentielle est représentative des données dans les traces.

Pour caractériser les paramètres de connectivité dans des environnements urbains, nous avons utilisé les traces de CRAWDAD en considérant des hypothèses équivalentes concernant la portée de transmission des véhicules et des point d'accès à l'infrastructure fixe. Nous avons observé que dans ce cas aussi la loi de *Pareto* offre une meilleure adéquation pour caractériser les intervalles de temps entre rencontres de véhicules-à-véhicules, et que la loi exponentielle reste acceptable

pour représenter la distribution des deux autres paramètres de connectivité. Ces analyses statistiques sont confirmées par les tests de Kolmogorov-Smirnof et le test du Khi-deux.

2.4 Conclusion

Dans ce chapitre, nous avons analysé certaines caractéristiques de connectivité entre deux véhicules et entre un véhicule et un point d'accès de l'infrastructure fixe qui sont importantes pour la conception et l'évaluation des performances et de la sûreté de fonctionnement d'applications véhiculaires utilisant des communications sans fil et des réseaux ad-hoc telles que celles présentées dans les chapitres suivants. Nous avons utilisé trois méthodes pour caractériser les paramètres de connectivité considérées dans le cadre de nos travaux : des preuves mathématiques, des simulations et des analyses statistiques de traces de mobilité.

Nous avons considéré principalement deux types de scénarios de mobilité : sur autoroute et dans une zone urbaine. Le tableau 2.1 résume les principales conclusions concernant la distribution des paramètres de connectivité considérés dans le cadre de notre étude. En particulier, concernant la distribution des intervalles de temps entre des rencontres de véhicule-à-véhicule, nous avons observé que cette distribution peut être décrite par une loi exponentielle dans le cas de scénarios de mobilité sur autoroute alors que la loi de Pareto est plus appropriée dans le cas d'un trafic urbain. Pour les deux autres paramètres de connectivité, la loi exponentielle offre une bonne adéquation. Ces résultats seront utilisés pour la modélisation de la sûreté de fonctionnement des trois cas d'études présentés dans les chapitres suivants.

Tableau 2.1: Caractérisation des paramètres de connectivité : Synthèse et conclusions

Scénario de mobilité hypothèses, technique d'analyse	Distribution temps entre rencontres véhicule-à-véhicule (C2C)	Distribution temps entre rencontres véhicule-à-infrastructure fixe (C2I)	Distribution des durées de connectivité C2C
Autoroute infinie avec trafic sur une voie et mouvements independants des véhicules – Preuve mathématique	• *exponentielle*	—	—
Autoroute de longueur finie, avec trafic sur deux voies dans deux directions opposées avec mouvements independants des véhicules - Simulation	• *exponentielle*	• *exponentielle*	• *exponentielle*
Trafic sur autoroute avec dependances entre les movements des véhicules – Traces MMTS	• *exponentielle/Pareto*	• *exponentielle*	• *exponentielle*
Trafic urbain avec dépendances entre les movements des véhicules – traces CRAWDAD	• *Pareto*	• *exponentielle*	• *exponentielle*

Chapitre 3 : Cohérence et disponibilité de données répliquées dans le domaine ad-hoc

La réplication des données est une technique bien connue et couramment utilisée pour assurer la sûreté de fonctionnement des applications en prenant en compte leur état interne. Un des challenges consiste à préserver la cohérence des états des différentes répliques. Ce chapitre traite de la modélisation et l'évaluation de la cohérence des états des différentes répliques et la disponibilité de ces répliques dans le contexte d'un service de réplication de données pour les applications véhiculaires qui utilisent des communications C2C. La modélisation est basée sur des réseaux d'activités stochastiques (SAN) et prend en compte les résultats du chapitre 2 qui sont relatifs à la caractérisation des paramètres de connectivité. Les résultats présentés dans ce chapitre constituent une extension des travaux préliminaires présentés dans [Matthiesen Moller *et al.* 2008] qui ont été effectués en collaboration avec l'Université d'Aalborg.

3.1 Problématique et contributions

Le cas d'étude présenté dans ce chapitre concerne un service de réplication de données dans le domaine *ad-hoc* en considérant un groupe de véhicules créé dynamiquement, contenant chacun une réplique des données. Ce service est décrit dans [Matthiesen Moller *et al.* 2008]. Un groupe est constitué d'un ensemble de répliques qui doivent avoir des états cohérents. Chaque membre d'un groupe peut lire les données partagées et également les modifier. Par conséquent, un mécanisme de mise en cohérence des données des différentes répliques est nécessaire. Par ailleurs, la défaillance des membres du groupe peut affecter la disponibilité des données partagées.

L'objectif des travaux présentés dans ce chapitre est d'élaborer des modèles de sûreté de fonctionnement permettant d'évaluer la disponibilité et la cohérence des données répliquées. Les modèles tiennent compte d'une part de l'évolution dynamique de la constitution des groupes de répliques provoquée par la mobilité et d'autre part des défaillances potentielles des véhicules sur lesquels les répliques sont en cours d'exécution.

Pour analyser l'impact de la mobilité des véhicules sur les mesures évaluées, nous avons considéré, conformément aux analyses effectuées dans le Chapitre 2, deux types de distributions pour décrire les temps entre les rencontres C2C : la loi *exponentielle* et la loi de *Pareto*.

3.2 Modélisation

Les modèles de sûreté de fonctionnement présentés dans ce chapitre constituent une extension de nos travaux décrits dans [Matthiesen Moller *et al.* 2008]. Dans ces travaux, la modélisation était basée sur les chaînes de Markov et les réseaux de Petri stochastiques généralisés (RdPSG) et tous les processus décrits dans ces modèles sont représentés par des lois exponentielles. Dans ce chapitre, nous considérons des scénarios de mobilité plus généraux. Pour ce faire, nous utilisons des réseaux d'activité stochastiques (SAN).

Les modèle de sûreté de fonctionnement développé pour le service de réplication prend en compte les les paramètres et processus et suivants :

(i) le taux avec lequel un nouveau membre rejoint le groupe ; ce paramètre est issu du modèle de mobilité décrivant les rencontres C2C dans le domaine ad-hoc ; nous avons analysé deux types de distributions : exponentielle et Pareto ;

(ii) le taux avec lequel un membre quitte le groupe ; le processus correspondant est décrit par une loi exponentielle.

(iii) le taux de modification des données de l'application ; le processus correspondant est décrit par une loi exponentielle.

(iv) le taux de mise à jour de l'état d'un réplique suite à une modification des données ; le processus correspondant est décrit par une loi exponentielle.

Le modèle décrit l'évolution de l'état du système résultant de l'occurrence de ces différents processus. Trois mesures principales sont évaluées à partir du modèle :

1) la probabilité que toutes les répliques d'un groupe aient des données cohérentes ;

2) la probabilité qu'il existe au moins une réplique disponible aux utilisateurs, en prenant en compte la possibilité de perte de répliques dues à la défaillance des véhicules sur lesquels ils s'exécutent ;

3) la probabilité que les utilisateurs puissent avoir accès à au moins un réplique et que les données correspondantes soient cohérentes avec l'état le plus récent de l'application.

3.3 Résultats et Conclusion

Nous avons effectué plusieurs études de sensibilité pour analyser l'évolution de ces mesures en fonction des paramètres du modèle et des scénarios de mobilités considérés. Ces analyses sont effectuées en considérant différentes valeurs pour le nombre maximal de répliques au sein d'un groupe, noté *n*, qui indique le niveau de redondance maximal souhaité. Elles permettent d'aider les concepteurs à dimensionner le niveau de redondance souhaiter en fonction des exigences de disponibilité et de cohérence des données requises par l'application.

Les études de sensibilité nous ont permis d'analyser dans un premier temps l'évolution dynamique du nombre de répliques au sein d'un groupe qui est influencé par la mobilité des utilisateurs. Nous avons ensuite étudié l'évolution des trois mesures de sûreté de fonctionnement considérées permettant d'analyser la disponibilité et la cohérence des données des différents répliques.

Nous avons observé que la probabilité d'atteindre la taille maximale d'un groupe diminue de manière significative pour des valeurs de *n* supérieure à 4. Néanmoins, même si le groupe n'atteint pas sa taille maximale, la probabilité d'avoir des répliques tout à fait cohérentes peut être très élevée. Par exemple cette probabilité est supérieure à 80% dans le cas où le taux de mise à jour des états des répliques est au moins dix fois plus rapide que le taux de modification des données de l'application. Typiquement, le taux de mise à jour des états des répliques sera d'autant plus élevé que le réseau ad-hoc interconnectant les véhicules sur lesquels tournent les répliques est rapide et non congestionné.

L'analyse combinée de la disponibilité et de la cohérence des données permet d'identifier les contextes qui offrent le meilleur compromis afin de satisfaire les exigences de sûreté de fonctionnement. En particulier, on peut observer que l'augmentation du nombre maximal de répliques au sein d'un groupe conduit à l'amélioration de la disponibilité de service, au détriment de la cohérence des données quand le nombre le répliques est supérieur à 4. Avec les paramètres considérés dans l'étude, le nombre optimal de répliques assurant le meilleur compromis entre la disponibilité et la cohérence des données est de l'ordre de 3 ou 4 répliques. Les tendances observées sont généralement similaires pour les deux types de distributions caractérisant les rencontres C2C dans le domaine ad-hoc (exponentielle et Pareto). Néanmoins, la disponibilité de service observée dans le cas où la mobilité est décrite par une loi de Pareto est plus faible, en comparaison avec le cas de la distribution exponentielle.

Chapitre 4 : Modélisation de la disponibilité d'une Boîte Noire Virtuelle pour les Systèmes Automobile

Le cas d'étude présenté dans le chapitre précédent a pris en compte la mobilité des véhicules et les rencontres de véhicules-à-véhicules (C2C) dans le domaine ad-hoc seulement. Le deuxième cas d'étude étudié dans ce chapitre qui concerne la Boîte Noire Virtuelle (BNV) étudie l'impact du point de vue de la de la sûreté de fonctionnement des rencontres de véhicules-à-véhicules (C2C) et de véhicules-à-infrastructures fixes (C2I). Une Boîte Noire Virtuelle permet de sauvegarder, de façon semblable à celle dans l'avionique, des informations importantes relatives à un véhicule afin d'être facilement récupérées en cas de problème. Cependant, au lieu d'utiliser un support matériel robuste sur le véhicule qui pourrait induire un coût élevé, une solution moins coûteuse serait de créer un mécanisme analogue en logiciel qui stockerait les données sur un serveur situé sur l'infrastructure fixe. Afin de protéger ces données contre d'éventuelles pertes avant qu'un accès à l'infrastructure fixe ne soit disponible, celles-ci peuvent être dupliquées et temporairement stockées sur les véhicules rencontrés qui utilisent les technologies de communication sans fil, avant d'être ensuite archivées sur le serveur. Le but de cette étude est d'analyser et d'évaluer la disponibilité des données sauvegardées dans la Boîte Noire Virtuelle (BNV), en comparant différentes stratégies de réplication des données possibles, et divers scénarios de mobilité. Les modèles sont basés sur les réseaux d'activités stochastiques (SAN).

4.1 Problématique et contributions

Les données importantes relatives à l'état d'un véhicule (telles que la vitesse et l'accélération du véhicule, la mise en action des freins, les indicateurs de direction, feux, et positions d'accélérateurs) peuvent être enregistrées périodiquement dans une boîte noire virtuelle afin d'être analysées en cas d'accident. D'un point de vue pratique, les données sont collectées en permanence sous forme d'enregistrements. La taille d'un enregistrement est d'environ 5 kilos octets. En cas d'accident, les derniers enregistrements de données successifs générés pendant les 15 à 30 dernières secondes avant l'accident sont généralement suffisants pour mieux restituer les conditions pendant lesquelles l'accident s'est produit.

En raison d'une capacité de stockage limitée sur le véhicule, et du fait qu'en cas d'accident les données stockées sur celui-ci pourraient être perdues, le contenu de la boîte noire virtuelle est sauvegardé sur un serveur sur l'infrastructure fixe.

Dans un premier temps, les enregistrements sont stockés dans le Véhicule lui-même, avant d'être livrés à l'infrastructure fixe dès qu'un point d'accès sera rencontré, pour compléter et mettre à jour la BNV. Entre deux mises à jour, l'information la plus récente, qui est cruciale en cas d'accident, ne se trouve que dans le Véhicule lui-même, et il y a un risque qu'elle soit perdue à cause de l'accident lui-même. L'idée est de tirer profit des véhicules rencontrés afin de répliquer les données en toute sécurité, avant qu'elles ne soient transmises à l'infrastructure fixe. Les données peuvent ainsi être transmises à l'infrastructure fixe par le biais soit du Véhicule lui-même soit par le biais des véhicules participants.

La réplication des enregistrements de données peut être réalisée soit en créant des copies complètes des enregistrements (réplication par duplication) soit par des mécanismes plus sophistiqués utilisant la stratégie de la réplication par fragmentation basée sur des codes d'effacement d'erreurs qui sont bien adaptés pour protéger les enregistrements de données contre des menaces aussi bien accidentelles que malveillantes. Considérons un enregistrement de données collecté par le Véhicule à un moment donné, qui doit être archivé sur l'infrastructure fixe, par l'intermédiaire de véhicules participants. Un code d'effacement d'erreurs, défini par les paramètres n et k ($n \geq k$), produit n fragments de l'enregistrement original qui sont ensuite dispersés parmi les véhicules environnants. Un code optimal permet de tolérer la perte de (n-k) fragments, c'est-à-dire, k fragments sont nécessaires et suffisants pour récupérer l'enregistrement de données original. La réplication par duplication des données correspond au cas n=k=1.

Quand un véhicule participant accède à l'infrastructure fixe, il transfère tous les fragments de données qu'il possède. En particulier, en cas d'accident, on peut s'attendre à ce que tous les fragments sur les véhicules participants soient délivrés à l'infrastructure fixe afin d'être utilisés pour l'analyse. Les derniers enregistrements successifs sont analysés en cas d'accident. Cependant, dans notre analyse, nous prenons en compte également le cas où un sous-ensemble parmi ces enregistrements est suffisant pour analyser l'accident.

Ce chapitre présente un modèle de sûreté de fonctionnement basé sur les réseaux d'activités stochastiques permettant de comparer différentes stratégies de réplication et de mieux comprendre l'impact de la mobilité sur la disponibilité des données. L'impact de la mobilité est étudié en considérant différentes distributions pour décrire le processus de rencontre des véhicules, en s'appuyant sur les résultats du chapitre 2.

4.2 Approche de modélisation

Le modèle SAN que nous avons développé permet de quantifier l'indisponibilité de la boîte noire virtuelle via l'évaluation de la probabilité de perte d'un ou plusieurs enregistrements de données collectés juste avant l'accident.

Le modèle est obtenu par composition de sous modèles représentant : i) la génération des enregistrements de données, ii) le comportement d'un enregistrement de donnée entre sa création et son transfert sur l'infrastructure fixe ou éventuellement sa perte, et iii) les conséquences de la perte de plusieurs enregistrements. Ce modèle tient compte du processus de rencontre entre véhicules et du processus de rencontre entre véhicules et infrastructure fixe, ainsi que de la stratégie de réplication des données (par duplication simple des enregistrements ou par fragmentation de l'enregistrement basée sur un code d'effacement d'erreurs). Le modèle ainsi construit permet de comparer différentes stratégies de réplication des données et d'effectuer des

analyses de sensibilité pour étudier l'impact des scénarios de mobilité et de facteurs environnementaux divers.

Un résultat important concerne le gain de disponibilité amené par la réplication de données comparé au cas où une telle réplication n'est pas employée. L'étude montre également que la distribution du processus de rencontres entre véhicules (selon une loi exponentielle ou de Pareto) peut avoir une influence significative sur la disponibilité.

4.3 Résultats et Conclusion

Cette étude s'inscrit dans le prolongement de nos travaux précédents liés à l'évaluation de la sûreté de fonctionnement d'un service de sauvegarde coopérative présentés dans [Courtès 2007], avec des extensions significatives. Les modèles présentés dans [Courtès 2007] sont basés sur des réseaux de Pétri stochastiques généralisés, et considèrent le cas d'un seul enregistrement de données, en supposant que le processus de rencontre des véhicules dans le domaine ad-hoc suivent une distribution exponentielle. Nous avons étendu le modèle initial afin d'analyser la disponibilité de l'application de la boîte noire virtuelle, en prenant en compte plusieurs enregistrements de données et des scénarios de mobilité pour lesquels il est plus représentatif de décrire le processus de rencontre de véhicule-à-véhicule par une distribution Pareto comme observé dans le chapitre 2. Ces extensions, nous ont conduit à utiliser comme formalisme de modélisation les réseaux d'activités stochastiques.

Les résultats obtenus dans ce chapitre montrent que la distribution des rencontres véhicule-à-véhicule (*exponentielle* ou *Pareto*) peut avoir une influence significative sur la disponibilité de l'application BNV, selon la valeur du rapport de connectivité c, qui représente le rapport entre la fréquence des rencontres dans le domaine *ad-hoc* et la fréquence des accès à l'infrastructure fixe. Un autre résultat important concerne le gain de disponibilité potentiel obtenu par la réplication de données dans le domaine *ad-hoc* comparé au cas où une telle réplication n'est pas employée. Pour un seul enregistrement de données, ce gain peut atteindre un niveau équivalent à la valeur du rapport de connectivité *c*, qui correspond au rapport entre le taux de rencontres entre véhicules et taux de connexion à l'infrastructure fixe (c = 100 dans notre étude). Ce résultat, valable pour les deux types de distributions considérées (*exponentielle* et *Pareto*), généralise la conclusion obtenue dans [Courtès 2007] où on a considéré uniquement le cas de la distribution *exponentielle*. De plus, notre étude a montré que le gain en disponibilité apporté par la réplication peut être nettement plus élevé quand on considère plusieurs enregistrements de données (de l'ordre de 104 quand c=100 et la distribution des rencontres véhicule-à-véhicule est *exponentielle*).

Enfin, les analyses présentées dans ce chapitre peuvent être affinées pour évaluer des mesures caractérisant la sécurité-innocuité du trafic, prenant en compte la distribution et le taux d'occurrence des accidents. En outre, le modèle peut être étendu afin d'évaluer la sûreté de fonctionnement de l'application BNV, en considérant les enregistrements recueillis sur une longue période et non pas seulement les données enregistrées juste avant un accident.

Chapitre 5 : Évaluation de la sécurité innocuité d'un Système Autoroute Automatisée

Le trafic automobile est de plus en plus congestionné, surtout dans les zones urbaines. Une des solutions étudiées est l'automatisation du trafic. De nombreux programmes de recherche ont été conduits, ou sont en cours, relatifs à l'assistance à la conduite automobile qui constitue, à long terme, une voie vers la route ou l'autoroute automatisée. Un exemple consiste à former des convois, ou des pelotons de véhicules (« *Platoon* » en anglais) pouvant évoluer de façon autonome. La mise en œuvre de routes automatisées vise à améliorer la fluidité du trafic et la sécurité routière (grâce à la diminution des accidents), tout en réduisant la consommation de carburant et les nuisances (en particulier la pollution). Dans ce chapitre, nous abordons le problème de l'évaluation de mesures quantitatives caractérisant la sûreté de fonctionnement dans le contexte d'un système d'autoroute automatisée, basé sur l'utilisation de pelotons de véhicules conduits par des agents plus ou moins autonomes, interagissant dans un même environnement *multiagent.*

5.1 Problématique et contributions

Dans ce chapitre, nous abordons le problème de l'évaluation de mesures quantitatives caractérisant la sûreté de fonctionnement dans le contexte d'un système d'autoroute automatisée, basé sur l'utilisation de pelotons de véhicules conduits par des agents plus ou moins autonomes, interagissant dans un même environnement multiagent. Notre travail porte sur le développement de méthodes et de modèles permettant d'évaluer la sécurité-innocuité de pelotons mis en œuvre dans un contexte mobile avec des réseaux *ad-hoc*. Plusieurs phénomènes, tels que l'occurrence de fautes accidentelles, la mobilité des véhicules, les déconnexions fréquentes des communications, sont à prendre en compte. Cette problématique est nouvelle dans le contexte d'applications et systèmes d'autoroutes automatisées mis en œuvre sur des réseaux *ad-hoc*, et il n'existe pas encore à notre connaissance de méthodologies et de résultats permettant d'évaluer la sûreté de fonctionnement dans ce domaine.

Les modèles développés et les mesures de sûreté de fonctionnement évaluées permettront de fournir aux concepteurs des éléments de décision objectifs pour l'analyse et la sélection des architectures les plus adaptées pour satisfaire les exigences de sûreté de fonctionnement. Dans nos travaux, nous considérons comme cas d'étude les architectures développées aux États-Unis dans le cadre du projet PATH. Ces architectures s'appuient sur la mise en œuvre de manœuvres automatiques permettant d'assurer la sécurité-innocuité (*safety*) des pelotons en présence de différents types de modes défaillance susceptibles d'affecter les véhicules, leur environnement ou la communication inter-véhicules. Les travaux présentés dans ce chapitre ont pour objectif de proposer une méthodologie et des modèles analytiques basés sur les réseaux d'activités stochastiques (SAN) permettant de modéliser ces manœuvres et d'évaluer leur impact sur la sécurité innocuité des pelotons. Plusieurs phénomènes, tels que l'occurrence de fautes accidentelles, la mobilité des véhicules, les moyens de communication entre les véhicules et les pelotons, et les déconnexions fréquentes des communications, sont à prendre en compte.

5.2 Approche de modélisation et résultats

Nous avons considéré deux voies d'autoroute automatisées avec un peloton dans chaque voie. Les véhicules dans chaque peloton peuvent changer d'une voie à une autre. Chaque peloton peut avoir jusqu'à *N* véhicules. Nous avons modélisé ce système en prenant en compte six modes de défaillances et leurs manœuvres de recouvrement associées. Selon le mode de défaillance considéré, certaines manœuvres nécessitent la communication entre les leaders (le premier véhicule dans le peloton) de plusieurs pelotons.

Nous analysons la sécurité d'un Système Autoroute Automatisée via l'évaluation de la probabilité d'avoir une situation catastrophique pour le système modélisée en fonction de temps (*t*). Cette mesure est appelée *insécurité* et est désignée par $\overline{S}(t)$.

Plusieurs paramètres et facteurs environnementaux sont pris en compte dans la modélisation : i) la mobilité des véhicules, ii) deux stratégies de coordination possibles (centralisée ou décentralisée) entre les véhicules appartenant au même peloton (intra-peloton) ou à des pelotons voisins (inter-peloton), et iii) différents modes de défaillances pouvant affecter les véhicules et leur communications, ainsi que les manœuvres de recouvrements associées.

Le modèle SAN est obtenu par composition de sous modèles représentant : i) la configuration des pelotons, ii) le comportement des véhicules prenant en compte les défaillances et les manœuvres de recouvrement, iii) l'évolution dynamique des pelotons en absence de défaillance, résultant des arrivées et des départs de véhicules, et iv) les conséquences de défaillances multiples affectant plusieurs véhicules. Le modèle ainsi construit permet de comparer différentes stratégies de communication inter et intra peloton et d'effectuer des analyses de sensibilité pour étudier l'impact des scénarios de mobilité et de modes de défaillances sur la sécurité du trafic. En particulier nous avons analysé l'évolution de la mesure de sécurité considérée pour différents durées de trajet (variant de 2 à 10 heures) et en étudiant également l'impact du nombre de véhicules par peloton et du taux de défaillance des véhicules. Par exemple, nous avons observé quand le taux de défaillance des véhicules est de l'ordre de 10^{-5}/heure, le système d'autoroute automatisée offre un niveau de sécurité élevé quand le nombre de véhicules par peloton est inférieur à 10. La comparaison des différentes stratégies de coordination a montré que la configuration optimale correspond au cas où les coordinations inter-peloton et intra-peloton sont décentralisées.

5.3 Conclusion

Certes, l'évolution vers une conduite automobile automatisée sur autoroute améliore la sécurité globale puisqu'elle fait intervenir les véhicules voisins d'un véhicule défectueux pour ramener l'ensemble vers un état sûr. Cependant, le risque d'accident n'est pas nul pour autant. Ce chapitre propose une approche d'évaluation du risque d'accident dans un environnement de conduite coopérative sur autoroute, faisant intervenir des communications entre véhicules/pelotons voisins de façon centralisée/décentralisée. Nos modèles prennent en compte les modes de défaillance affectant les véhicules, leur niveau de gravité et les manœuvres permettant de ramener le système dans un état sûr. L'approche de modélisation basée sur des SANS est conçue pour prendre en compte l'évolution des configurations des pelotons durant le trajet. Elle est modulaire et compositionnelle, en développant des modèles génériques caractérisant le comportement des

véhicules qu'on peut interconnecter facilement quand une nouvelle configuration ou un nombre de véhicules plus important doit être pris en compte.

Pour illustrer l'approche et le type de résultats que l'on peut obtenir, nous avons considéré des exemples simples. Nous avons effectué des études de sensibilité permettant d'analyser l'impact de quelques paramètres sur la sécurité d'un système d'autoroute automatisée et d'aider les concepteurs à trouver les configurations optimales offrant le meilleur niveau de sécurité espéré, en déterminant en particulier, le nombre maximal de véhicules par peloton, la durée de trajet et la stratégie de coordination des véhicules inter et intra pelotons.

Nous avons considéré dans la modélisation que les processus intervenant dans le modèle suivent des distributions exponentielles. Ces hypothèses peuvent s'avérer discutables dans certains contextes, en particulier pour ce qui concerne le processus de sortie ou d'entrée dans un peloton qui dépend fortement des caractéristiques de mobilité des utilisateurs. Les extensions envisagées pour nos travaux ont pour objectif d'étendre les modèles actuels pour prendre en compte d'autres types de coordination caractérisant les méthodes d'interaction entre les véhicules. On peut envisager également la prise en compte de scénarios impliquant un nombre significatif de véhicules et de pelotons, et ensuite la généralisation de l'approche à des scénarios plus complexes.

Conclusion générale et perspectives

La sûreté de fonctionnement de services et applications utilisant des communications mobiles suscite un intérêt croissant dans le domaine de l'automobile. Les réseaux sans fil *ad-hoc*, et les technologies de communication sans fil sont de plus en plus intégrées dans nos véhicules pour améliorer la sécurité et également pour offrir de nouveaux services de loisirs etd e confort aux passagers. L'évaluation de la qualité de service et du niveau de sûreté de fonctionnement fournie par ces applications est important dans ce contexte.

Les travaux présentés dans ce mémoire de thèse portent sur l'élaboration d'approches de modélisation visant à l'évaluation de la sûreté de fonctionnement d'applications mises en œuvre sur des véhicules utilisant des communications de véhicule-à-véhicule (C2C) via des réseaux ad-hoc, et des communications de véhicules-à-infrastructure fixe (C2I) pour accéder à des serveurs distants. A notre connaissance, il existe peu de travaux dans ce domaine et il y a un manque d'exemples concrets permettant d'illustrer comment l'évaluation de sûreté de fonctionnement peut être effectuée dans ce contexte. Trois cas d'études sont présentés à savoir un service de réplication de données dans le domaine *ad-hoc*, une application boîte noire virtuelle, et un système d'autoroute automatisée. Ils nous ont permis d'illustrer différentes caractéristiques de sûreté de fonctionnement des applications véhiculaires utilisant des communications C2C et C2I et de développer des approches de modélisation adéquates permettant de prendre en compte différents scénarios de mobilité et des paramètres de la conception. Nous avons principalement considéré l'impact des menaces accidentelles sur les services fournis. Plusieurs propriétés de sûreté de fonctionnement et de mesures quantitatives ont été étudiées en fonction de l'étude de cas, avec un accent particulier sur la *disponibilité* et la *sécurité*.

Un aspect important de notre travail concerne la caractérisation de scénarios de mobilité afin d'estimer les distributions des processus de rencontres C2C et C2I en utilisant des techniques

analytiques, de la simulation et des traitement de traces issues d'observations réelles. Ces analyses ont montrée en particulier que les rencontres C2C peuvent être décrites par une loi exponentielle ou une loi de Pareto en fonction des environnements et des caractéristiques de trafic considérés (autoroute ou trafic urbain). Les résultats issus de ces analyses sont ensuite intégrés dans des modèles de sûreté de fonctionnement pour évaluer les mesures de disponibilité ou de sécurité recherchées.

Plusieurs orientations de recherche sont possibles pour compléter et étendre les contributions présentées dans cette thèse, relatives à la modélisation et l'évaluation de la sûreté de fonctionnement des applications mobiles.

Notons d'abord que les modèles présentés dans cette thèse peuvent être étendus à l'analyse de cas d'utilisation et de scénarios de mobilité plus complexes. Par exemple, les modèles peuvent être affinés pour tenir compte des défaillances survenant sur l'infrastructure fixe afin de fournir des mesures de sûreté de fonctionnement de bout-en-bout. Les modèles de sûreté de fonctionnement peuvent être également structurés de façon hiérarchique en distinguant différent niveaux d'abstraction à savoir : i) l'utilisateur, l'application incluant la description des fonctions et services fournis, iii) l'architecture matérielle et logicielle mettant en œuvre l'application et iv) le support de communication. Un cas d'étude plus détaillé est nécessaire pour illustrer ces différents niveaux. Dans ce contexte, la combinaison de différents types de formalismes et techniques d'évaluation, incluant des modèles analytiques, des simulations et des mesures expérimentales peut s'avérer nécessaire afin de fournir des évaluations de sûreté de fonctionnement détaillées de bout en bout, en tenant compte de la décomposition et des différents niveaux d'abstraction du système.

Dans cette thèse, nous avons considéré uniquement l'impact sur la sûreté de fonctionnement de défaillances d'origine accidentelle. L'impact des menaces malveillantes devrait également être étudié dans les travaux futurs afin de fournir une estimation plus réaliste et aider les concepteurs à faire des choix d'architecture en conciliant les exigences relatives à la disponibilité et la sécurité-innocuité des services fournis ainsi qu'à l'intégrité ou la confidentialité des données.

Enfin, dans nos travaux, nous avons considéré des cas d'étude issus du domaine automobile. Néanmoins nos résultats peuvent être étendus également à d'autres domaines d'application.

Abstract

This thesis focuses on developing methods and models making it possible to evaluate quantitative measures characterizing the dependability of mobile services as perceived by the users. These models and measures are aimed at providing support to the designers during the selection and analysis of candidate architectures that are well suited to fulfill the dependability requirements. We consider the case of vehicular applications using inter-vehicle communications based on *ad-hoc* networks and may have access to services located on fixed infrastructure.

We propose an approach combining: 1) dependability models based on stochastic activity networks, in order to describe the system failure modes and associated recovery mechanisms, and 2) simulation and analytical models allowing the estimation of connectivity characteristics, taking into account different mobility scenarios and environment. This approach is illustrated on three case studies including a virtual black box based on cooperative data replication and backup, and an automated highway system (Platooning application).

Table of Contents

List of Figures

List of Tables

Introduction

"*Success is ninety nine percent failure."*
Soichiro Honda, Japanese, Businessman Quotes

"*Fortunately science, like that nature to which it belongs, is neither limited by time nor by space. It belongs to the world, and is of no country and of no age. The more we know, the more we feel our ignorance; the more we feel how much remains unknown; and in philosophy, the sentiment of the Macedonian hero can never apply, — there are always new worlds to conquer."*
Sir Humphrey Davy, Cornish chemist, 1825

With the increasing deployment of wireless networking and mobile computing technologies at home/office/open public spaces/highways and the proliferation of wireless enabled portable devices, e.g., palm-size computers, cellular phones and pagers, a number of new services and applications taking advantage of these technologies have already been introduced. The implementation of such services is more and more based on mobile *ad-hoc* networks with the possibility of connection to remote servers in the infrastructure domain. The automotive sector is an example of application area where the use of such wireless technologies have opened new opportunities for the development of innovative services in vehicular applications. Such services cover a large variety of domains including information and entertainment (voice and video streaming, online gaming, contextual information delivery, etc.), as well as safety and dependability critical services (hazard warning, safety and traffic management, etc.). This fast growing area provides great opportunities but also poses significant challenges from the dependability point of view.

Clearly, the success of such mobile-based applications and services heavily relies on their perceived dependability and resilience. Several constraints and challenges have to be taken into account in the design and evaluation of these new applications, in particular with respect to the dependability and resilience to accidental to malicious faults: mobility of the users, frequent disconnections, low reliability of wireless links, etc. Indeed, a period of service unavailability of such applications may lead to a catastrophic situation depending on the service interruption duration and on the criticality of the service.

The main objective of this dissertation is to develop methods and models making it possible to evaluate quantitative measures characterizing the dependability of mobile services as perceived by the users, considering vehicular applications as an example. These models and measures are aimed at providing support to the designers during the selection and analysis of candidate architectures that are well suited to fulfill the dependability requirements.

The topic addressed in this thesis has been seldom addressed in the literature. Indeed, the current state-of-the art dealing with the evaluation of the quality of service in the context of mobile applications using *ad-hoc* networks is mainly focused on the evaluation of performance related metrics, considering in particular communication and routing protocols. Our primary objective is to focus on the application and middleware layers, taking into account the fault tolerance and recovery mechanisms implemented at these levels to provide the continuity and the availability of service.

Model-based dependability analysis and evaluation methods are useful to help the users and the designers of mobile-based applications to compare various design alternatives from the dependability point of view and to understand how some parameters related to fault occurrence and recovery, as well as to the environment (mobility scenario, car speed and density, transmission range, etc.) may affect relevant dependability properties at the application and user level (availability, safety, etc.)..

To fulfill these objectives, two critical aspects need to be addressed: 1) describe mobility related characteristics and analyze their impact on the dependability of the considered applications taking into account different mobility scenarios and environments, and 2) find efficient means to master the complexity of the models both for model construction and model processing.

System and application designers are generally interested in understanding and quantifying the impact of failures at different levels of decomposition of the systems considered. A single technique is generally not sufficient to perform such evaluations and it is generally necessary to combine different techniques and tools. This is due in particular to the complexity of the analysis and the need to take into account different characteristics of the target systems that cannot be captured efficiently by a single method (description of mobility patterns, modeling of components failures and recovery strategies, etc). Compositional modeling combining analytical and simulation techniques is an interesting approach for controlling the models complexity and capturing mobility related characteristics. This type of methods has been privileged within the modeling approach and the case studies investigated of this thesis. In particular, the evaluation of quantitative dependability measures is based on the combined use of: i) analytical models based on Stochastic Activity Networks, ii) simulation techniques and iii) the statistical processing of real traces. The two latter techniques are used to characterize connectivity characteristics under various mobility scenarios. The parameters derived from these analyzes and associated probability distributions are used as an input in stochastic activity networks based models allowing the evaluation of dependability related characteristics at a higher level of abstraction. Stochastic Activity Networks are well suited to master the complexity of the models through the composition and replication of sub-models.

We have considered three case studies in our research namely, a replication service in the ad-hoc domain, a virtual black-box application, and a platooning application. These case studies exhibit various dependability characteristics and challenges that are typical of vehicular applications using Car-to-Car and Car-to-Infrastructure communications.

This work has been partially carried out within the framework of the European project (HIDENETS: *HIghly Dependable ip-based NETworks and Services*) [HIDENETS 2006a]. This project addressed the development and the validation of architectures and methods making it possible to ensure the dependability of applications and services that are implemented in a mobile context with *ad-hoc* networks. Selected use-cases of *ad-hoc* Car-to-Car communication with infrastructure service support have been used as illustrative examples in the context of this project.

This dissertation is organized as follows. Chapter one presents the motivation and the context of the thesis work and provides background information. We provide a discussion on dependability modeling starting by the main dependability concepts. We present a brief overview of dependability evaluation and the methods that can be used to evaluate mobile-based applications.

A hierarchical modeling approach developed in the context of the HIDENETS project for end-to-end dependability evaluation of such applications is also presented. The last part of this chapter presents background information about vehicular applications using wireless technologies and discusses the corresponding challenges from the dependability evaluation perspective. A brief description of the three case studies presented in the following chapters of this dissertation is also included. We finalize chapter one by giving a brief overview of the topics covered in the dissertation. Related work is discussed in the context of each chapter.

Chapter two is devoted to the estimation of connectivity dynamics in vehicular *ad-hoc* wireless networks and analyzes the distribution of car-to-car and car-to-infrastructure encounter times using analytical proofs, simulation experiments and the statistical processing of real traces. These distributions are used in the dependability modeling of the case-studies which are presented in the dissertation. The following three chapters illustrate the dependability modeling and evaluation of our three case-studies, respectively. Finally, Chapter six concludes on the contributions of this dissertation and identifies future research tracks.

Chapter One: Context and Background

"*Its how you deal with failure that determines how you achieve success.*"
Charlotte Whitton, Canadian, Politician Quotes

This chapter presents background information related to the topics developed in the dissertation. We first provide some definitions related to dependability, focussing in particular on evaluation and assessment activities. Then, we outline the main characteristics of vehicular mobile applications using Car-to-Car and Car-to-Infrastructure communications, and the challenges raised by such applications from the dependability evaluation perspective. Finally, we outline the main topics and contributions developed in the dissertation. Specific related work is discussed in the context of each chapter.

1.1 Dependability Concepts

The work presented in this dissertation aims to assess the *dependability* of mobile-based computing systems, considering vehicular applications as a case study. Dependability concepts and terminology for computing systems have been an active research area since the eighties.

According to the terminology presented in [Avizienis *et al.* 2004, Laprie *et al*. 1995], *dependability* is the ability to deliver a service that can justifiably be trusted. The service delivered by a system is its behavior as perceived by its user(s); a user is interacting with the system as another system (human or physical).

Dependability is a global concept which includes various notions that can be grouped into three classes: the *threats*, the *attributes* and *the means* by which dependability is attained.

The *threats* to dependability are: *faults*, *errors*, and *failures* that might affect the service(s) delivered by the system. They are undesired, but not in principle unexpected circumstances causing or resulting from undependability.

- *A service failure*: is an event that occurs when the delivered service deviates from correct service, *i.e.*, it is a transition from correct service to incorrect service delivery.

- *An error*: is the part of the system state that may lead to its subsequent service failure

- *A fault*: is the adjudged or hypothesized cause of an error.

Depending on the applications considered, different facets of dependability may be important, *i.e.*, different emphasis may be put on different attributes of dependability. The dependability attributes allow to: *a)* express the properties expected from the system; and *b)* to assess the quality of the service delivered, as resulting from the threats and the means used to avoid them.

Basic dependability attributes are defined as follows:

- *Reliability*: continuity of correct service.

- *Availability*: readiness for correct service.
- *Safety*: absence of catastrophic consequences on the user(s) and the environment.
- *Confidentiality*: absence of unauthorized disclosure of information.
- *Integrity*: absence of improper system alterations.
- *Maintainability*: ability to undergo modifications and repairs.

Several other dependability attributes can be obtained as combinations or specialization of the primary attributes listed above. In particular, *security* is defined as the concurrent existence of *a*) availability for authorized users only, *b*) confidentiality and *c*) integrity, where 'improper' means 'unauthorized'.

Several complementary means are needed to specify, design, and analyse systems in which a fault is natural, predictable, and tolerable. These are the methods and techniques that enable one *a)* to provide the ability to deliver a service on which reliance can be placed, and *b)* to reach confidence in this ability. Four major categories of dependability means can be distinguished:

- *Fault prevention*: aims to prevent the occurrence or introduction of faults.
- *Fault tolerance*: aims to avoid service failures in the presence of faults.
- *Fault removal*: aims to reduce the number and severity of faults.
- *Fault forecasting*: aims to estimate the present number, the future incidence, and the likely consequences of faults.

Fault prevention and fault tolerance aim to provide the ability to deliver a service that can be trusted, while fault removal and fault forecasting aim to reach confidence in this ability by justifying that the functional and the dependability and security specifications are adequate and that the system is likely to meet them.

This dissertation mainly addresses *fault forecasting* in the context of mobile-based systems, considering vehicular applications as an example. We mainly take into account the impact of accidental threats on the delivered services. Several dependability properties and quantitative measures will be investigated depending on the case study under investigation, with a particular focus on *availability* and *safety*.

1.2 Dependability evaluation

Fault forecasting is performed using evaluation of the behavior of a system relative to the occurrence of faults and their activations. By adopting a structural view of a system, evaluation consists in reviewing and analyzing the failures of its components and their consequences on the system's dependability. The analysis can be performed iteratively considering different levels of

decomposition and abstraction of the system. Evaluation can be conducted in two ways, depending on the target of the evaluation [Laprie *et al.* 1995]:

- *Ordinal evaluation*: aims at identifying, listing and ranking failures, or the methods and techniques used for avoiding them.
- *Probabilistic evaluation*: aims at evaluating in terms of probabilities the degree of satisfaction of certain attributes of dependability.

The methods and techniques designed to carry out dependability evaluations are either specific to each type of analysis (ordinal and probabilistic evaluation) or can be used for both. For example, FMEA (Failure Modes and Effects Analysis) is specific for ordinal evaluation while Markov chains are generally used for probabilistic evaluation. On the other hand, fault trees and reliability block diagrams can be used in both forms of evaluation. It is noteworthy that the two types of evaluation techniques are complementary, since probabilistic evaluation requires, in a first step, the identification of the failures modes to be taken into account in the assessment of the quantitative dependability measures.

1.2.1 Quantitative measures

Quantitative probabilistic measures are defined to assess the behavior of a system at different abstraction and decomposition levels (from the perspective of the users, for a given function, a subsystem, a particular fault tolerance mechanism, a communication protocol, etc.). These measures can be expressed statistically by evaluating the mean, the variance and the different associated quantities. Parametric probability distributions can also be associated to these measures and can be evaluated based on data collected from experimental measurements, from simulations, or from the processing and calculation of analytical models.

A wide range of measures can be considered to assess the dependability attributes defined in Section 1.1: *reliability*, *availability*, *safety*, *maintainability*, etc. The selection of the relevant measures depends on the application considered and the impact of failures on the users and the environment.

To define the dependability measures, two classes of states of the service delivered should be distinguished: *correct* and *incorrect*. The *failure* processes (from a correct service to an incorrect one) and the *restoration* processes (from an incorrect service to a correct one) govern transitions between these classes of states.

The main measures aim at characterizing the time of correct service delivery. Two main categories of measures are generally distinguished:

- Measures that characterize the sojourn time in the states where the correct service is being delivered (before reaching the incorrect service states): these correspond for example to *reliability* and *MTTF* (*Mean Time To Failure*), which measure the time of correct service delivery prior to a failure.

- Measures that characterize the delivery of correct service with respect to the alternation of correct and incorrect services: these encompass the various forms used to measure *availability* (time instant, time interval, or asymptotic).

When systems feature several performance levels, several modes of services (correct and incorrect) can be distinguished [Laprie *et al.* 1995]. According to the viewpoints considered to evaluate dependability, there exist two main "*extreme*" cases:

- The system features several modes of correct service completion and one single mode of incorrect service.

- The system features one single mode of correct service delivery and several modes of incorrect service

A particularly interesting case corresponding to the second category is that of systems exhibiting two modes of incorrect service with different levels of failure severity (*benign* and *catastrophic*). This allows the measures linked to the safety evaluation of these systems to be obtained within the very same framework. Thus, *safety* represents the measurement of time in the *safe states* (correct service delivery and benign failure) prior to a catastrophic failure. *Safety* is defined by the absence of catastrophic consequences on the user(s) and the environment

In addition to the previously discussed measures, it is sometimes relevant to carry out combined performance and dependability analyses, i) to assess the impact of failures on the performance of the system, ii) to study how performance degradations may lead to failures, and iii) assess the combined impact failures and performance degradation on system behavior. The *performability* concept defined in [Meyer 1980] is well suited for such analyzes as illustrated in several work e.g., [Ciciani & Grassi 1987, Smith *et al.* 1988, Ciardo *et al.* 1990, Rabah & Kanoun 2003, Grassi *et al.* 2007, Wang *et al.* 2007]

1.2.2 Dependability evaluation techniques

Dependability evaluation encompasses both measurement-based and model-based techniques.

Measurement encompasses both the observation of a real-life system in operation (or field measurement) and controlled experimentation on the real system or a prototype to characterize its behaviour under specific forced conditions. Measurement-based evaluation provides helpful support for understanding real phenomena (information on actual error/failure behaviour and on possible system bottlenecks) and for quantifying dependability measures. Even if there is no better way to understand the system dependability than by field measurement, analysis of field data can be performed only when the system is already in operation. In which case, feedback from field data from previous systems is very helpful for providing realistic estimates of certain parameters characterizing the behavior of specific mechanisms and components. Both dependability and performance related characteristics can be measured. These parameters can be used as inputs in dependability models describing the system behavior at a higher level of abstraction to estimate system level dependability measures. In the case of a system that does not exist yet, the parameters could be estimated from measurements on similar systems.

Modeling is useful to guide the development of a system during its design phase by providing quantitative measures characterizing its dependability. Various design alternatives can be compared in order to choose the final solution that provides an acceptable dependability, taking into account various impacting factors. In particular, models can enable the designers *i*) to understand the impact of various design parameters from the dependability point of view, and *ii*) to identify potential bottlenecks. These reasons make the modeling attractive in the probabilistic evaluation.

We can distinguish essentially two main modeling techniques: *simulation* and *analytical* modeling. Discrete-event simulation can be applied to almost all problems of interest, but the main drawback is that it can take a long time to run, mainly when large and realistic systems are described at a detailed level and also in the presence of rare events, even if rare events can be speeded up in a controlled manner using for example importance sampling techniques [Nicola *et al.* 1990]. On the other hand, simulation is usually the only possible solution when one needs to analyze the behavior of some mechanisms under some specific conditions and complex behavior that cannot be easily modeled using analytical techniques.

Analytical modeling is a more cost effective solution in terms of model processing time. However, it often requires simplifying assumptions, generally considering only essential aspects of the system at a high level of abstraction, in order to make models tractable. Nevertheless, we should note that considerable advances in the last twenty years have significantly extended the capabilities of analytic models.

Several evaluation techniques are generally needed to analyze and assess the dependability of complex systems at different abstraction and decomposition levels. In our dissertation, the three evaluation techniques (measurement, simulation, analytical models) will be combined to analyze and assess the dependability of vehicular mobile applications.

The next section presents a brief summary of modeling formalisms for the probabilistic evaluation of dependability.

1.2.3 Modeling formalisms for dependability evaluation

Two main types of models are extensively used for dependability evaluation: *combinatorial models* and *state-space* models.

Combinatorial models (e.g., Fault trees, Reliability Block Diagrams, Reliability Graphs) are concise, rather easy to build and have efficient processing methods. They capture static combinations that result in system failure in terms of the structural relationships between the system components and their states. These models are limited in expressiveness as strong dependencies between components and between modeled events typically cannot be taken into account. However, they can be used as parts (sub-models) in hierarchical and composite system-level models to describe specific processes or subsystems (e.g., a fault tree can describe how component failures may result in subsystem level failures and evaluate the corresponding probability that can then be used as an input in the system level model).

State-space models constitute the prevailing type of model for evaluating dependability measures, particularly with respect to stochastic dependencies between system components [Fota *et al*. 1999, Kanoun & Borrel 2000, Rabah & Kanoun 2003].

Markov chains are the most commonly used state-space models for system dependability and performance evaluation. They are based on the assumptions that all events described in the model are exponentially distributed. When other probability distributions need to be considered more complex processes (e.g., semi-Markov, Markov Regenerative or even more general processes) must be used.

To facilitate the generation of state-based dependability models (in particular Markov Chains), higher-level formalisms are generally used, mainly based on Stochastic Petri Nets and their extensions (Generalized Stochastic Petri Nets, Stochastic Activity Networks, Stochastic Well-formed Nets, etc.). The state space (called reachability graph) generated from these models can be mapped to a continuous time stochastic process (e.g., a Markov chain) that is suitable for the evaluation of the dependability measures. These models offer several interesting features such as the ability to represent graphically the concurrency and synchronization of events and activities, as well as well as resource contention, priorities, etc. Also some of these formalisms offer the possibility to carry out structural verifications on the model through the identification of transition and place invariants.

With Stochastic Petri Nets (SPN) only exponentially timed transitions are allowed [Florin & Natkin 1985, Juanole 2003]. Generalized Stochastic Petri Nets (GSPN) [Marsan *et al.* 1995] offer more flexibility by allowing both exponentially timed transitions and transitions with zero time delay (i.e., immediate transitions). GSPNs are recognized as a powerful modeling formalism for performance and dependability evaluation of concurrent and distributed systems. Another step forward in alleviating the drawback of having only exponentially timed transitions was the introduction of the Deterministic and Stochastic Petri Net (DSPN) [Marsan & Chiola 1987], Markov Regenerative Stochastic Petri Nets [Choi *et al.* 1994, Marie *et al.* 1997], or Stochastic Activity Networks (SANs) [Meyer *et al*. 1985, Sanders & Meyer 2001] which allow deterministic or more general distributions to be considered in the models. As regards SANs, they also provide powerful composition mechanisms that are very useful to master model complexity when dealing with large state space models including symmetries. It is noteworthy that several solutions have been proposed to cope with the largeness of the state space at the model construction and model processing level. A recent discussion of the state of the art can be found in [Nicol *et al*. 2004, Kaâniche *et al*. 2008].

In the case of Petri nets and their extensions, there are a number of tools suitable for modeling and evaluation. Among these tools are the following: SURF-2 [Béounes *et al.* 1993], SPNP [Ciardo *et al*. 1989], Möbius [Sanders & Meyer 2001, Dali *et al.* 2000], and DEEM [Bondavalli *et al.* 2000].

In this dissertation, we rely on Markov chains and Stochastic Activity Networks models as an abstraction model for evaluating the dependability of vehicular applications using Car-to-Car and Car-to-Infrastructure communication. The use of SANs in our context, instead of GSPNs for example, is mainly motivated by the need to take into account different types of distributions, not only exponential distributions. Also, as will be discussed in the following sections, combined

approaches using analytical, simulation and measurement techniques are needed to address the challenges raised by these applications from the dependability assessment point of view.

1.3 Context of the study: vehicular *ad-hoc* based applications

In this section, we present an overview of the main characteristics of vehicular *ad-hoc* based applications used as a main case study in our dissertation, focussing on background information needed for dependability modeling and evaluation presented in the next chapters.

1.3.1 Operational context and networking scenarios

Figure 1.1 presents a high-level view of the dissertation study context. It introduces the relevant components and network domains considered in our work. The main system components are the vehicles equipped with wireless network interfaces that can communicate with other vehicles or with a set of servers on the fixed infrastructure. The latter can be accessed through various road-side infrastructure units, and possibly through the Internet protocol core backbone using different service providers. The vehicular applications that are addressed here can be based on Car-to-Car (C2C) communications, Car-to-Infrastructure (C2I) communications, or on a hybrid scenario using a combination of both types of communications.

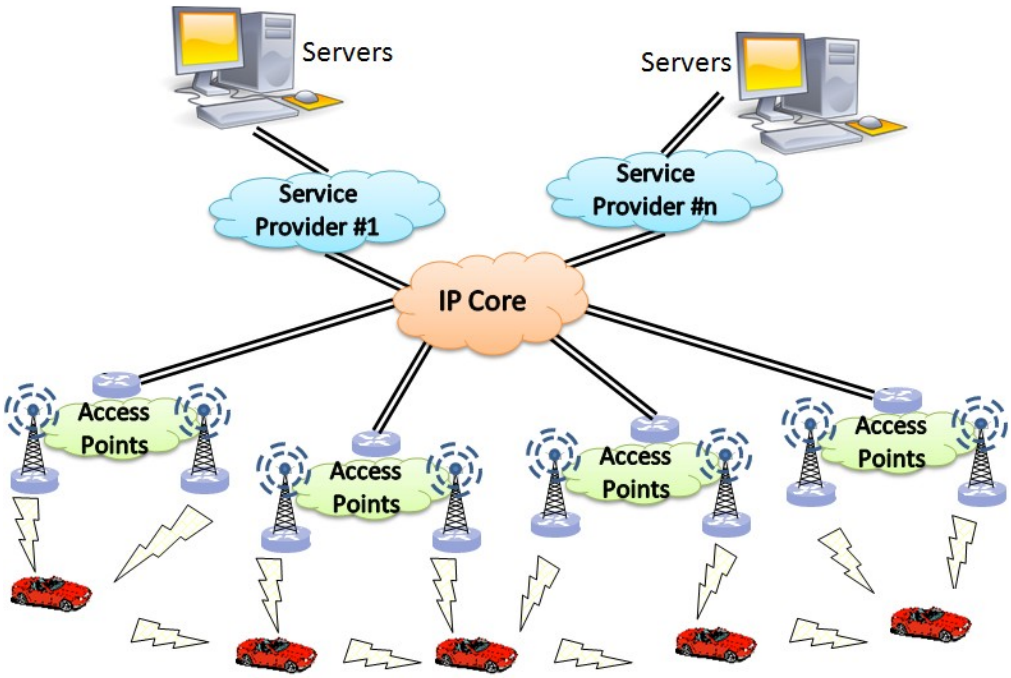

Figure 1.1: Operational Context of the Study

As abstracted by Figure 1.2, we can distinguish two fundamentally different domains:

1. the *ad-hoc domain* in which service access and service deployment are performed in a wireless setting,

2. the *infrastructure domain* that consists of road-side infrastructure units and, depending on the applications considered, the back-bone IP network connecting service providers as well as service clients in the *ad-hoc* domain.

Parts of the *ad-hoc* domain may be connected to the infrastructure domain via wireless service access points or other cellular technologies such as GPRS, UMTS. The communication between the *ad-hoc* domain and the infrastructure domain can be unidirectional or bi-directional depending on the application.

When the vehicles communicate directly without an infrastructure, using C2C communications *i.e.*, within the *ad-hoc* domain, they form an *ad-hoc network.*

Depending on whether the information is retransmitted at intermediate hops or not, two types of C2C communications can be distinguished: *single-hop* and *multi-hop* [Schitiu & Kihl 2008]. In single-hop communication, information source and destination are within transmission range of each other. On the other hand, information exchange over distances larger than *the transmission range* of a single vehicle can be achieved in multi-hop C2C communication system. In such a scenario, information source and destination are connected by one or more intermediate vehicles "*hops*" that forward the information.

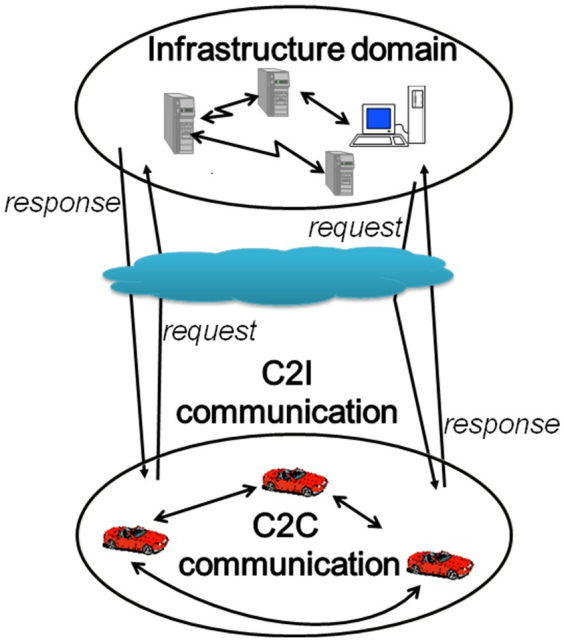

Figure 1.2: *Infrastructure* and *ad-hoc* domains

1.3.2 Wireless access technologies

Various wireless access methods can be used in vehicular environments, namely DSRC (*Dedicated Short-Range Communication*)/WAVE (*Wireless Access in a Vehicular Environment)*, Cellular, WiFi 802.11a/b/g, WIMAX/802.16e, etc. [Lee & Gerla 2010, C2CCC 2007]. Key differences exist between these technologies in terms of: *i*) performance, throughput, transmission range, latency, etc., and *ii*) possible networking scenarios that can be achieved.

While conceptually there are no restrictions on the wireless technology that can be used for C2C communication, there have been recent efforts to standardize protocols tailored to vehicular environments and safety applications developed for such environments [Seada 2008]. The main protocols in this direction are DSRC (*Dedicated Short-Range Communication*) [Jiang *et al*. 2006] and the corresponding standard[1] IEEE 802.11p ("WAVE") [Report_802.11p].

DSRC/802.11p can operate in two modes: *i*) *ad-hoc* mode for C2C communications and *ii*) infrastructure mode for C2I communications. This standard supports vehicle speed up to 190*km*/*h*, nominal transmission range of 300*m* (up to 1000*m*) and data rate of 6*Mbps* (up to 27*Mbps*) [Lee & Gerla 2010, C2CCC 2007]. As regards the transmission range, the maximum value of 1000*m* is theoretical and it is expected that the quality of service of the communication will significantly decrease as the distance between the information source and the destination increases. This is illustrated for example by the simulation results reported in [Torrent-Moreno 2007, Hartenstein & Laberteaux 2008].

It is noteworthy that traditional WiFi 802.11a/b/g standards used in more general mobile *ad-hoc* networking scenarios have a lower nominal transmission range (30*m*-300*m*) [Lee & Gerla 2010, Jr 2009]. Although these standards were not tailored to vehicular environments, their ubiquitous deployment made them attractive for their use in such environments. Some examples of measurement studies for vehicular applications using WiFi 802.11a/b/g standards are reported in [Bychkovsky *et al*. 2006, Gass *et al*. 2006, Hadaller *et al*. 2007, Seada 2008]. The authors measured various indicators characterizing the quality of the communication and the data transfer performance achieved between a moving vehicle and an access point, or between vehicles considering different speeds and various traffic scenarios. Here also it is shown that the quality of service of the communications and the nominal transmission range that can be achieved in practice are variable and depend on the characteristics of the considered vehicular scenarios.

1.3.3 Applications

Several applications are enabled by C2C and C2I communication systems. In this section, we present an overview of such applications. The set of applications presented is not meant to be comprehensive, but rather representative of major classes of applications.

One of the earliest studies on C2C and C2I communication systems was started by JSK (Association of Electronic Technology for Automobile Traffic and Driving) of Japan in the early

[1] At this date, the standard is still under development and it is planned to be released by November 2010.

1980s. Later, California PATH [Hedrick *et al.* 1994] and Chauffeur of EU [Gehring & Fritz 1997] have also demonstrated the technique of coupling two or more vehicles together electronically to form a train of vehicles. Recently, the European project CarTALK 2000 [Reichardt *et al*. 2002] investigated problems related to the safe and comfortable driving based on C2C and C2I communications. Since 2002, with the rapid development of wireless technologies, the number of studies related to C2C and C2I communication systems has noticeably increased. In the following, we provide some synthesis of safety applications which are considered by the various activities and projects which have been introduced here before. This is certainly not an exhaustive list of applications.

1.3.3.1 Local Danger Warning

The local danger warning application is the capability provided to a given vehicle to signal (or broadcast) in real time to the surrounding vehicles, any detected danger. Such messages can also be relayed by other vehicles (a single-hop C2C communication system) and by the sparse/ubiquitous C2I communication systems. Examples of local dangers are:

- Some vehicles in accident. In this case, the vehicle(s) in accident will issue the local danger warning.
- An immobilized vehicle due to a breakdown situation. The concerned vehicle will then issue a local danger warning.
- A vehicle in fire, particularly in risky areas (e.g., in a tunnel or a parking). In this case, the concerned vehicle can start the broadcasting of a local danger warning as soon as an excessive temperature is detected. Such alert can then be repeated by the infrastructure or by other vehicles.
- When a bad weather condition is detected (e.g., some slippery icy condition) by a given vehicle.
- When a bad road condition is detected by a given car (e.g., signaling some oil on the road making it slippery: use of the ESP car function for such detection).
- etc.

1.3.3.2 Public safety applications

This is a generic application to be used by other applications (e.g. Driver advice on safety margin) based on the capabilities for a given vehicle to detect some accident risks. Public safety applications are geared primarily toward avoiding accidents and loss of life of the occupants of vehicles [Yang *et al*. 2004], and outside the vehicle. Collision warning systems have the potential to reduce the number of vehicle collisions in several scenarios.

On highways, frontal collisions with slow moving (or stopped) vehicles are one of the most common types of accidents, often with serious consequences. A vehicle with its airbags deployed, or stopped or a rapidly decelerating vehicle can transmit warning signals to all other approaching

vehicles. Intermediate relays may be used to increase the dissemination range of the warning beyond the direct transmission range.

At intersections, vehicles running red stop lights often result in side crashes [E2213 2003]. If both vehicles to be involved in the accident are equipped with vehicular communication systems, such accidents can be prevented. A similar situation may occur with other types of vehicles (e.g., trains) [E2213 2003]. In some cases, if a collision is imminent, the system may be able to prepare the vehicles for collision (inflate air bags, tighten seat belts, etc.) [E2213 2003].

Safety applications have obvious real-time constraints, as drivers have to be notified before the information is no longer useful. Either a multi-hop C2C communication system or a ubiquitous C2I communication system (sparse C2I communication system for intersections) can be used for these applications. It is possible that, depending on the communication range, a single-hop C2C communication may be sufficient for these applications.

1.3.3.3 Traffic management applications

All intelligent vehicle functionalities rely on the vehicle's knowledge of its surrounding environment. Traffic management applications are focused on improving traffic flow, thus reducing both congestion and accidents resulting from congestion, and reducing travel time [Wischhof *et al*. 2005].

Traffic monitoring can provide high-resolution, localized, and timely information regarding the traffic for several miles around the current location of the vehicle. For this application each vehicle in the system will act as a sensor (determining its current speed), as a relay (if the information is to travel for more than the direct transmission range) as well as a destination (using information from the other vehicles in the system). The information can be used to simply inform the driver or, in more complex systems, to reroute, estimate the time to destination, or even control the traffic by using adaptive speed limits, ramp metering, and so on.

Traffic light scheduling can be significantly improved by using a sparse C2I communication system. Currently, many traffic lights are scheduled either statically or only considering limited information (e.g., by sensing the presence or absence of a vehicle in front of a traffic light). A sparse C2I communication system can provide additional information, such as the length of the queues at the traffic light as well as the number of vehicles expected to arrive in the near future, which can improve the efficiency of schedules.

1.3.3.4 Traffic coordination and assistance applications

Traffic coordination and traffic assistance have been the main research topics of many C2C system projects [Varaiya 1993], even though "*People say they never want the car to take control*," as stated by Cisschke, Group Vice President, Sustainability, Environment and Safety Engineering, Ford Motor Company.

Platooning (*i.e.*, forming tight columns of vehicles closely following each other on highways) has the potential to radically increase the capacity of existing highways. High-speed closed loop control is of paramount importance for this application. Passing and lane change assistance may

reduce or eliminate risks during these maneuvers, since they are often the source of serious accidents.

Clearly these applications require close-range C2C communication system with tight real-time constraints and can be implemented with either a single-hop C2C communication system or an ubiquitous C2I communication system. Both systems can offer similar real-time guarantees and delays if properly designed, although a single-hop C2C communication system may have a slight advantage as it faces reduced contention and direct links.

1.3.3.5 Comfort applications

The main focus of comfort applications is to make travel more pleasant [Sugiura & Dermawan 2005]. This class of applications may be motivated by the desire of passengers to communicate with either other vehicles or ground-based destinations such as Internet hosts or the public service telephone network.

Blackboard application shall serve to distribute information which is relevant for a certain geographic area or to make a communication targeted with another vehicle. This type of applications allows localized communications (potentially multi-hop) between two vehicles. Voice, instant messaging, or similar communications may occur between the occupants of vehicle caravans traveling together for long distances, or between law enforcement vehicles and their "victims." Note that this application does not scale to large network sizes. The notion behind is that a lot of information can be broadcast into the network but the user only wants to see the information which is relevant for him. A major share of information is relevant for a certain geographic area only, such as speed traps, fuel prices, restaurant offers, warning about a slippery road, etc.

This special type of application assumes that the information is not necessarily permanently repeated by its source but sent only once (or with larger intervals). This means that the information is to be distributed by the cars on the road, and that the cars on the road have to store the data and to make sure it will not be lost. A car may have received the message outside the relevant geographic area already but it is displayed only upon entering the area. This may be relevant for cases where no car is in the considered region but the message shall still be preserved in the network.

Finally, there are many other comfort applications, and many nonstandard systems do exist and work well. For example: i) tolls for roads and bridges can be collected automatically, ii) parking payments can be made promptly and conveniently, iii) repair and maintenance can be recorded at the garages performing them, iv) multimedia files such as DVDs, music, news, audio books, pre-recorded shows can be uploaded to the car's entertainment system while the car is in the garage.

Communications in such applications may be implemented using either a C2C or an ubiquitous C2I communication system. Some comfort applications such as tolls for roads and parking can be enabled by a sparse C2I communication system.

1.3.4 Challenging characteristics of vehicular ad-hoc network environments

Vehicular applications based on *ad-hoc* networks exhibit several challenging characteristics. While some of these characteristics are shared with traditional *Mobile Ad-hoc NETworks* (MANETs) [Chlamtac *et al*. 2003], there are some characteristics that differentiate C2C and C2I communication systems from the common assumptions made in the MANET literature. A detailed discussion on these characteristics can be found in [HIDENETS 2007a, HIDENETS 2007c, and Li & Wang 2007].

Examples of difficulties raised by some of these challenging characteristics from the dependability modeling and evaluation perspective are discussed in [HIDENETS 2007a] and are summarized in the following.

- *Highly dynamic topology and intermittent connectivity*: Due to high speed of movement of vehicles, the topology of vehicular *ad-hoc* based networks is always evolving. The rate at which topology changes occur is much higher compared to traditional MANETs. For example, if we assume that the wireless transmission range of each vehicle is 250*m*, in the worst case, if two cars with the speed of 90*km/h* (25*m/sec*.) are driving in opposite directions, the link will last only for at most 10*sec*. Clearly, the rate of topology changes and the duration of connectivity between vehicles highly depend on the characteristics of the traffic scenarios and the communication environment under investigation (type of environment —freeway *vs*. urban-city, vehicle speed, density, transmission range, etc.).

- *Heterogeneity of the communication environments*: Vehicular networking environments as well as MANETs include wireless *ad-hoc* networks, wireless infrastructure based networks, and also wired networks. The characteristics of these network domains differ in terms of dynamicity of the topology as discussed above, and also from the point of view of the dependability related properties and performance of the communication links. Heterogeneity also concerns differences between traffic conditions in various mobility environments. For example, in highway traffic scenarios, the environment is relatively simple and straightforward (e.g., constrained one dimensional movement), while in urban cities the topology of the environment is more complex and the presence of buildings, trees and other obstacles may significantly affect the quality of the wireless communications.

- *Mobility patterns*: The definition of realistic mobility models is one of the most critical and difficult aspects of the simulation and analysis of applications and systems designed for mobile environments. Two possible types of mobility patterns can be used and combined: *i*) *traces* obtained by means of measurements of deployed systems, and *ii*) *synthetic models* that correspond to mathematical models abstracting specific characteristics of nodes movements in particular environments. Recent surveys on mobility models and examples of publicly available data repository of mobility traces are reported in [Camp *et al*. 2002, Musolesi & Mascolo 2009]. However, most of the popular mobility models such as Random Waypoint Model (RWP) are not suitable to describe mobility in vehicular environments. Recently, this has motivated research aiming at the development of more realistic mobility models and simulators that take into account the specific characteristics of vehicular environments in terms of topology, velocity, traffic

density, signal propagation characteristics, etc. (see e.g., the survey published in [Härri *et al*. 2009]).

- *Complexity*: The analysis of vehicular *ad-hoc* based networks and applications generally requires the investigation of a large number of components and interaction scenarios. The number of components to be taken into account depends on the level of detail of the analysis and the quantitative measures to be evaluated. The complexity of the evaluation can also result from the existence of a large number of failure modes and recovery and maintenance scenarios to be taken into account.

- *Variety of threats*: MANETs in general, including vehicular mobile *ad-hoc* systems, have to cope with a variety of threats and address both accidental and malicious faults (attacks and intrusions). Such faults may lead to a variety of failure modes and affect the behaviour of the systems under investigation at different levels of the architecture (hardware and software processing and storage components, communication links, middleware, application level services, etc.). Besides addressing these faults from the design perspective, it is also important to take them into account in the dependability analysis and evaluation activities.

- *Energy and storage capacity*: A common characteristic of nodes in vehicular *ad-hoc* networking environments is that nodes have ample energy and computing power (including both storage and processing), since nodes are vehicles and not handheld devices. As a consequence the problems related to energy generally considered in MANETs literature are less critical in the context of vehicular environments.

The characteristics identified above call for suitable and efficient approaches to address the corresponding challenges at the architectural level as well as from the perspective of dependability assessment. In this dissertation, we focus on the latter aspect. Some examples of architectural solutions for providing error detection, recovery mechanisms, and fault tolerance, at the communication level or the middleware level close to the applications, have been developed in the context of the HIDENETS European project. More details can be found in [HIDENETS 2007b, HIDENETS 2008].

Considering dependability evaluation, system and application designers are generally interested in understanding and quantifying the impact of failures at different levels of decomposition of the systems considered. A single technique is generally not sufficient to perform such evaluations and it is generally necessary to combine different techniques and tools. This is due in particular to the complexity of the analysis and the need to take into account different characteristics of the target systems that cannot be captured efficiently by a single method (description of mobility patterns, modeling of components failures and recovery strategies, etc.). In the following section, we present the holistic evaluation approach that is aimed at fulfilling this objective that has been developed in the context of the HIDENETS European project in which our study has been partially carried out. Also, we outline our main contributions in the context of the holistic approach that are detailed in next chapters of this dissertation.

1.4 The HIDENETS holistic dependability evaluation approach

HIDENETS "*HIghly DEpendable ip-based NETworks and Services*" [HIDENETS 2006a] addressed the development and the validation of architectures and methods making it possible to ensure the dependability of applications and services implemented in a mobile context with *ad-hoc* networks. The investigations included networking scenarios consisting of *ad-hoc*/wireless (multi-hop) domains as well as infrastructure network domains. Applications and use case scenarios from the automotive domain [Bondavalli *et al.* 2010], based on Car-to-Car communications with additional infrastructure support, have been used as case studies to support the development of: *i*) architectural dependability solutions at the middleware and communication levels, and *ii*) efficient methods for the verification and quantitative assessment of the corresponding applications from the dependability view point.

In this section, we outline the modeling approach proposed in the context of HIDENETS to perform the quantitative dependability evaluation of mobile – based applications.

1.4.1 Overview

The key challenges discussed in section 1.3.4, call for the use of a holistic evaluation framework where the synergies and complementarities among several evaluation methods can be fruitfully exploited. To cope with the complexity of the target applications, dependability evaluation can be carried out at different decomposition levels, and different techniques can be used to cover various aspects that need to be taken into account (description of the mobility and the dynamic evolution of the system, modeling of the components and subsystems failure modes to be analyzed and their impact, etc.). The holistic framework is needed in this context to integrate and to combine the results provided by the individual techniques. The combination of different evaluation techniques can be used for different purposes.

- *Cross validation:* A partial solution validates some assumptions introduced to solve another sub-problem, or validates another partial solution (e.g., a simulation model can be used to check if the duration of an event in an analytic model has an *exponential* or a *Pareto* distribution).

- *Solution feedback:* A partial solution (or a part of it) obtained by applying a solution technique to a sub problem is used as input to solve another sub-problem possibly using a different technique (e.g., a critical parameter in an analytic model is obtained using experimental evaluation).

- *Problem refinement:* A partial solution gives some additional knowledge that leads to a problem refinement (e.g., the architecture of a component changes since it is recognized to be a system bottleneck).

In the holistic evaluation approach, the experimental evaluation techniques and the simulation approaches are used to capture the low level details of the system (at the architecture and communication layers), while the model-based approaches are used to assess the system at higher abstraction levels and to provide quantitative measures for end-to-end scenarios.

It is clear that the system decomposition is not unique, as we can identify different system decompositions corresponding to different levels of abstraction. The higher the level of detail required for capturing the system behavior, the higher is the complexity of the system to be modeled and solved. Therefore the choice of particular system decomposition is of primary importance, and it is always a *trade-off* between faithfulness of representation of the real system behavior (with respect to the measures of interest) and the capability to solve the corresponding models. In the following we present a decomposition approach that statically focuses on the various levels of abstractions that can be used to represent a system. It is noteworthy that a similar type of decomposition has been proposed e.g., in [Kaâniche *et al*. 2003, Martinello 2005] for the dependability modeling of web-based applications running in the fixed infrastructure domain.

1.4.2 Abstraction-based system decomposition

As illustrated in Figure 1.3, the system under study can be analyzed at different levels of abstraction: each level captures a specific aspect of the overall system behavior and it "*interacts*" with the other levels through some well-specified interfaces. Such interfaces mainly define the inputs they require from other abstraction levels, as well as their outputs. The dependability evaluation step is directly related to the hierarchical abstraction-based system description as presented in Figure 1.3. A submodel is associated to each abstraction level to evaluate dependability measures characterizing the quality of service delivered by the different entities of the corresponding level taking into account the outputs evaluated from lower level submodels. The outputs of a given level are used in the next immediately upper level to compute the dependability measures associated to this level. In this figure, it is assumed that the dependability assessment at the different levels is based on modeling (using analytical models or simulation). However, as explained in the previous section, experimental techniques can also be used to estimate some dependability related parameters or to validate some assumptions considered in the different submodels. Also, various modeling techniques (Fault trees, Markov chains, SANs, etc.) or simulation tools (NS2 [Network_Simulator_2], VanetMobiSim [Fiore *et al*. 2007], etc.) can be used for the evaluation of the quantitative measures associated to each level. The selection of the most appropriate technique depends primarily on the granularity of the analysis, the assumptions considered, the kinds of dependencies between the elements of the considered level, and the quantitative measures to be assessed. A detailed overview of evaluation techniques, methods and simulation tools that can be used in the context of mobile based systems can be found in [HIDENETS 2007a].

In the following section, we describe the abstraction levels depicted in Figure 1.3. It is noteworthy that the overall framework can be used to provide end-to-end dependability measures taking into account lower level components of the architecture up to the user level.

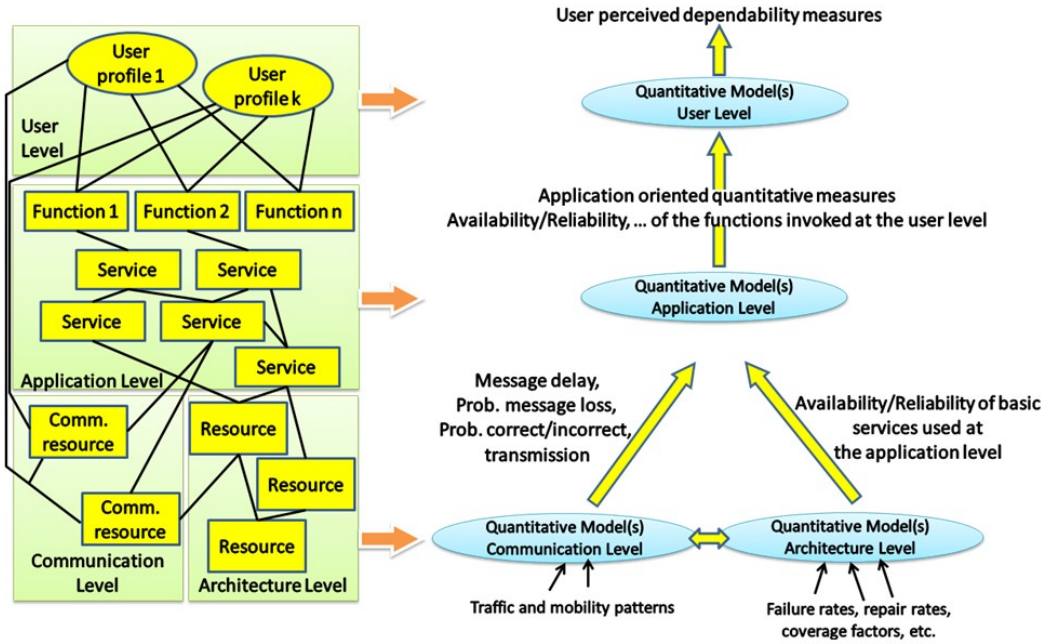

Figure 1.3: Abstraction based decomposition and hierarchical dependability assessment

1.4.2.1 User Level

This level provides dependability attributes as perceived by the users. It describes the users' profiles, that is, how the users interact with the application and how their requests are mapped to the different functions of the application. Accordingly, the dependability attributes perceived by the users depend on the dependability attributes of the corresponding functions. A user level is needed to account for different classes of users having different behaviours and different requirements. Mobility scenarios and application utilization profiles are just some examples of users' characteristics that can differentiate a user's class from another.

- The expected inputs are the outputs produced by the application level.
- The expected outputs are high-level dependability attributes related to the user's perspective. Examples of user-oriented measures include availability, safety or reliability reflecting the considered scenarios and operational profiles.

1.4.2.2 Application level

This level describes the system behaviour from the application point of view. The applications differ in their technical properties, their mechanisms, their interfaces, and they can impose different communication and middleware level requirements. The user-interfaces consist of a set of functions, and each function corresponds to a set of middleware) services offered by the architecture for its implementation. Each function may depend on several services and the services may depend on each other.

- The expected inputs are the outputs produced by the architecture and communication levels.

- The expected outputs are QoS and dependability related measures associated to each function invoked at this level, like availability, reliability, performance, etc. Some of these measures could be provided as input to the user level.

1.4.2.3 Architecture level

This is the part of the system capturing the behaviour of the main hardware and software components (resources) that can affect the application level measures, including the error detection and recovery mechanisms implemented in the system to support dependability. It describes how the functions and services of the application level are implemented on these resources. This layer also includes the middleware that abstracts some details of the underlying layers for the application running on top of it.

- The expected inputs are some low level parameters concerning hardware, software or basic services, such as failure rate, error latency, repair rate, error propagation probability.

- The expected outputs are some medium-level dependability-related attributes, like availability and reliability of some services used at the application level by various applications.

1.4.2.4 Communication level

It captures the communication aspects of the system that can affect the application level. It addresses the link layer (possibly considering several types of networks (e.g., WLANs, UMTS and GPRS) and considering several networks' properties).

- The main expected inputs are the traffic and mobility patterns and a set of assumptions introduced to hide low-level system details that are not the target of the analysis.

- The expected outputs are communication level measures like: message delay, probability of lost message, probability that a message is incorrectly emitted or is omitted. Such measures could be mean values or complete distributions, and could be used at the application-level or at the architecture (Middleware) level.

Indeed, the communication level can be seen as a special case of the architecture level focusing on some communication related aspects that might affect the dependability-related attributes perceived at the application and the user levels. Also, it is noteworthy that sometimes we might need to analyze and assess some characteristics of the application level as a function of communication related aspects without explicitly modeling in detail the architecture level. Similarly, depending on the level of detail considered in the description of the studied system, it could be sometimes more efficient to model the user and the application levels as a single abstraction level.

More detailed information about this approach is presented in [HIDENETS 2007a, Bondavalli *et al*. 2010]. It can be noticed that an evaluation workflow based on UML modelling has been also proposed to support the implementation of the holistic approach and the integration of several evaluation tools and model transformation steps [Kovacs *et al*. 2008]. Several examples are illustrating the application of the proposed framework for end-to-end.

1.5 Contributions of the dissertation work in the context of HIDENETS

The holistic evaluation framework presented in Section 1.4 results from a collective work by HIDENETS partners that have been involved in the activities related to dependability assessment during the project. Besides participating to the elaboration of the conceptual framework, our contributions concerned the illustration on selected case studies of the feasibility and usefulness of the proposed framework for the quantitative dependability assessment of mobile-based applications relying on C2C and C2I communications. As illustrated in Figure 1.4, one of the main objectives pursued here is to take advantage of the cross fertilization and interactions among the different individual techniques, thus concretizing the holistic approach, considering two complementary aspects:

1) The characterization of mobility scenarios and the estimation of connectivity parameters for C2C and C2I communications. Various techniques are investigated including simulation, analytical connectivity models and the statistical processing of real connectivity traces. The connectivity parameters of interest correspond for example to the rate at which vehicles meet, the duration that such vehicles stay in connectivity range, and the rate at which a vehicle meets an access point of the fixed infrastructure.

2) The development of high-level dependability models allowing the description and quantitative assessment of the impact of component failures and restorations on system level dependability properties, taking into account the information derived from the connectivity analysis of mobility scenarios. The construction and processing of the dependability models are based on the Stochastic Activity Networks formalism, and the associated Möbius tool.

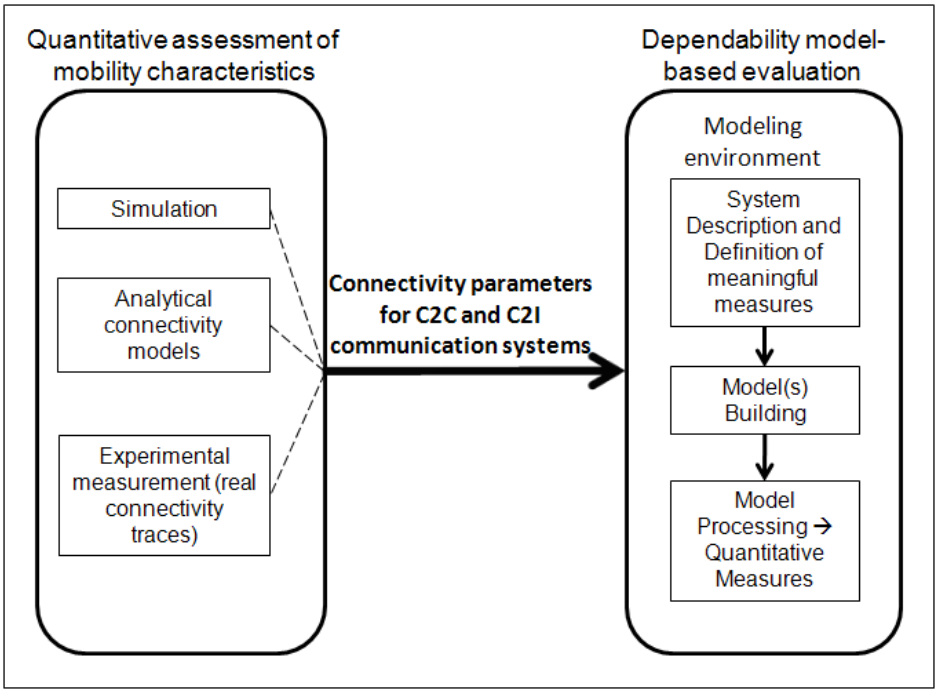

Figure 1.4: A simplified view of the modeling process and its interactions with the quantitative assessment of mobility characteristics

We have considered three case studies in our research namely, a replication service in the ad-hoc domain, a virtual black-box application, and a platooning application. These case studies exhibit various dependability characteristics and challenges that are typical of vehicular applications using C2C and C2I communications.

A chapter is dedicated to the dependability modeling and evaluation of each of these case studies. In the following, we summarize the main characteristics of the three case studies that are considered in the dissertation.

1.5.1 Replication service

The replication service is a middleware service designed to improve application dependability by providing failover support. The replication service is mainly designed for use within the *ad-hoc* domain.

The replication service provides the application with a memory area that is shared between the cooperating nodes. It has the ability to select nodes in the network capable of establishing a data sharing cooperation, and the selection is done to provide a stable set of replica nodes. The replication service is able to handle crash faults of service provisioning nodes. To improve the dependability of an application, the replication service dynamically forms groups. These groups are formed and kept coherent by changing group memberships. Group members (nodes) are selected based on geographical properties and communication properties. The group members are

selected in a way that optimizes the data consistency. Towards the application, the replication service provides a simple interface with functions to initiate a shared memory area and to stop using a previously initiated shared memory area. Furthermore it has functions to read and write data to and from the memory area. An example of an application using the replication service is a shared whiteboard server. The users of the whiteboard send messages to the whiteboard server containing their changes to the board. To have a consistent whiteboard, all changes need to be committed to the board in the shared memory area where it will be replicated to the replica servers in the group. Communications in this case study may be implemented using either a C2C or a ubiquitous C2I communication system.

The main goal of the evaluation models for this case study presented in Chapter 3 is to characterize the service availability and the data consistency among the replica of a group, taking into account the dynamic evolution of the topology of the network in the *ad-hoc* domain.

1.5.2 Virtual black-box application

Recent developments in automotive systems recommend storing historical information related to a vehicle and its environment, that can be retrieved in case of a problem, in a manner that is similar to the black box of an aircraft. Typically, the data recorded just before the accident is critical and is very valuable for accident investigation. However, instead of using a classical hardware based black-box that could induce a high cost, a more cost effective solution would be to use a software based mechanism that consists in storing the recorded data on a dedicated server at the fixed infrastructure. To protect the data against potential losses before an access to the fixed infrastructure is available, the data can be replicated and temporarily stored on neighboring encountered vehicles using wireless communication technologies, before being permanently saved on the server.

The aim of the dependability modeling study described in Chapter 4 is to analyze and evaluate the availability of the historical data recorded in the virtual black-box, taking into account different possible data replication strategies, and various mobility scenarios.

1.5.3 Platooning application

The third case study addresses the safety modeling and evaluation of Automated Highway Systems (AHS), based on automatically controlled *platoons* of vehicles. The objective of the AHS is to reduce congestion and increase the throughput and safety of the highway by adding automation to the vehicles and the roadside without having to build new roads. A *platoon* is a series of coordinated vehicles that are moving in the same direction on a highway [Godbole *et al*. 1996], controlled by the vehicle at the head of the platoon. The vehicles are driven by more or less automated agents, interacting in a multi-agent environment [Hallé & Chaib-draa 2005]. Switching to manual driving is possible under specific circumstances. Platooning requires C2C communication and may include C2I communication. The application combines vehicle data with position and map data. Longitudinal control of the vehicle is provided in order to maintain the short-range headway following within a platoon similar to adaptive cruise control. Lateral control via automated steering provides lane-keeping and lane-change maneuvers of platoon vehicles in a coordinated manner.

In chapter 5, we address the safety of AHS based on platooning applications implemented in a mobile context with ad-hoc networks. We analyze the impact on safety of the strategy used to coordinate the vehicles' operations, inside each platoon and between platoons, when vehicles enter or exit the highway, or when manoeuvres are carried out to recover from failures affecting the vehicles or their communication. To cope with the complexity of the studied system, a compositional modeling approach based on stochastic activity networks is developed. The proposed models take into account the configuration of the platoons and their dynamic evolution and describe different failure modes affecting the vehicles and associated recovery manoeuvres considering various possible coordination strategies between the vehicles, inside each platoon and between platoons.

1.6 Conclusion

This chapter introduced the concepts of dependability as well as the required background including some approaches and techniques for mobile-based application dependability evaluation, considering the particular case of vehicular applications. Our work relies on modeling in order to provide a quantitative approach for evaluating the dependability of aforementioned case studies. The models presented along the chapters are based on Stochastic Activity Networks (SANs) that are well suited to master the complexity of the models through the composition and replication of sub-models.

An overview of a hierarchical and stepwise modeling approach aimed at the evaluation of mobile applications was briefly discussed. This approach is based on the evaluation of the target application at different abstraction and decomposition levels using a combination of various evaluation techniques an integrated way in order to provide end-to-end evaluations. In this dissertation, we mainly focus on the combined modeling of connectivity dynamics under various mobility scenarios and the integration of such information in higher level dependability models using SANs. Three case studies with different characteristics and dependability properties are presented in the following chapters for illustration

As long as these case studies are specific and cover different domains, a discussion of related work is included in each of the corresponding chapters.

The following chapters of the dissertation are organized as follows. Chapter 2 deals with the characterization of the stochastic processes underlying connectivity parameters for different mobility scenarios considering freeways and urban city environments. The parameters derived from these analyzes are used in Chapter 3 and Chapter 4 for the dependability evaluation of the replication service, and the virtual black box application. The dependability modeling and evaluation of the platooning application is addressed in Chapter 5.

Chapter Two: Estimation of Connectivity Dynamics in Vehicular *ad-hoc* Wireless Networks

"*If you want me to play only the notes without any specific dynamics, I will never make one mistake.*"
Vladimir Horowitz, American, Musician Quotes

This chapter focuses on parameters that may affect the connectivity between the vehicles in dynamic vehicular *ad-hoc* wireless networks. Its aim is to characterize the distribution of some connectivity parameters based on analytical proofs, simulations, and real life traces. The results of this chapter will be used for the dependability assessment of the applications presented in the following chapters.

2.1 Introduction

The recent few years have witnessed noticeable effort in research, development and standardization of the protocols and systems needed for realizing C2C applications. The quality of service and the dependability provided by such applications highly depend on the connectivity characteristics of C2C and C2I communications. Relevant connectivity parameters include in particular the rate at which connections can take place and the duration of the communications. These connectivity parameters depend heavily on the considered mobility environment (*highway*, *urban-city*, etc.). Characterizing C2C and C2I connectivity is a challenging problem both theoretically and experimentally due to the high mobility of the participants and varying connectivity conditions in such environments. Most of the existing connectivity analysis studies are based on theoretical techniques and simulations due to the difficulties of conducting real-world experiments on the road.

In this chapter we will focus on the analysis of the stochastic processes describing the moments in time at which a particular vehicle enters a new *k*-hop connectivity relation with another vehicle or with an access point of the fixed infrastructure. Due to requirements of route stability and low end-to-end delay, low values of *k* are of particular interest. Hence, large parts of the chapter focus on direct neighborhood relations with single hop communication ($k = 1$). Link-layer connectivity can thereby be established as soon as the distance between two vehicles is below a certain value [Khadar & Simplot-Ryl 2007].

We will use three complementary techniques: analytical proofs, simulation, and publicly available real life traces.

This chapter is organized as follows. Section 2.2 discusses work related to the quantitative evaluation of mobility connectivity dynamics in vehicular traffic scenarios. Section 2.3 and 2.4 present the connectivity analysis results starting with a simple 1-dimensional lane highway traffic scenario with independent movement of vehicles and then addressing more complex 2-dimensional multiple lane highway scenarios, with and without the independent movement of vehicles assumption. Section 2.3 outlines the reference highway mobility scenarios and the main stochastic processes investigated in our study. In particular, in Section 2.3, it is shown that under the assumption of independent movement of vehicles at opposite directions but with constant speeds and initial spatial *Poisson* placement of vehicles on a 1-dimensional space, the time

between two cars encounters (C2C encounter times) is *exponentially* distributed. The encounter rate can be expressed as the product of the vehicle density and the expected value of the absolute relative speed of the other vehicles. Subsequently, at the end of Section 2.3, we investigate, using simulation, more complex 2-dimensional multiple lane highway traffic scenarios with independent movement of cars, that deviate from the original assumptions of the proof in the beginning of Section 2.3. It is concluded that the *exponential* distribution still provides a good fit to C2C encounter times. Additional results are derived as regards the impact of the vehicle connectivity range, and the distribution of C2I encounter times and C2C connectivity duration. Finally, Section 2.4 analyses the impact of the car movement independence assumption by considering more realistic traffic scenarios with dependencies based on publicly available real life traces in *highway* and in *urban-cities* mobility environment. In case of *urban-cities* environment, it appears that the C2C encounter times are best described by a *Pareto* distribution. A summary of the scenarios investigated in this chapter and the corresponding conclusions are given in Section 2.5.

2.2 Related work

Most of the existing work on connectivity analysis in mobile *ad-hoc* networks focuses on static snapshots of the node placement. For example, [Miorandi & Altman 2006] analyses different connectivity metrics under the assumption that the node placement can be described by a spatial renewal process. A more general setting using percolation theory is presented in [Cheng & Robertazzi 1989] to analyze broadcast propagation in mobile *ad-hoc* networks. [Hansen *et al.* 2008] examined one dimensional vehicular freeway scenario, considering cases of spatial correlation described by *Markovian Arrival Processes*. The metrics thereby include distribution and moments for the number of direct neighbors (single-hop) and for the number of reachable nodes *via* multi-hop connections. Furthermore, spatial distances and hop-count distributions are characterized and analyzed *via* numerical results.

Considering the particular case of vehicular applications, a few studies have addressed the characterization of connectivity parameters. Most of these studies have been carried out in the perspective of providing insights for the assessment and analysis of vehicular networks routing protocols. As a consequence, connectivity is generally analyzed considering multi-hop connections. The survey presented in [Hui 2005] provides some insights about some existing work dealing with this topic. Some examples of related work in this context are discussed in the following.

To analyze the impact of different mobility models on the connectivity characteristics of mobile *ad-hoc* networks and the performance of routing protocols, [Bai *et al*. 2003] proposed the IMPORTANT framework. This framework proposes various protocol independent metrics to capture interesting mobility characteristics, including spatial and temporal dependence and geographic restrictions. In addition, a set of mobility models including Random Way Point [Hyytiä *et al*. 2006], Reference Point Group mobility [Hong *et al.* 1999], and a Freeway model are used to generate test cases scenarios to assess the proposed metrics. In particular, the Freeway mobility model generates traffic for multilane two directional highways with temporal and spatial dependencies between car movements. This framework inspired the work carried out in [Sadagopan et al 2003] to derive statistically based on the simulation of the probability distributions of a single hop and multi-hop connection duration between mobile nodes, under the

four mobility models included in IMPORTANT. These metrics are called *link* and *path* duration respectively. This study suggests that at moderate and high velocities of the nodes (10 to 30 *m*/*sec*), the path duration can be described by an *exponential* distribution for the considered mobility models. For lower velocities, the distribution exhibits a multi-modal behavior.

Highway mobility scenarios have also been considered in the study reported in [Artimy *et al*. 2004] in which the impact of various road dynamics on the connectivity of vehicles has been investigated. Using a traffic micro-simulator that generates vehicles movement in a multi-lane unidirectional highway, the authors analyzed *i*) the ability of the network to maintain an active connection between a pair of vehicles and the connection lifetime as a function of vehicle density, velocity, and connectivity range and *ii*) the number of the highway lanes. In particular, it is observed that the probability distribution of connection lifetime follows a *Power Law* function, where a connection between two vehicles exists whenever there is an active route between them. In [Artimy *et al*. 2004], the characterization of the duration for single-hop connections is not addressed.

In the papers discussed above, it is assumed that the vehicle density is constant. However, such assumption may not reflect real traffic situations as observed e.g., in [Bai & Krishnamachari 2009]. Based on statistical analysis of empirical data and analytical modeling, the authors studied the effect on connectivity in highway mobility scenarios of traffic variation over time, geographic locations and over technology adoption phases. The empirical data used in this study corresponds to traffic measurements from two sources: *i*) a freeway in an urban area of Berkeley, United States and *ii*) an expressway in suburban area of Toronto, Canada. It is observed that while traffic in a single direction exhibits a homogeneous behavior, this behavior differs when considering opposite directions.

In comparison to this existing work, only a few studies have investigated the characterization of the encounter times and connectivity duration parameters. Research in this area is recent, in particular in the context of vehicular *ad-hoc* networks. It is mostly simulation based, complemented with the processing of real life traces and in some cases with analytical proofs. Some examples of related work in this area are discussed in the following.

Recent works have studied encounter times based on mobility traces. The main objective is to understand the connectivity characteristics under realistic mobility scenarios and improve the fidelity of mobility models compared to the traditional mobility models such as Random Way Point. Indeed, as observed in [Fiore & Härri 2008] traditional mobility models generally used in the analysis of the performance characteristics of mobile based applications and protocols are not representative in particular of vehicular mobility scenarios.

[Chaintreau *et al*. 2007] characterized all-node pairs aggregate encounter times from six mobility traces of pedestrians carrying Bluetooth devices. They observed that encounter times can be approximately distributed according to a *Power Law*. Based on this observation, they analytically studied the performance of different data forwarding algorithms under such a model.

Considering vehicular *ad-hoc* networks, [Zhang *et al*. 2007] studied the distribution of the encounter time based on mobility traces collected from the UMass DieselNet, an operational vehicular Disruption-Tolerant Network (DTN) consisting of *WiFi* nodes attached to buses [Burgess *et al*. 2006]. The results are based on data collected during 55 *days* of operation. The

authors derived a model of encounter times that can be used to generate synthetic traces for simulation based analyses. However, no specific analytical distribution was derived from the data. A particular focus in this study concerned the analysis of the granularity level of the model and its impact on the DTN performance.

As regards analytical-based modeling studies, the results reported in [Karagiannis *et al*. 2007, Spyropoulos *et al*. 2008] apply a methodology to derive closed form equations for popular mobility models for encounter-based protocol and urban mesh networks. Nevertheless, as mentioned in [Karagiannis *et al*. 2007, Spyropoulos *et al*. 2008], it appears that there are still different opinions about which of the distributions (*Pareto*, *exponential* tail, or *Poisson*) applies for describing the encountering process, and the debate regarding this point is ongoing.

Related work shows that the analysis of the connectivity characteristics of vehicular *ad-hoc* based applications and protocols is still an open area of research. The studies focusing at the characterization of encounter times and encounter durations probability distributions are quite recent. These studies have been mainly carried out during the last five years in parallel of our research presented in this thesis. Most of the results of these studies have been used to analyze the performance characteristics of mobile *ad-hoc* routing protocols and Delay-Tolerant Networks. In this thesis, our objective is to perform detailed analyses of the connectivity characteristics of different vehicular mobility scenarios taking into account, on one side, the impact of several parameters such as the connectivity range of the vehicles, their density, and speed as well the density of the C2I access points. On the other side, we are taking into account wide specifications to cover most of the *ad-hoc* network technologies (e.g., about 5*m* to 300*m* as vehicle transmission range and 20*m/sec* to 35*m/sec* as vehicular speed). The results are aimed to be used to analyze the dependability of various vehicular applications as illustrated in the following chapters. In our analysis, we combine the different techniques including simulations, traces and analytical proofs.

2.3 Freeway mobility and connectivity model

To analyze the dependability of the vehicular applications investigated in our study, we are interested in the characteristics of C2C and C2I encounter opportunities, *i.e.*, how often do they occur, and how long. These characteristics can be analyzed through the modeling of the three following connectivity parameters and the statistical analysis of their probability distribution:

- *C2C encounter time*: corresponds to the time interval between two cars encounters in the *ad-hoc* domain. An encounter occurs when a vehicle gets within the encountered vehicle's connectivity range.

- *C2C connectivity duration*: corresponds to the time during which an encountered vehicle remains in another vehicle's connectivity range.

- *C2I encounter time*: corresponds to the time during which a vehicle is out of the connectivity ranges of two successive access points.

It is worth to mention that other terminologies for these parameters do exist in the literature. For example, [Chaintreau *et al*. 2005] used the concepts of *contact times* and *contact duration* to analyze connection opportunities in the field of opportunistic mobile *ad-hoc* networks and delay

tolerant networks. In [Bai *et al.* 2003, Sadagopan *et al.* 2003] the authors use the concept of *link duration* to refer to the duration of a connection between two mobile nodes. This parameter is called *link lifetime* in [Samar & Wicker 2004]. Also, we can mention the more complete set of connectivity and mobility metrics for *ad-hoc* networks presented in [Xu *et al.* 2007], and the proposed analysis framework that is aimed at clarifying the relationships between the various presented metrics. Such metrics are generally used to support the performance analysis of mobile *ad-hoc* networks considering routing protocols, in particular.

The rest of this section is organized as follows. Section 2.3.1 presents the mobility scenario used for the simulation to analyze and estimate the aforementioned vehicle connectivity parameters. Connectivity parameters analysis using an analytical proof for a straight highway is investigated in Section 2.3.2. Section 2.3.3 introduces the finite lane 2-dimentional freeway scenario, and then the later three successive subsections 2.3.4, 2.3.5, and 2.3.6 outline and discuss the results obtained from the simulation for the three vehicle connectivity characteristics, respectively.

2.3.1 Mobility scenario[2]

Figure 2.1 shows an abstraction of a vehicular freeway scenario, used in our context for the estimation, based on simulation, of the properties of the connectivity parameters characterizing vehicular communication scenarios. We consider a long straight piece of freeway (of width 2*W*) with movements in two directions. The considered piece of freeway has a finite length *L*, with (*L* » *W*).

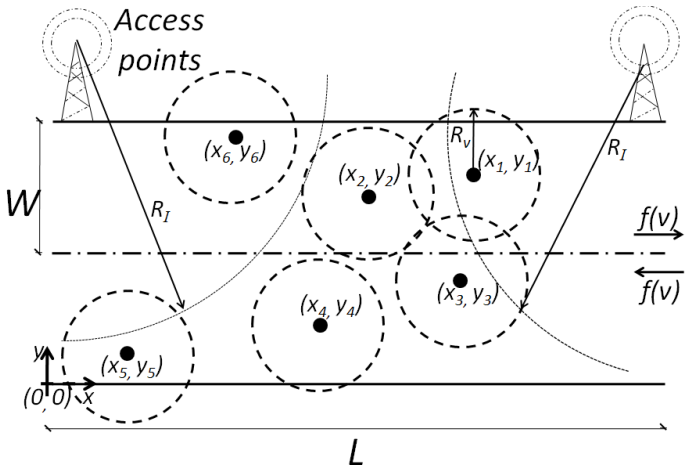

Figure 2.1: A two lane freeway mobility scenario

To characterize connectivity dynamics, we consider a reference vehicle located at position (x_1, y_1), and we analyze the parameters defined above, and refined in the following:

[2] The work presented in this section is a joint work with Jakob Rasmussen, Erling M. Møller, and Hans-Peter Schwefel during my stay as visiting researcher at Aalborg University, Denmark (see [Hamouda *et al.* 2009b]).

- α: is the C2C encounter rate which equals to the inverse of the mean of the C2C encounter time. The time between two C2C encounters is represented by the time instances at which other vehicles enter (single-hop or *k*-hop) in connectivity to the reference car. For *k*=1, these instances correspond to new vehicles coming into connectivity range R_v of the reference vehicle. The mean time between such encounters is α^{-1}.

- ω^{-1}: is the mean of the connectivity duration, *i.e.*, the time during which an encountered vehicle remains in the reference vehicle connectivity range R_v. The connectivity duration is the random variable corresponding to the time-interval starting when another vehicle enters into *k*-hop connectivity range of the reference vehicle and ending when this vehicle leaves *k*-hop connectivity. Figure 2.2 shows two examples of how this connectivity duration T_{on} could be obtained. Figure 2.2 *(a)* shows T_{on2} for a vehicle (x_2, y_2) in the same lane and direction as the reference vehicle (x_1, y_1). Figure 2.2 *(b)* shows the T_{on3} for a vehicle (x_3, y_3) in the opposite direction.

- β: represents the C2I encounter rate which equals to the inverse of the mean of the C2I encounter time. The C2I encounter time corresponds to the time during which a vehicle is out of the connectivity ranges of two successive access points (the connectivity range is R_I). The mean time between encounters of two successive access points is β^{-1}. It corresponds to the mean of the variable T_{off} for a given vehicle, as represented in Figure 2.3. T_{off} depends on the *y-position* of the vehicle and its *speed*.

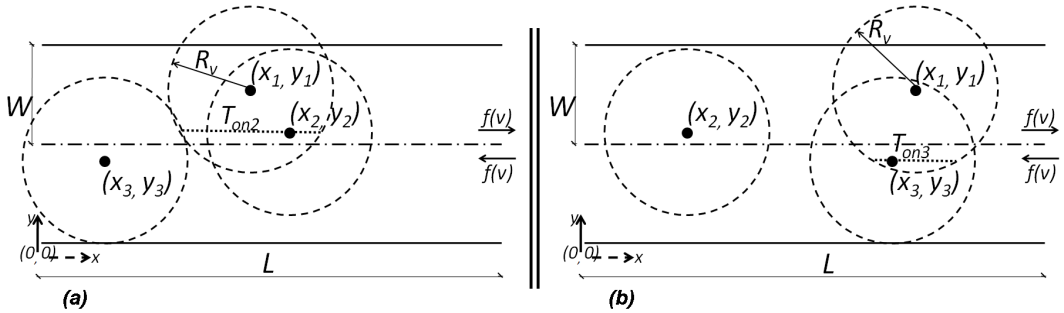

Figure 2.2: C2C connectivity range R_v, C2C encounter rate, and C2C connectivity duration

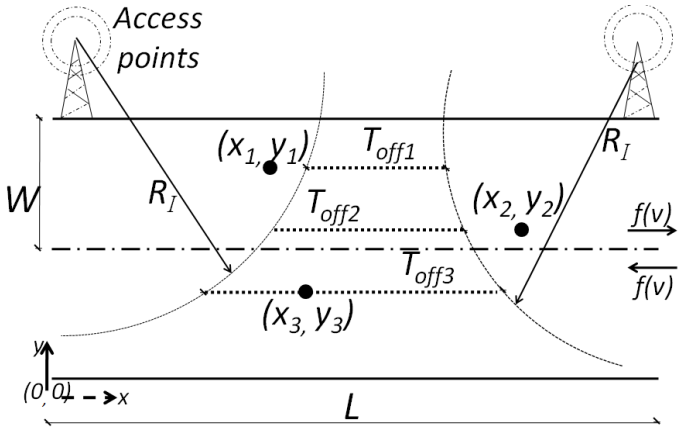

Figure 2.3: C2I encounter rate

In order to avoid edge effects in the simulations, we assume that cars that leave on one side enter at the corresponding point on the other side, *i.e.*, the long piece of road can be seen wrapped around a cylinder. Vehicles have a constant (*absolute*) speed v_i which is assumed to be *independent* and *identically distributed* (*i.i.d*) with *probability density function* (*p.d.f.*) *f(v)*. In order to focus on understanding the impact of the geographic mobility model, we adopt the approximation that two nodes can communicate on a direct link when their geographic *Euclidean* distance is less than a connectivity range R_v, where R_v is constant and has the same value for all nodes regardless of their speed (so neglecting *Doppler* shifts). Hence, we assume a homogeneous communication technology. In order to avoid technicalities in the derivations, we later on also utilize squared connectivity areas, *i.e.*, two cars with coordinates (*"*x_1,y_1*"* and *"*x_2,y_2*"*) can communicate if their distance with respect to the *supermum-norm* is less than R_v, *i.e.*, $max\{x_1\text{-}x_2, y_2\text{-}y_1\} \leq R_v$.

In the following, we present the analytical results developed in [Hamouda *et al*. 2009b] showing that under certain assumptions, the vehicles encounter process is a *Poisson* process with rate α and come up with a relation to calculate α. Then, we compare this result with simulations of more general mobility models as reflected by Figure 2.2, investigating first the case of independent movements of vehicles and then taking into account dependencies between movements of different vehicles.

2.3.2 Infinite lane one-dimensional freeway scenario

In this section, we show that under certain scenario assumptions, the encounter process is a *Poisson* process with rate α, where α can be calculated from the car density and the expected absolute value of the relative speed distribution.

Assuming a quadratic connectivity range for the mobility scenario with independent movements in Figure 2.2, the vertical dimension of the scenario is irrelevant and we can simplify the freeway model to a 1-dimensional movement scenario. We now assume that the road is infinitely long and the initial placement of the vehicle is a spatial *Poisson* process. We place the reference vehicle at position *0* at time *0*, and let the other vehicles, say x_i, move with constant velocities v_i drawn from

a continuous distribution V with *p.d.f.* $f(v)$ (*i.e.*, we consider their relative velocity to the reference vehicle). The following theorem then shows that the single-hop encounter process E of meeting new direct neighbours in this case is a *Poisson* process.

Theorem 1: *let X be a Poisson process with intensity $\rho(x)$ for $x \in \Re$, and v_i be independent identically distributed (i.i.d.) continuous random variables with density f. Then the encounter process E is a Poisson process with intensity:*

$$\alpha(t) = \int_0^\infty \rho(R_v - vt) f(v) v dv - \int_{-\infty}^0 \rho(-R_v - vt) f(v) v dv$$

provided α can be locally integrated.

In the case where X is stationary with intensity ρ, the encounter process E is also stationary and has intensity:

$$\alpha = \rho . E|V| \tag{2.1}$$

where V is a random variable with density f provided $E|V| < \infty$.

Proof: Let X_1 and X_2 denote the point processes of those $x_i \in X$ with $v_i > 0$ and $v_i < 0$, respectively. Since v_i are *i.i.d.* and they are independent of X, X_1, and X_2 are independent thinning of X, so they are *Poisson* with intensities $p\rho(x)$ and $(1-p)\rho(x)$, where $p = P(v_i > 0)$ (see e.g., [Møller & Waagepetersen 2004]).

Consider first X_1. Since the vehicle in X_1 has a higher velocity than the vehicle at zero, these will get into connectivity range at position $-R_v$. The time at which this occurs for a vehicle with position x_i at time zero is given by $t_i = -(x_i - R_v)/v_i$; denote the process of these t_i by E_i. Since v_i is distributed with density $f_1(x) = f(x)/p$ on $(0,\infty)$ (*i.e.*, f restricted to the positive half-line), the conditional distribution of t_i given x_i has density:

$$f_1\left(-\frac{x_i - R_v}{v_i}\right) . \left|\frac{x_i - R_v}{v_i^2}\right|$$

By Proposition 3.9 in [Møller & Waagepetersen 2004], E_i is a *Poisson* process with intensity:

$$\alpha_1(t_i) = \int_{R_v} p\rho(x_i) f_1\left(-\frac{x_i - R_v}{t_i}\right) . \left|\frac{x_i - R_v}{t_i^2}\right| dx_i$$

$$\alpha_1(t_i) = \int_0^\infty \rho(R_v - vt_i) f(v) v dv$$

where the transformation $v = -(x_i - R_v)/t_i$ has been used.

Similarly it can be shown that E_2, the process of t_i resulting from $x_i \in X_2$, is a *Poisson* process with intensity:

$$\alpha_2(t_i) = \int_0^\infty \rho(-R_v - vt_i) f(v)(-v) dv$$

Since X_1 and X_2 are independent, E_1 and E_2 are also independent, so their superposition is a *Poisson* process with intensity $\alpha(t) = \alpha_1(t) + \alpha_2(t)$ (see e.g. [Møller & Waagepetersen 2004]), from which the first assertion follows. The stationary case follows directly by letting ρ be constant in the equation for $\alpha(t)$.

2.3.3 Finite lane, 2-dimensional freeway scenario

The rigorously proven theorem in the previous section relies on three key assumptions:

i) Infinite, 1-dimentional stretch of freeway.

ii) Initial *Poisson* placement of vehicles.

iii) Independent movements of vehicles over time at constant speed selected from some given distribution.

In this section, we will check the sensitivity of the *Poisson* results towards slight deviations from assumptions (*i*) and (*ii*) by investigating slightly more complex freeway mobility scenarios as described in Section 2.3.1.

Subsequently, we perform simulation experiments in order to compare with the *Poisson* results of the previous section and to validate to what extent Equation (2.1) provides a useful quantitative approximation even in more complex mobility models. We therefore consider the 2-dimensional spatial scenario described in Figure 2.2. Vehicles move in two different directions with the following example corresponding to uniform speed distribution and uniform initial placement of the vehicles:

- Each vehicle has a constant speed selected at the beginning of the simulation uniformly distributed between v_{min} and v_{max}.
- Each vehicle is assigned x coordinate between $(0, L)$ and y coordinate between $(0, 2W)$ according to a uniform distribution.
- The reference vehicle at the beginning of the simulation is assigned initial coordinates (x_1, y_1) and speed v_1.

Unless specified differently, we use the default values of the parameters given in Table 2.1.

Table 2.1: Default freeway mobility scenario parameters

Parameter	Definition
Freeway length L	5000m
Freeway width W	15m
Vehicle density ρ	1car/100m
Bounds of the initial speed following a uniform distribution	v_{min} = 80km/h, v_{max} =130km/h, average speed of v_i =105km/h
Vehicles connectivity range R_v	300m
Access points connectivity range R_l	400m
Initial placement of the reference vehicle and its initial speed	(x_1, y_1) = (2500m, 22.5m), v_1 = 108km/h

Note that for this set of default parameters, the connectivity range of the reference car covers the full width of the freeway in both driving directions. Also, the expected absolute value of the relative speed in comparison to the reference car results as $Ev_{rel,1}$ = 8.7 km/h.

The geographic movement of the cars is simulated considering fixed time steps of granularity 0.1sec. Time steps at which k-hop connectivity with the reference car are established newly or vanishing are recorded. The results from the simulation model are used to investigate the distribution of the encounter times. In the second part of this section, we analyze the impact of the connectivity range on the encounter process rate both by detailing Equation (2.1) and by simulations. Finally, we analyze the distribution and expected value of the connectivity durations. We first focus on single-hop connectivity (k=1), then we analyse the case of multiple hops.

2.3.4 Marginal distribution of single-hop C2C encounter times

Let us denote by $\{t_1, t_2, t_3 ... t_n\}$, the time intervals observed between consecutive encounters. Figure 2.4 plots the empirical probability distribution function for the C2C encounter time in the single-hop case. This simulation is obtained from samples of 300 simulation runs, each entailing *5hrs* simulated time (approximately 600 encounter samples in each run). Considering a car density ρ =1car/100m and a connectivity range R_v=300m.

Note that with this set of parameters, the connectivity range of the reference vehicle covers the full width of the freeway in both driving directions.

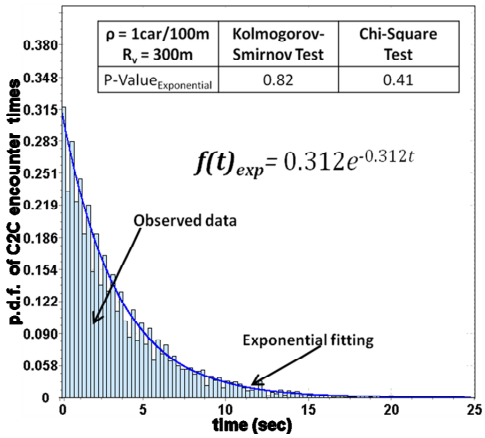

Figure 2.4: Empirical probability density function of the single-hop C2C encounter time: simulation results and comparison to an *exponential* distribution

Statistics for the times between encounters observed from the simulation show a *mean* = 3.34*sec* corresponding to a rate estimate $\hat{\alpha}$ = 0.29/*sec* and a *variance* of the C2C encounter times = 10.38sec^2. The encounter rate value that results from the fitting to an exponential distribution is α=0.312/*sec*, so rather close to the simulation estimate. This result is confirmed by the "*best*" *p-values* associated to the *Kolmogorov-Smirnov* and χ^2 statistical tests shown in Figure 2.4 by using *Least Squares Estimates* (LSE) as parameter estimation method used. The distribution fitting is done with the EasyFit tool[3].

Note that the actual simulated mobility model deviates from the assumptions of the proof in:

i) The road length is finite.

ii) The road width is not zero; however as cars keep their *y*-coordinate during the simulation runs and the connectivity range covers the full road width, there is no expected impact on the results.

iii) The initial placement is done using uniformly distributed coordinates, which however for large number of cars converges to a spatial *Poisson* distribution.

With decreasing node density, the impact of (*i*) and (*iii*) becomes stronger and deviations from the *Poisson* properties of the encounter process become more pronounced in the case of straight highway mobility model.

[3] http://www.mathwave.com/products/easyfit.html

2.3.4.1 Impact of the connectivity range R_v

The encounter rate α as resulting from Equation (2.1) does not show a direct dependence on the connectivity range R_v. However, if R_v is so small that it does not cover the full width of the freeway, there is an indirect dependence, as not all vehicles driving in the opposite direction can come into range, hence both the effective vehicle density ρ connected to the reference vehicle and also the expected value of the relative speed will depend on the connectivity range in that case. In the following, we quantify such dependence for the example case of a uniform speed distribution. The calculation can be performed analogously for any speed distribution.

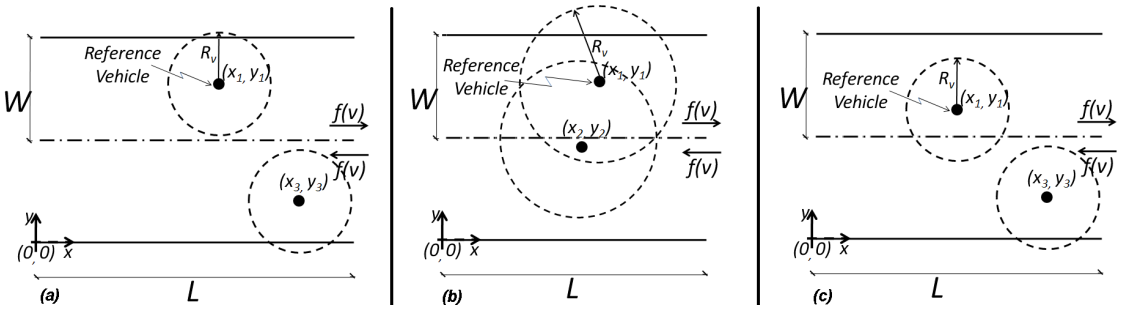

Figure 2.5: Three different cases considered for the reference vehicle

Let us consider the reference vehicle at position (x_1, y_1) shown in Figure 2.5. As mentioned in the beginning of Section 2, each vehicle has a connectivity range R_v. Let us denote by P_1 the fraction of the upper lane (in which the reference vehicle is placed) that is covered by R_v. The upper lane is completely covered ($P_1 = 1$), when y_1 satisfies the condition: $max\ \{2W\text{-}y_1,\ y_1\text{-}W\} \leq R_v$. For other cases ($P_1 < 1$), two situations should be distinguished:

1) If the connectivity range R_v of the reference vehicle does not cross the other lane ($R_v < y_1\text{-}W$), $P_1 = \frac{R_v + (2W - y_1)}{W}$, see Figure 2.5 (***a***).

2) Otherwise, $P_1 = \frac{R_v + (y_1 - W)}{W}$, see Figure 2.5 (***b***), and (***c***).

Similarly, let us define by P_2 the fraction of the lower lane covered by R_v. it can be shown that P_2 is obtained by: $P_2 = \min(1,\ \max(0, \frac{R_v + W - y_1)}{W}))$.

As a consequence, the effective car density which only takes into account cars that can come into the connectivity range of the reference vehicle is:

$$\rho(R_v) = \rho \frac{P_1 + P_2}{2} \tag{2.2}$$

Given that the speed of the cars is distributed according to *f (v)*, we need to go through some technicalities in order to compute the expected relative speed, which we illustrate here for the case of the uniformly distributed speed.

Considering the case where $v_1 < \frac{v_{max}-v_{min}}{2}$, we need to distinguish the three following situations:

1) Vehicles have a larger speed than the reference vehicle and are in the same lane where the reference vehicle is. Let us denote by S_1, the average relative speed corresponding to this case:

$$S_1 = \int_0^{v_{\max}-v_1} v.f_1(v).dv \text{ where, } f_1(v) = \frac{P_1}{P_1+P_2} f(v)$$

2) Vehicles have smaller speed than the reference vehicle and are in the same lane where the reference vehicle is. Let us denote by S_2, the average relative speed corresponding to this case:

$$S_2 = \int_0^{v_1-v_{\min}} v.f_1(v).dv$$

3) Vehicles are in the other lane where the reference vehicle is. Let us denote by S_3, the average relative speed corresponding to this case:

$$S_3 = \int_{v_{\min}+v_1}^{v_{\max}+v_1} v.f_2(v).dv \text{ where, } f_2(v) = \frac{P_2}{P_1+P_2} f(v)$$

Taking into account the $v_1 < \frac{v_{max}-v_{min}}{2}$ and by combining the three averages S_1, S_2, and S_3 together, the average relative speed, denoted by $E|V|$, is obtained as follows:

$$E|V| = \int_0^{v_1-v_{\min}} f_1(v).v.dv + \int_{v_1-v_{\min}}^{v_{\min}+v_1} f_1(v).v.dv + \int_{v_{\min}+v_1}^{v_{\max}-v_1} v.dv + \int_{v_{\max}-v_1}^{v_{\max}+v_1} f_2(v).v.dv$$

Let us consider the case of uniform distributed speeds:

$f(v) = \frac{1}{v_{max}-v_{min}}$, for $v_{min} \le v \le v_{max}$, the average relative speed is given by:

$$E|V| = -\frac{(v_{\max}^2+2v_1(v_{\max}-v_{\min})-v_{\min}^2)P_2+((v_1-v_{\min})^2+(v_1-v_{\max})^2)P_1}{2(v_{\min}-v_{\max})(P_1+P_2)} \quad (2.3)$$

The same procedure and final results can be applied and obtained for the other case where $v_1 \ge \frac{v_{max}-v_{min}}{2}$.

Finally, by applying the relation given in Equation (2.1), we get:

$$\alpha(R_v) = \rho(R_v).E|V| \tag{2.4}$$

For the remaining of this section, we will consider the case where $v_l < \frac{v_{max}-v_{min}}{2}$.

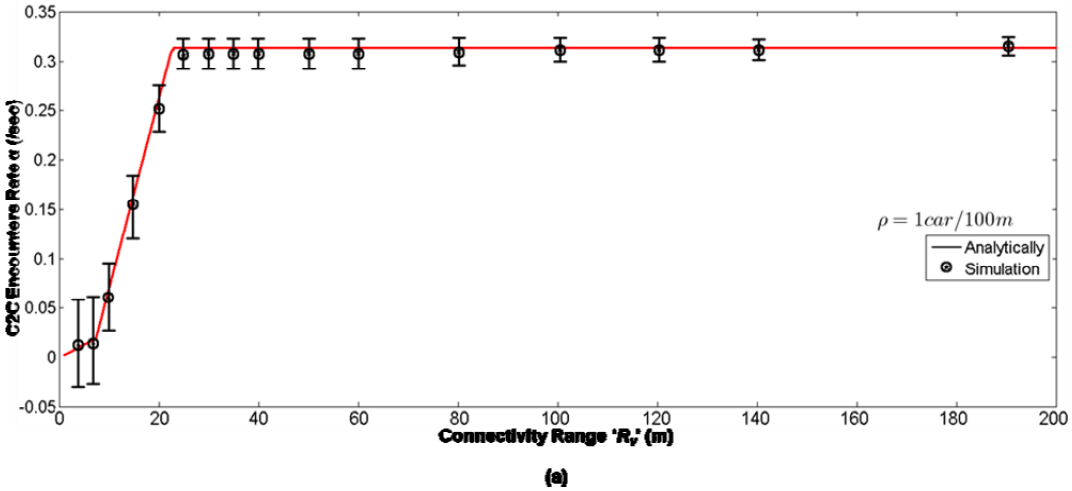

Figure 2.6: Connectivity range impact on the C2C encounters rate α.

Figure 2.6 *(a)* illustrates the impact of the connectivity range on the C2C encounter rate $\alpha(R_v)$ using Equation (2.4) and compares the result obtained analytically to the estimated C2C encounter rate from our simulations. Despite the small deviations of the mobility model from the underlying assumptions of Equation (2.1), calculated values are always within the 95%

confidence intervals of the simulation estimates. We can note that for connectivity ranges with $P_2 = 0$, the C2C encounter rate is increasing slowly then dramatically increases, during the period $P_2 > 0$, until it reaches a stable state when P_1 and P_2 are equal to one. We can notice that the encounter rate stabilises after certain value R_v around *30m*. This corresponds to the case where the two lanes are covered by the connectivity range of the reference vehicle.

The impact of the vehicles density ρ on the encounter rate α for different values of R_v is illustrated in Figure 2.6 *(b)*. We can see that a *10 times* increase of ρ leads to encounter rate α that is *15 times* higher.

2.3.4.2 Multi-hop connectivity dynamics

In this section, we extend our simulation analysis to also count *k*-hop connectivity via relaying nodes ($k>1$). Besides the vehicles that are directly in the connectivity range of the reference vehicle, we count those that can be reached via (*k-1*) relaying nodes. Figure 2.7 plots the Empirical distribution of the time between encountering new (single-hop or 2-hop) neighbours. The simulation estimate of the encounter rate from the simulation data is $\hat{\alpha}$ =0.44/*sec*. In this case also, the *exponential* distribution provides a good fit to the observed Empirical distribution. This is confirmed by the *p-values* computed with the *Kolmogorov-Smirnov* and the *Chi-Square* tests.

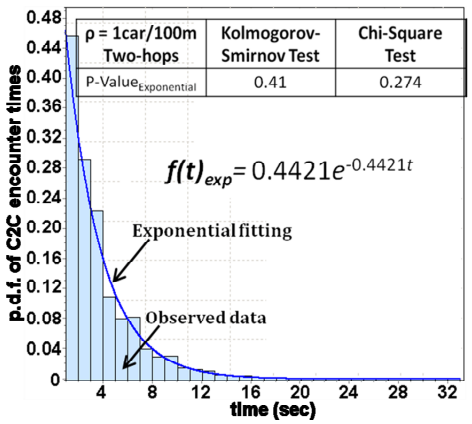

Figure 2.7: Empirical probability density function of the 2-hop C2C encounter time: simulation results and comparison to an *exponential* distribution

The increased encounter rate for the 2-hop case compared to the single hop case results from the fact that: (*a*) obviously, any car can only communicate with the reference car if its distance is below $2R_v$; (*b*) while being in that distance range however, a car can get into and out of connectivity to the reference car multiple times caused by relay nodes gaining or loosing single-hop connectivity. It is noteworthy that, for three and four hops connectivity scenarios, the *exponential* distribution is also a good approximation in the considered cases, which led to estimated rates $\alpha(k=3)=0.485/sec$ and $\alpha(k=4)=0.552/sec$, respectively (see Figure 2.8). With the analogous arguments as for the 2-hop case, the meet rate increases slightly when increasing the number of hops.

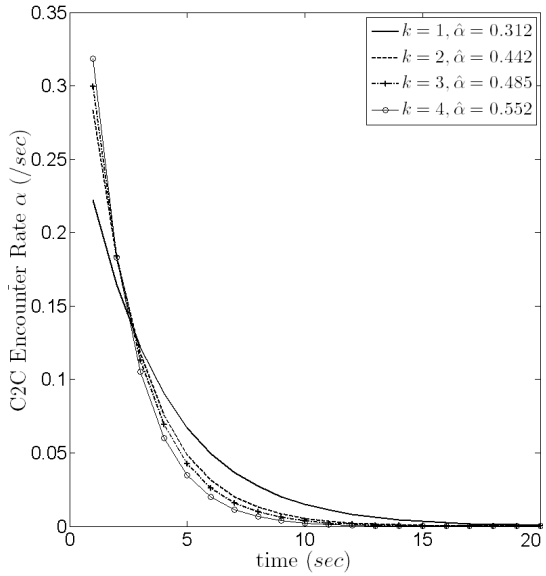

Figure 2.8: *k*-hop C2C encounter time distribution (*k=1,2,3,4*) : *exponential* fitting

2.3.5 Connectivity duration characterization

For many application cases, not only the process of meeting new vehicles in connectivity range is relevant, but also the duration of the connectivity duration. Let us consider again the freeway 2-dimensional mobility scenario described in Figure 2.2. For the single-hop connectivity case with squared connectivity range (*i.e.*, whenever the distance of two vehicles with respect to the supermum-norm is smaller than R_v, the two vehicles can communicate), under the corresponding assumptions, the random variable T_{on} describing the connectivity duration with the vehicle has an expected value given by:

$$\mathrm{E}(T_{on}) \approx 2R_v / |V| \tag{2.5}$$

$|V|$ is the expected absolute value of the relative speed of the vehicle in comparison to the reference vehicle, whose analytical expression is given by Equation (2.3). We remind that $|V|$ also depends on the connectivity range R_v.

From Equation (2.5), we can also derive the disconnection rate ω given by:

$$\omega = [\mathrm{E}(T_{on})]^{-1} \approx |V| / 2R_v \tag{2.6}$$

If we denote by *f(v)* the density function of the absolute value of the relative speeds compared to the reference vehicle, the connectivity duration has a density function *g(t)* given by:

$$g(t) = f_{rel}(2R_v / t)\frac{2R_v}{t^2} \tag{2.7}$$

For circular connectivity range, the calculation of the density becomes slightly more complicated as the distribution of the vertical distance between the vehicles has to be taken into account. On the other hand, the difference in calculations for circular or square connectivity range could be neglected as shown later in this section.

Let us denote by $\{t_{on1}, t_{on2}, t_{on3}... t_{on(n)}\}$, the observed connectivity duration of each encounter with the reference vehicle for each vehicle. Figure 2.9 plots the empirical probability density function obtained from the same simulation experiments as used for the encounter process; the histogram is based on 180000 samples. The figure also plots the probability density function resulting from Equation (2.7) when considering uniform distributed relative speeds. Also shown in the figure is an *exponential* fitting of the Empirical distribution. The estimate of the expected value of the connectivity duration from the simulation data is 1.901*sec*, so the *exponential* fit has rate parameter $\hat{\omega} = 0.53/sec$. We can observe that the results corresponding to the *exponential* distribution fit and those obtained from the application of Equation (2.7) are very close to the simulation results. This is also confirmed by the *Kolmogorov-Smirnov* and the *Chi-Squared* statistical tests.

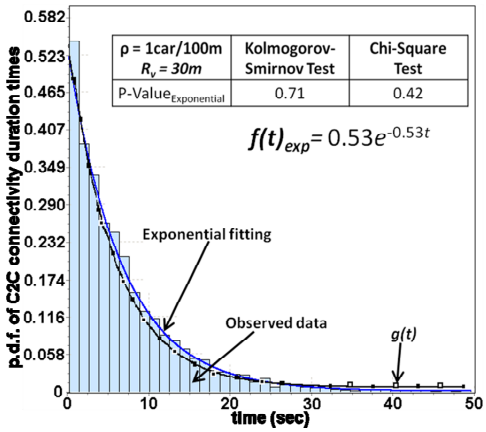

Figure 2.9: Empirical probability density function of connectivity duration: simulation results and comparison to an *exponential* distribution

The results plotted in Figure 2.9 are obtained with a connectivity range R_v = 30m. The impact of the connectivity range on the expected connectivity duration is plotted in Figure 2.10, using simulation results and also analytically based on Equation (2.5).

The connectivity range in the simulation is of circular shape, which shows its impact only for small values of R_v. It can be observed that the average connectivity duration $E(T_{on})$ first increases with the connectivity range R_v till the connectivity range covers the upper lane entirely, and decays subsequently (as vehicles from the lower lane with high relative speed contribute to the connectivity events). Starting from ($R_v = y_1 = 22.5m$), $E(T_{on})$ increases linearly with R_v.

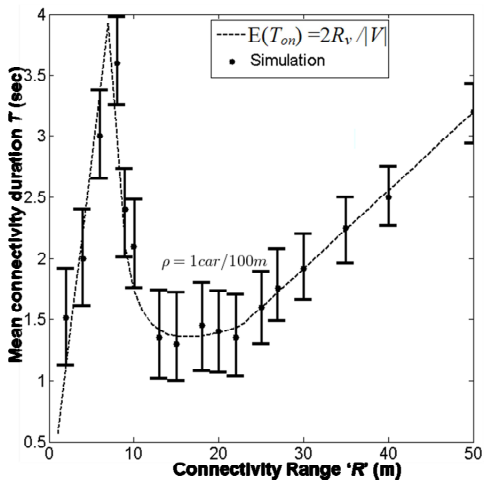

Figure 2.10: The connectivity range impact on the expected connectivity duration

2.3.6 Distribution of C2I encounter times

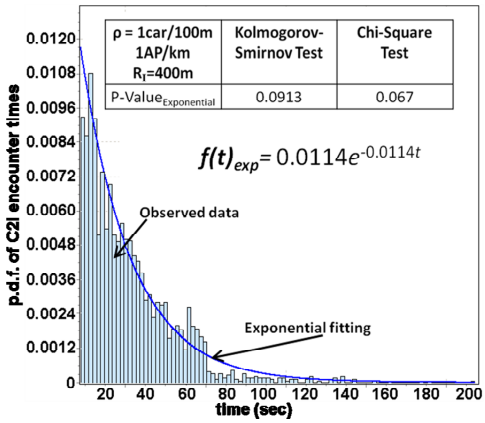

Figure 2.11: Empirical probability density function of C2I encounters times: simulation results and comparison to an *exponential* distribution

Besides the distribution of C2C encounter times, it is also relevant to analyze the distribution of C2I encounter times. Such distribution mainly depends on the density of the access points and the corresponding connectivity range R_I. Figure 2.11 presents simulation results obtained with a density of access points equal to $1/km$ and a connectivity range $R_I=400m$. The other simulation parameters are the same as those used in the previous sections.

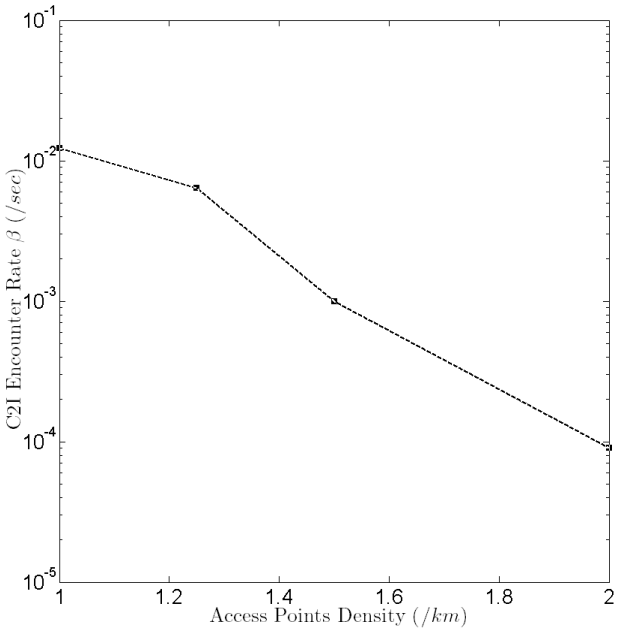

Figure 2.12: C2I encounter rate as a function of the density of access points

Let us denote by β the rate at which the vehicles access to the fixed infrastructure via the access points, and let us denote by $\{t_{off1}, t_{off2}, t_{off3}... t_{off(n)}\}$, the observed times associated with the random variable T_{off} for each vehicle. Statistics from the simulation for the C2I encounter times show a *mean*=81.1*sec* corresponding to a rate estimate $\hat{\beta}$= 0.0123/*sec*. The C2I encounter rate that results from the fitting to an exponential distribution is β=0.011/*sec*. The good quality of fit of the exponential distribution to the simulation results is confirmed by the *p-values* associated to the *Kolmogorov-Smirnov* and χ^2 statistical tests. It is noteworthy that a similar conclusion is obtained when considering a lower density of access points, e.g., *1/2.5km* or *1/5km*. In this case, smaller C2I encounter rates are observed (β=0.0064/*sec* and β=0.00099/*sec*, respectively) as illustrated in Figure 2.12.

2.4 Mobility scenarios with dependencies between vehicles

One of the key assumptions in the simulation scenarios of the previous sections was that the vehicles move independently of each other at fixed speed (though different between vehicles) and they do not change lanes. As we saw in the simulation results, we are in such settings close to the assumptions of the mathematical theorem in Section 2.3.2, and hence a *Poisson* process is a close approximation for the C2C encounter process, even in multi-hop connectivity settings. We could also derive and analyze the distribution of the connectivity duration associated with the C2I encounter process.

In this section, we remove the assumption of independence between vehicles by utilizing publicly available realistic mobility traces. Our main objective is to use realistic mobility traces of vehicular networks that take into account the dependencies between vehicles resulting e.g., from the traffic situations (e.g., slow down in case of congestion), or user behaviours (change lanes,

exit highway, etc.), considering different mobility environments, in particular highway and urban cities.

Several mobility traces are currently publicly available. We can mention for example those generated by the Mobile Node Trace generator (MoNoTrac) [Blywis *et al*. 2009] or by CORSIM [CORSIM], or the mobility traces for buses of public transport systems discussed in [Jetcheva *et al*. 2003].

A thorough examination of these traces led us to select the following traces that seem to us to be the most realistic ones to estimate our connectivity parameters for the highways and urban cities environments respectively: 1) the MMTS traces based on a Multi-agent Microscopic Traffic Simulator developed at ETH Zurich [Traces_MMTS] to analyse the distribution of C2C and C2I connectivity parameters two traces for highway environment and 2) the CRAWDED traces [Kotz & Henderson 2005] based on real life measurements of vehicular mobility scenarios in urban cities using taxi cabs in San Francisco.

In the following, we summarize the main results obtained based on these traces.

2.4.1 Connectivity parameters distribution for highway networks

The Multi-agent Microscopic Traffic Simulator (MMTS) developed at ETH Zurich is capable of simulating public and private traffic over real regional road maps of Switzerland with a high level of realism. MMTS models the behavior of people living in the area, reproducing their movement (using vehicles) within a period of 24 hours. The decision of each individual depends on the area it lives in. The individuals in the simulation are distributed over the roads according to statistical data gathered by a census. Within the 24 hours of simulation, all individuals choose a time to travel and the mean of transportation according to their needs and environment. Travel plans are made based on road congestion; congestion in turn depends on the travel plans.

To estimate our C2C connectivity parameters from the simulation traces, we used a 24 hour detailed vehicle traffic trace file generated by MMTS of granularity 0.1*s*. The file contains detailed simulation of the area in the canton of Zurich; this region includes the part where the main country highways connect to the city of Zurich, the largest city in Switzerland. Around 260000 *vehicles* are involved in the simulation with more than 25 000 000 recorded vehicles direction/speed changes in an area of around 250 km x 260 km. The traces describe a vehicular freeway scenario containing also two driving directions, where each direction has either two or three lanes. The traces provide for each time unit of the simulation, the position (x_i, y_i) of the vehicles.

Trace files are captured with different levels of activity of vehicles in the region and select three time periods that correspond to high density rush hour (more than 50 *cars/km* of road), medium density (30 to 40 *cars/km*), and low vehicle density (less than 15 *cars/ km*), see Table 2.2.

The results presented in this section are for these three different densities and were taken for 5 different areas. From each of these areas, 10 files (each one day) are created. The dimensions of the highways are 2.5*m* per lane and a 2*m* gap between the two sets of lanes. Speed values are in the range of 33.4*km/h* up to 118.8*km/h*.

Table 2.2: MMTS settings for vehicle densities in Highway scenarios

High density	Medium density	Low density
53-90 *vehicles/km*	24-41 *vehicles/km*	11-17 *vehicles/km*

In the following, we characterise the distribution of C2C encounter times and connectivity durations in the case of the highway mobility scenarios using the MMTS traces. The estimation of the connectivity parameters is carried out by assuming that each vehicle has a constant connectivity range of circular shape as considered in Section 2.3.3. The impact of the connectivity range R_v is also analyzed.

2.4.1.1 C2C encounter times (rate α)

Figure 2.13 plots the probability density function of C2C encounter times derived from the simulated data in the *high* car density scenario and the fitted *exponential* distribution using Least Squares Estimates (LSE) for parameter estimation. We have assumed first that the vehicles connectivity range does not cover all lanes (R_v *=10m*). In this case, the statistics for the times between encounters observed from the simulation show a *mean* = 1.12*sec* corresponding to a rate estimate $\hat{\alpha}$ = 0.9/*sec* and a *variance* of encounter times = 1.26sec^2. The encounter rate value that results from the fitting to an *exponential* distribution is α=0.89/*sec*, so rather close to the fitting estimate. This result is confirmed by the *p-values* associated to the *Kolmogorov-Smirnov* and χ^2 statistical tests shown in Figure 2.13.

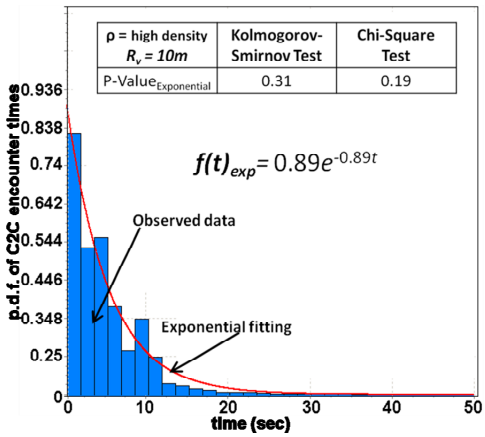

Figure 2.13: Empirical probability density function of the single-hop C2C encounter time: comparison to an exponential distribution – high car density

In addition, with other car densities (*medium* and *low* car densities), the *exponential* distribution is always adequate for modeling the C2C encounter time with rate equal to 0.84/*sec* and 0.81/*sec* for medium and low car densities, respectively. These rates are close because the car densities are close too compared to those used in the simulations discussed in Section 2.3.3. Moreover, it could noticed that the C2C encounter rates are of the same order of magnitude of those observed in

Section 2.3.3 in our simulations of 2-dimensional freeway scenarios with independent movements of cars.

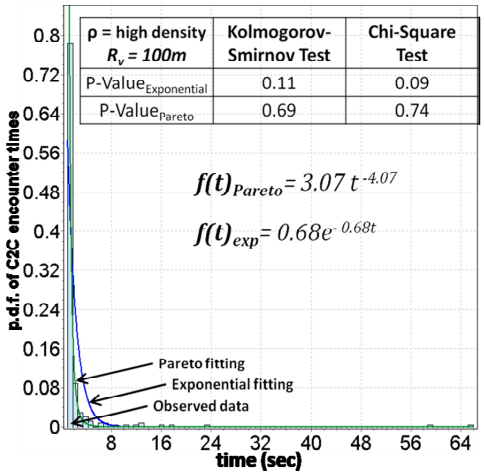

Figure 2.14: Empirical probability density function of the single-hop C2C encounter times and statistical fit to a *Pareto* and an *exponential* distribution – high car density

To analyze the impact of the connectivity range R_v on the results, Figure 2.14 plots the probability density function of C2C encounter times derived from the simulated data in the *high* car density scenario in the case where the connectivity range covers all lanes and both highway directions (R_v=*100m*). Statistics for the times between encounters observed from the simulation show a *mean* = 1.47*sec* corresponding to a rate estimate $\hat{\alpha}$ = 0.68*/sec* and a *variance* of encounter times = 2.2sec^2. As regards the distribution that best fits the simulated C2C encounter times, we provide the results obtained when assuming an *Exponential* distribution and a *Pareto* distribution.

The *probability density function* of the *Pareto* distribution has typically the following expression:

$$f(t) = k_1 . t_m^{\ k_1} . t^{-(k_1+1)} \tag{8}$$

k_1 is the shape parameter and t_m is the scale parameter. When k_1>*1*, the *mean* of the Pareto distribution exists, it is given by: $k_1 . t_m /(k_1-1)$.

As it can be seen visually from the curves plotted in Figure 2.14 and based on the *p-value*s associated to the *Kolmogorov-Smirnov* and χ^2 statistical tests, it appears that the *Pareto* distribution with shape parameter k_1 = 3.07, and scale parameter t_m = 1 (corresponding to a *mean* = 1.482*sec*), provides a better fit in this case than the *Exponential* distribution.

As illustrated in Figure 2.15 (a) and (b), similar results are observed for the *medium* and *low* car densities traffic traces, respectively. For *medium* and *low* vehicular densities, the mean is equal to 2.82 *sec* and 2.67*sec*, respectively. Moreover, in both cases the *exponential* distribution does not provide an acceptable fit to the simulated data.

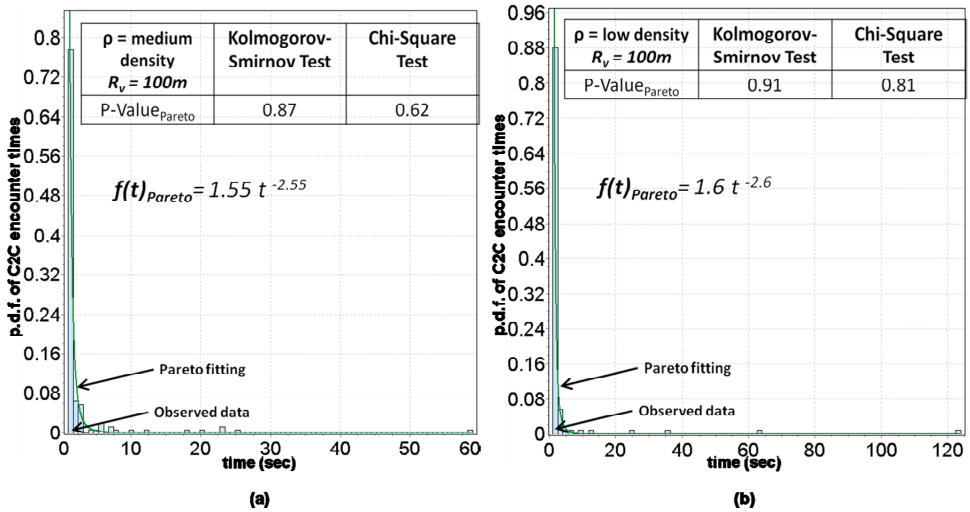

Figure 2.15: Empirical probability density function of the single-hop C2C encounter times and statistical fit to a *Pareto* distribution – (*a*) medium car density, (*b*) low car density

2.4.1.2 C2I encounter times (rate β)

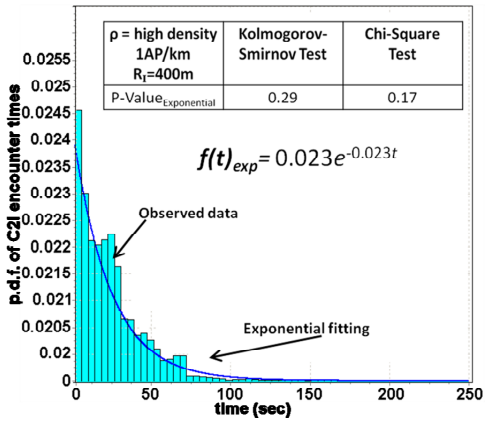

Figure 2.16: Empirical probability density function of C2I encounter times: comparison to an *exponential* distribution – high car density, R_v=100*m*

For the trace files represented above, we assumed a density of access points equal to 1*/km* and an associated connectivity range R_I=400*m*. Also, it is assumed that the vehicles connectivity range is equal to R_v=100*m*. Figure 2.16 presents the results corresponding to these settings which concern the distribution of the C2I encounter times. Statistics for these results show a *mean*=43.5*sec* corresponding to a rate estimate $\hat{\beta}$ = 0.023*/sec*. The C2I encounter rate that results from the fitting to an *exponential* distribution is β=0.0227*/sec*. The good quality of fit of the *exponential* distribution illustrated visually is confirmed by the *p-values* associated to the *Kolmogorov-Smirnov* and χ^2 statistical tests shown in Figure 2.16 It is noteworthy that a similar conclusion –

exponential fitting distribution –is obtained when considering a lower density of access points e.g. *1/2.5km* or *1/5km*, with smaller C2I encounter rate β=0.019*/sec* and β=0.0087*/sec*, respectively.

2.4.1.3 C2C connectivity duration characterization (rate ω)

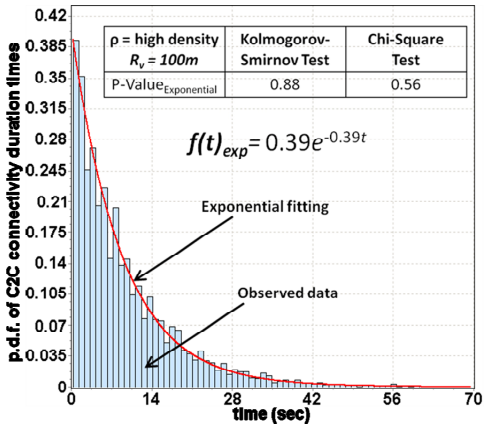

Figure 2.17: Empirical probability density function of connectivity duration: simulation results and comparison to an *exponential* distribution

Figure 2.17 plots the Empirical probability density function obtained from the same aforementioned trace files settings as used for the C2C and C2I encounter process. Also shown in the figure is an *exponential* fitting of the Empirical distribution. The estimate of the expected value of the connectivity duration from the simulation data is 2.6*sec*, so the *exponential* fit has rate parameter $\hat{\omega}$ = 0.391*/sec*.

2.4.2 Connectivity parameters distribution for urban (in-cities) networks

In this section, we characterise our connectivity parameters in the case of the urban mobility scenario. For more realistic vehicles' mobility traces, we use a large GPS-based trace prepared by the CRAWDAD team [Kotz & Henderson 2005] considering mobility traces of taxi cabs in San Francisco, USA[4]. It contains GPS coordinates of approximately more than 500 taxis collected over 30 days in the San Francisco Bay Area. Each taxi is equipped with GPS receiver and sends a *location-update* (timestamp, identifier, geo-coordinates) to a central server. Each timestamp is for a granularity of 0.4*sec*. The San Francisco Bay Area covered by the trace file is about 10*km*x10*km* large. With 500 taxis concerned this corresponds to a vehicle density of about *7vehicles/km*. Compared to the values reported in Table 2.2. for the MMTS trace, this density is lower than the value corresponding to the low density category (*11-17 vehicles/km*).

[4] http://www.cabspotting.org/

As in the highway mobility scenario of Section 2.4.1, the estimation of the connectivity parameters is carried out by assuming that each vehicle has a constant connectivity range of circular shape, and the impact of the connectivity range R_v on the distribution on C2C encounter times and C2C durations is also analyzed.

2.4.2.1 C2C encounter times distribution (rate α)

Figure 2.18 plots the observed and the estimated probability density functions considering C2C encounter times in the case where the vehicle connectivity range R_v is equal to *100m*. Statistics for the observed times between encounters from the traces show a *mean* = 3.33*sec* corresponding to a rate estimate $\hat{\alpha}$ = 0.3*/sec* and a *variance* of encounter times = 11sec^2. Similarly to the highway mobility scenarios based on the MMTS traces with low and medium traffic densities, it appears that the *Pareto* distribution provides the best fit to the observed data. The fitted *Pareto* distribution has the following parameters: k_1=1.32 and t_m =0.8 corresponding to a *mean=3.3 sec*.

It is noteworthy that the observed C2C encounter rates in the urban mobility trace is two times lower than the encounter rate estimated from the simulated highway traces corresponding to the high density scenario. This is related to the low density of the vehicles involved in the urban mobility trace as mentioned in the beginning of this section (*7 vehicles/km* compared to *53-90 vehicles /km* in the highway scenario).

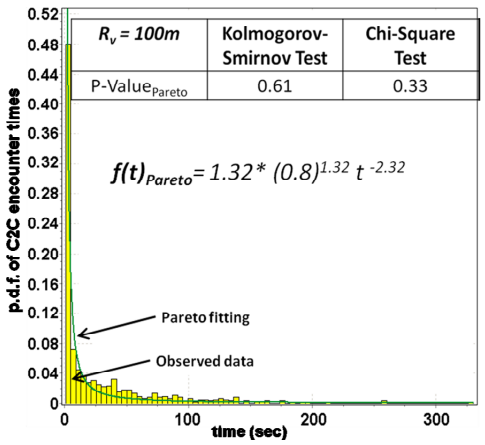

Figure 2.18: Empirical probability density function of the single-hop C2C encounter times and fitting to a *Pareto* distribution

2.4.2.2 C2I encounter times (rate β)

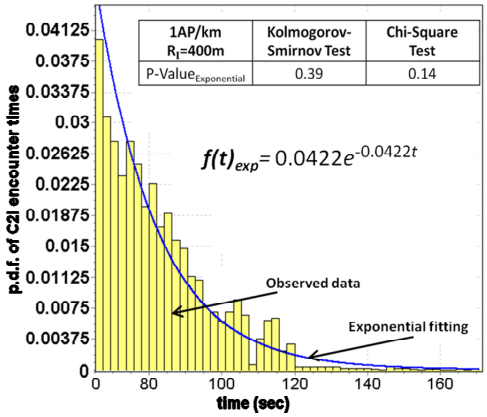

Figure 2.19: Empirical probability density function of C2I encounter times: comparison to an *exponential* distribution – R_v=100*m*

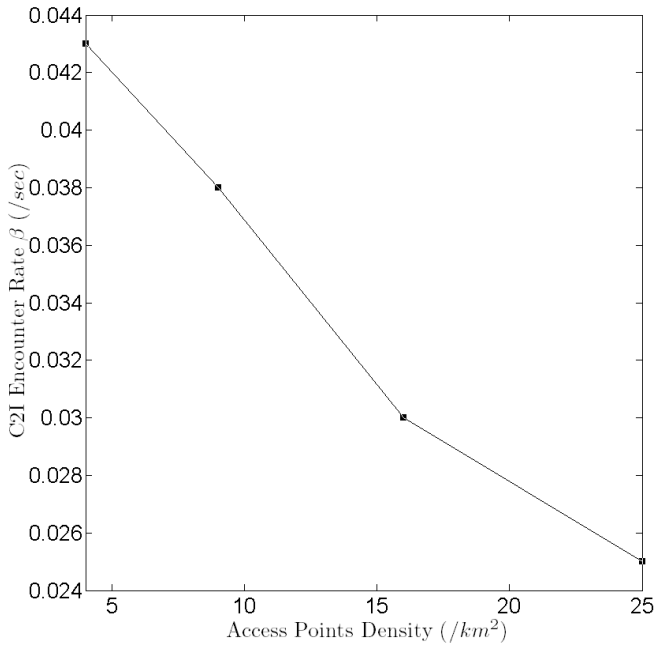

Figure 2.20: C2I access rate as a function of the density of access points

The San Francisco Bay Area covered by the trace file has an area of about 10*km*x10*km*. We placed an access point each about 4km^2 to characterize the C2I encounter process. Figure 2.19 shows the results corresponding to these settings which concern the distribution of the C2I encounter times. Statistics for these results show a *mean*=23.5*sec* corresponding to a rate estimate $\hat{\beta}$= 0.043/*sec*. The C2I encounter rate that results from the fitting to an *exponential* distribution is

β=0.0422/*sec*. This result is confirmed by the *p-values* associated to the *Kolmogorov-Smirnov* and χ^2 statistical tests shown in Figure 2.19. It is noteworthy that a similar conclusion – good fitting of the *exponential* distribution –is obtained when considering a lower density of access points e.g. 1/(9km^2) or 1/(16km^2), with smaller C2I encounter rate β=0.038/*sec* and β=0.03/*sec*, respectively (see Figure 2.20).

2.4.2.3 Connectivity duration characterization (rate ω)

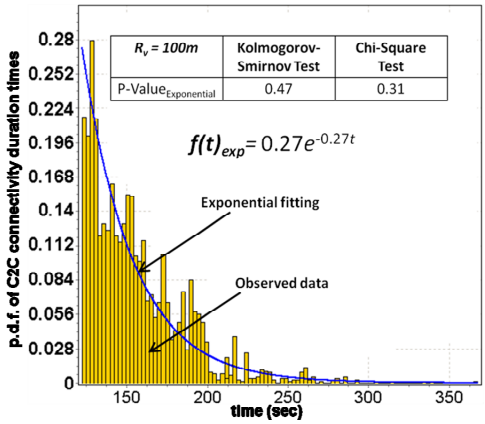

Figure 2.21: Empirical probability density function of connectivity duration derived from the data of the trace and fitting to an *exponential* distribution

Figure 2.21 plots the Empirical probability density function obtained from the same aforementioned trace files settings as used for the C2C and C2I encounter process. Also shown in the figure is an *exponential* fitting of the Empirical distribution. The estimate of the expected value of the connectivity duration from the simulation data is 3.7*sec*, so the *exponential* fit has rate parameter $\hat{\omega}$= 0.27/*sec*. This result confirms once again that the *exponential* distribution is adequate for modeling C2C connectivity duration. Compared to the highway mobility scenario, the estimated rate associated to connectivity duration is almost 1.5 times lower for the urban mobility trace.

2.5 Summary and Conclusion

In this chapter, we analyzed some connectivity characteristics in dynamic vehicular communication scenarios that are important for the design and the performance and dependability assessment of such applications as will be presented in next chapters. Table 2.2 summarizes the different cases investigated and the corresponding analysis technique (analytical, simulation, traces) and the main conclusions derived.

Table 2.3: Connectivity parameters characterization: Summary of analyzed mobility scenarios and conclusions

Mobility scenario, assumptions, analysis technique	C2C encounter times distribution (rate α)	C2I encounter times distribution (rate β)	C2C connectivity duration distribution (rate ω)
Infinite lane highway with independent car movements – analytical proof	• *exponential*	-	-
2-dimensional finite lanes highway with independent car movements - simulation	• *exponential*	• *exponential*	• *exponential*
Highway mobility scenarios with dependent car movements – MMTS traces	• *exponential/Pareto*	• *exponential*	• *exponential*
Urban cities mobility scenarios with dependent car movements – CRAWDAD trace	• *Pareto*	• *exponential*	• *exponential*

Starting from a simple mobility scenario, we have noted that under certain assumptions it can be analytically proven that the C2C encounter time follows an *exponential* distribution. With respect to the C2C encounter time, the rate of the *exponential* distribution can be calculated from the car density and the average relative speed.

Subsequently, simulation experiments with more complex freeway mobility models have been used to check the sensitivity of this result to the underlying assumptions. In particular, we have shown that the *exponential* distribution assumption is also a good approximation in more general freeway scenarios of independent car movements, even in multi-hop connectivity situations. The influence of the connectivity range R_v on the C2C encounter rate was also analyzed. We have observed that when the traffic density is constant, the C2C encounter rate is sensitive to the connectivity range only for low values of R_v (i.e., when all the lanes of the freeway are not entirely covered).

Finally, to analyze the impact of the independent movement assumption of the distribution of the connectivity parameters, we have studied vehicular mobility scenarios with dependent car movements considering two publicly available traces corresponding to two different mobility environments: highways and urban cities. In this case, we have observed that the Pareto distribution generally provides a better approximation for the C2C encounter time distribution and the exponential distribution is no longer acceptable (excepting the case when the vehicle connectivity range is very low).

As regards the distribution of C2I encounter times and C2C durations, we have observed that in all the scenarios that we have investigated for highways and also for urban cities mobility environments, it can be well approximated by an *exponential* law.

These results provide useful insights for the dependability analysis of mobile based applications as will be illustrated in the next chapters. In particular, the dependability modelling formalism to be used should allow us to take into account exponential distributions and also Pareto distribution for describing different mobility scenarios. For this reason, we have decided to use Stochastic Activity Networks for the dependability modelling of the case studies presented in the following chapters (see Chapter 1 Section 1.2.3).

Chapter Three: Consistency and Availability of a Replication Service

"*The only completely consistent people are the dead."*
Aldous Huxley, 1894-1963, British Author

In order to increase the dependability of *stateful* applications, redundancy, provided by replication in cluster architectures, is a well-known and frequently utilized approach. This chapter addresses the modeling and evaluation of data consistency and availability in the context of a replication service middleware for C2C applications in dynamic *ad-hoc* networks. The modeling is based on Stochastic Activity Networks and takes into account the results obtained in Chapter 2 concerning the characterization of connectivity parameters, to analyze the impact of the users' mobility on the probability of data consistency and availability. This chapter builds on our preliminary work, carried out in collaboration with Aalborg University presented in [Matthiesen Moller *et al.* 2008].

3.1 Introduction

As described in Chapter 1, many of the future networking scenarios consist both of wireless *multi-hop* parts and infrastructure based network components. For new application types and future service platforms, server-based applications access is not only offered by the infrastructure network part, but also by the potentially mobile nodes in the *ad-hoc* domain. Dependability requirements for such applications from the user perspective may include several properties such as safety and high availability. The case study investigated in this chapter mainly addresses availability concerns in the *ad-hoc* domain using replication.

Traditional solutions for high-availability rely on redundancy offered by cluster implementations [SAF 2007], in which Middleware services support the timely replication and *fail-over* in case of crash failures of individual cluster nodes [Chen *et al.* 2005]. For stateful applications, such *fail-over* capability typically involves timely replication of application state, which could be implemented by a redundant distributed shared memory [Bozinovski *et al.* 2004].

In mobile *ad-hoc* networks the lifetime of a communication path may be short [Artimy *et al.* 2004]. Communication delays in principle are unbounded due to the unpredictability of the Medium Access procedures on the link-layer and potential re-routing delays in dynamic multi-hop scenarios [Chen *et al.* 2001]. Replication strategies for dynamic data such as application state need to take these communication properties into account [Matthiesen Moller *et al.* 2007]: Larger communication delays can increase the probability of inconsistent replica state, so that dynamic cluster member selection can lead to substantial improvements in mobile scenarios [Olesen *et al.* 2006]. The replication service case study defined in [HIDENETS 2007c] is an example of such a middleware component that provides applications with a resilient shared memory area and performs management of such a dynamic cluster of vehicles. Heuristic algorithms as investigated in [Matthiesen Moller *et al.* 2007], based on measured communication delays and geographic positioning and speed information, can be utilized to trigger membership reconfiguration in a mobile *ad-hoc* network setting.

Various vehicular applications using C2C communications could benefit from the replication service investigated in this case study [Helal *et al.* 1996]. We can consider as an example, an *ad-hoc* network based road-traffic information service, whose information base is dynamically updated. Utilizing the replication middleware service, the application state (traffic information) is kept in a memory area which is shared among the participating servers in the replica set. In case the network topology (connectivity graph) changes significantly, *e.g.*, when a server vehicle (node) exits the highway, the intra-cluster communication with that vehicle will experience increasing delays as the geographic distance and eventually the multi-hop communication path length grows. As the communication with a vehicle deteriorates (*e.g.*, increase of communication delays or of message losses), the probability of observing an inconsistency between the replica states and the application real state will increase.

In [Casimiro Costa *et al.* 2007], a replication service is defined as a middleware service designed to improve dependability of services by service replication and *fail-over* support. The replication service is mainly designed for use within the *ad-hoc* domain. It provides the application with a memory area that is shared between the cooperating vehicles. It has the ability to select vehicles in the network capable of establishing a data sharing cooperation, and the selection is done to provide a stable set of replica vehicles. The replication service is able to handle the failure of service provisioning vehicles. It is aimed at improving the dependability of an application, by creating replicas of the service in dynamically formed groups. These groups are formed and kept consistent by changing group memberships. Group members (vehicles) are selected based on geographical properties and communication properties. The group members are selected in a way that optimizes the data consistency.

In this chapter, we evaluate the dependability of a replicated *stateful* application service based on dynamic cluster of vehicles formation as provided by the Replication Service presented in [HIDENETS 2007c]. Stochastic models are built to analyse and quantify the evolution of the number of participating vehicles in a cluster, the consistency of the states of the replicas composing the cluster, and the application availability taking into account potential failures of the nodes on which the replicas are running. Increasing the number of participating vehicles in a cluster can improve *client-access* to the replicated service in delay-constrained scenarios, as *client-vehicles* can potentially select the server instance with the shortest communication delay to the clients [Hansen *et al.* 2010]. Replica consistency is expressed as the number of vehicles that store the correct "*up-to-date*" data, where correctness here refers to the *real-time* ordering of "*write*" operations to the distributed memory area storing the changing application state.

Using numerical and simulation results from Stochastic Activity Networks models of this dynamic replication scenario, we experiment with different geographic mobility scenarios, and different degrees of dynamicity of the application state. In particular, we consider the case where C2C encounters times are represented by an *exponential* or a *Pareto* distribution depending on the mobility model used (based on Chapter 2 results).

The rest of this chapter is organized as follows. Section 3.2 describes basic background about: *a*) typical scenarios that illustrate the context of the case study and the main concepts behind the design of the replication service investigated in this chapter, and *b*) the dependability challenges and relevant properties to be quantified. The dependability models allowing the assessment of the probabilities characterizing the evolution of the number of replicas in the network and their

consistency are presented in Section 3.3. Numerical and sensitivity analysis results are discussed in Section 3.4. Finally, Section 3.5 summarizes the main conclusions and discusses future work.

3.2 Background

In this section, we present necessary background information needed to understand the scenarios investigated for this case study, both with respect to replication and vehicular *ad-hoc* networks. First, we present a high level view of the replication service design as defined in the context of the HIDENETS project, and then we discuss the problem of data consistency in a distributed shared memory area and also the problems involved with service replication in *ad-hoc* networks.

3.2.1 Design concepts

With respect to the application, the replication service middleware provides a simple interface with functions to initiate a shared memory area and to stop using a previously initiated shared memory area. Furthermore, it has functions to read and write data from and to the memory area. An example of an application using the replication service is a shared whiteboard server. The users of the whiteboard send messages to the whiteboard server containing their changes to the board. To have a consistent whiteboard all changes need to be committed to the board in the shared memory area where it will be replicated to the replica servers in the group.

The replication service middleware consists of four different parts, as shown in Figure 3.1 [Matthiesen Moller *et al.* 2008]. Going clockwise in the figure, the *inconsistency evaluation* part is responsible for estimating data consistency based on communication metrics. The *membership part* is responsible for determining the members of the replica group or cluster. The *state manager* is responsible for sending the state update messages in case of a state change, to store received state variables and to send an acknowledgement when a state update message has been received. The state is the shared memory area provided to the application. The *retrieval part* is responsible for locating the latest and most *up-to-date* state. This part is mainly used in the user nodes to locate a *fail-over* server candidate.

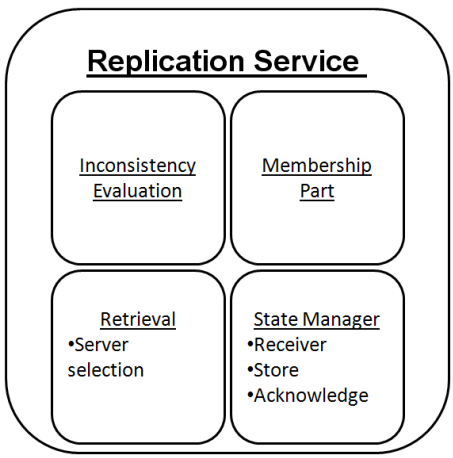

Figure 3.1: Parts of the replication service design

In application scenarios using the replication service, three different nodes and roles can be distinguished: *user nodes*, *relay nodes* and *service nodes*, where a node refers to a vehicle. The user nodes are the clients of the application service which is provided by the service nodes. Although different replica sets can co-exist and overlap, the subsequent discussions will, without loss of generality, focus on a single replicated service instance. In order to start using the replicated application service, the replication service middleware will need to identify the corresponding replica node set, *e.g.* via a dynamic naming service. Upon a failure of the service node or degraded/disconnected communication, the application in the user node can *fail-over* to other service nodes in the replica group. Relay nodes are network nodes that relay packets in the network in a multi-hop communication scenario. In principle, all service nodes within the reach of the multi-hop connectivity are eligible to act as application service replicas; however, in order to limit communication delays between replica nodes, the communication path-lengths may be bounded to a maximum hop-count H. In the extreme case H=1, only direct neighbors (within link-layer connectivity) would be eligible to act as replicas. Any vehicle corresponding to a service replica node can modify the data in the shared memory area (*e.g.*, in reaction to processing requests by users or of the local environment).

The dependability models presented in this chapter and the sensitivity analysis results are intended to quantify and investigate the impact of different traffic situations and dynamic behaviors on the resulting application availability and the probability of state inconsistency of the service replicas.

3.2.2 Dependability challenges

In a C2C *ad-hoc* network, the topology may change so frequently that the amount of signaling needed to keep a group of replicas consistent is potentially large [Killijian *et al.* 2004]. The smaller the number of active replicas, the higher the unavailability of the application service due to crash failures of the server that is undistinguishable from the server leaving the connected network. On the other hand, service replication by broadcasting state information to all vehicles in the network will increase wireless bandwidth consumption and may lead to congestion. The number of replicas to select is hence a tradeoff between *overhead* and *availability*. The design goal of the replication service middleware investigated in the context of the HIDENETS project is to provide services to the user with a high perceived quality of service, while keeping replication and reconfiguration overhead as low as possible.

The properties with the highest impact on user perceived quality of service are *service response time*, *service availability*, and *the correctness of the service* [Avizienis *et al.* 2004]. Correctness is influenced by the consistency of the replicated data. If a user is getting a service which is provided based on *out-of-date* data, the delivered service is not correct. Service availability is directly related to the availability of service nodes that may be impacted by crash failures or other types of failures. The faults that affect the user perceived quality of service are closely related to the properties already described. For example packet loss or excess delay affects the timeliness of the service and also the correctness. A lost update packet is extending the time where the replica server is in an inconsistent state until a successful retransmission or a new update message is received. The same is true for large delay. If a single server in the group is experiencing these types of faults, it can be excluded from the group and a new service node –if any eligible ones reachable in the network –can be included. The replicas are selected in order to achieve stable

clusters as this is preferred to minimize reconfiguration overhead; see [SAF 2007] for strategies to increase stability.

3.3 Dependability modeling

In this section, we present the dependability model that we have developed to evaluate quantitative measures characterizing the consistency of the application state shared by the replica in dynamic clusters, and the availability of the application taking into account crash failures affecting the service nodes. The aim is to understand how different mobility models that match different mobility scenarios can affect the data consistency in the context of this case study. The effect of mobility is studied by considering different distribution laws (*exponential* and *Pareto*) to describe the vehicle-to-vehicle encountering process. These distributions are derived from the results detailed in Chapter 2, based on the statistical processing of real and simulated mobility traces. The results obtained from the models and the sensitivity analysis give preliminary indications and estimations about the expected behaviors and trends that can be observed.

This chapter is based on our previous work related to the dependability evaluation of the replication service middleware reported in [Matthiesen Moller *et al.* 2008]. The dependability model presented in [Matthiesen Moller *et al.* 2008] is based on Markov chains and generalized stochastic Petri nets (GSPNs), assuming a constant rate exponential distribution for describing the encounter process in the *ad-hoc* domain. Here, we discuss these results in the context of the same case study considering more general mobility scenarios. For this purpose, we use stochastic activity networks (SAN). We analyze the dependability provided by the replication service as a function of (i) the various environmental parameters (the rate at which the replica group members join or leave the group, the update rate of the data shared by the group, and network and processing delay of data update messages); and (ii) different distributions for new member join rate (*Pareto* and *exponential* distribution).

In the following, we represent the SAN model and the quantitative measures evaluated from the model to assess the data consistency and application availability. Finally, we discuss the main parameters that are considered in the sensitivity analysis.

3.3.1 Stochastic Activity Network model

Figure 3.2 presents the SAN model describing the evolution of the replication group topology as a result of members joining and leaving the group and potential inconsistency between the local states of the group members, especially during the period when the application state changes and this change is not yet reflected by the group members. In this model, we do not take into account the failures that might affect the different vehicles involved in the scenario. Such failures are considered in section 3.4.2.3 when assessing the service availability.

Let us denote by n the maximum number of eligible replica vehicles in the network and by k, the number of service vehicles that belong to the replica group (also called cluster). Besides these parameters, the model has four input parameters, defined as follows:

1. The first parameter denoted by α (C2C encounter rate) is the rate at which new eligible service vehicles enter the (*multi-hop*) communication range of the existing network. This

parameter depends on the geographic mobility model, the *link-layer* characteristics of the wireless technology (*e.g.*, expressed by a communication radius R_v), and the criterion on '*eligibility*' to act as a cluster node, *e.g.* the maximum *hop-count H* of the multi-hop connection.

2. The second parameter, ω, is the rate with which eligible service vehicles leave the communication range of the replica set, regardless of whether this vehicle is leaving because of a crash failure or another cause; ω is influenced by the same three factors as stated above for α. Note that $1/\omega$ corresponds to the average duration of a C2C connection.

3. The third parameter, γ, is the rate with which the application state changes. This rate depends on the application type and the service usage scenario. Upon the occurrence of an application state change event, all the service replica states become inconsistent and need to be updated.

4. The fourth parameter, δ, is the rate that represents the delay for a replica vehicle to update its state and reflect the application state change.

In the SAN model presented in Figure 3.2 five timed activities are associated to the processes corresponding to the parameters defined above: alpha (when a new vehicle is encountered and joins the network), gamma (when the application state changes leading to the need to update the state of the different replica), delta (describing the time needed to update the state of a replica from an inconsistent state to a consistent state), omega1 and omega2 (describing the processes when a service node holding a replica with a consistent state, respectively with an inconsistent state, leaves the network).

The SAN model includes six places; the two most important ones are labelled IC and CO. These two places denote the number of inconsistent and consistent replica in the group, respectively. The place NS is initially marked with a number of tokens corresponding to the desired network size *n*, *i.e.*, the maximum number of vehicles of the network. The initial marking of the place NR is given by *k*, the desired number of replicas in the group. The marking of this place is decremented when a new vehicle joins the network (increasing the marking of place my_network). The latter saves the number of eligible vehicles that may join the group of service replica. When a new replica joins the group, it has initially an inconsistent state (instantaneous activity t0 is fired and place IC is marked). This state is updated after some delay represented by the timed activity delta, leading to the marking of place CO.

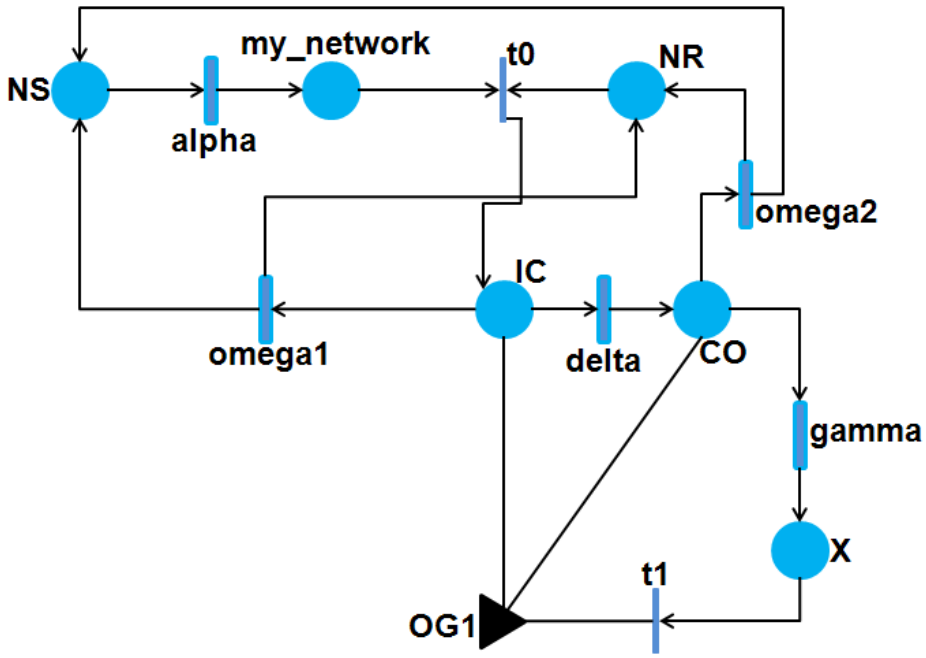

Figure 3.2: SAN model of network size *n* and replication group size *k*

When the state of the application changes, all group replica move to the inconsistent state until they update their local state. This is managed by the place X, the instantaneous activity t1 and the associated output gate OG1. The place X is used to move all the tokens from CO to IC when the gamma transition fires. When place X gets marked, all the tokens in the place CO move to the place IC by activating the instantaneous activity t1. The specification of the function associated to the output gate OG1 is given in Figure 3.3.

```
IC->Mark() = IC->Mark()+ CO->Mark() + 1;
CO->Mark() = 0;
X->Mark() = 0;
```

Figure 3.3: Specification of the function associated to the output gate OG1

The activities t0 and t1 are instantaneous whereas all the other activities are timed and are *exponentially*/*Pareto* distributed. It is important to note that the firing rates of the activities delta and omega1 are weighted by the marking *m*(IC) of the input places IC and the firing of omega2 is weighted by the marking *m*(CO) of the input place CO. The timed activities definitions of the SAN model are summarized in Table 3.1.

Table 3.1: SAN model timed activities distribution

Transition	Distribution	Weight	Description
alpha	α (*exponential/Pareto*)	-	The C2C encountering process with rate α
omega1	ω (*exponential*)	*m*(IC)	The rate at which a replica leaves the group weighted by the marking of place IC
omega2	ω (*exponential*)	*m*(CO)	The rate at which a replica leaves the group weighted by the marking of place CO
gamma	γ (*exponential*)	-	Application state change rate
delta	δ (*exponential*)	*m*(IC)	Replica state update rate weighted by the marking of place IC

The evaluation of the replica consistency can be done through the computation of the probability that place CO is marked. The analysis of the evolution of such a probability for various parameter value ranges, and considering *exponential* and *Pareto* distributed encountering processes is carried out in Section 3.4.

3.3.2 State graph generated from the SAN model

The state graph derived from the SAN model depicted in Figure 3.2 when all the timed activities are *exponentially* distributed corresponds to a Markov chain that can be processed analytically using traditional Markov chain processing techniques [Kemeny & Snell 1960]. When the C2C encounter process is described by a *Pareto* distribution, the corresponding state graph can be processed using the Discrete-event simulation techniques included in the Möbius tool. In the following, we describe the state graphs generated from the SAN model considering the case $n=k$ and $n>k$, respectively.

3.3.2.1 State graph when $n=k$

The Markov chain describing the evolution of replica group members and their consistency when $n=k$ is given in Figure 3.4. Each state is labelled by the triplet (*s, i, m*) where *s* denotes the current size of the network, *i* the number of inconsistent replicas and *m* the number of free places in the group, *i.e.*, the number of network members that can still be accepted to reach the group size *k*.

The encountering of new vehicles is reflected by an increasing number of replica group members when moving to the right within the columns. The number of *consistent* group members is depicted in the rows, increasing from top to bottom. Note that for illustration purposes, a 3-dimensional state-space labelling set is actually fully determined by the first two components.

The state in the lower right hand side is the optimal state of the replica group, as the maximum number *n* of consistent replica nodes is present. Fully consistent replica groups (of not necessarily maximal size) are represented by the whole lower diagonal.

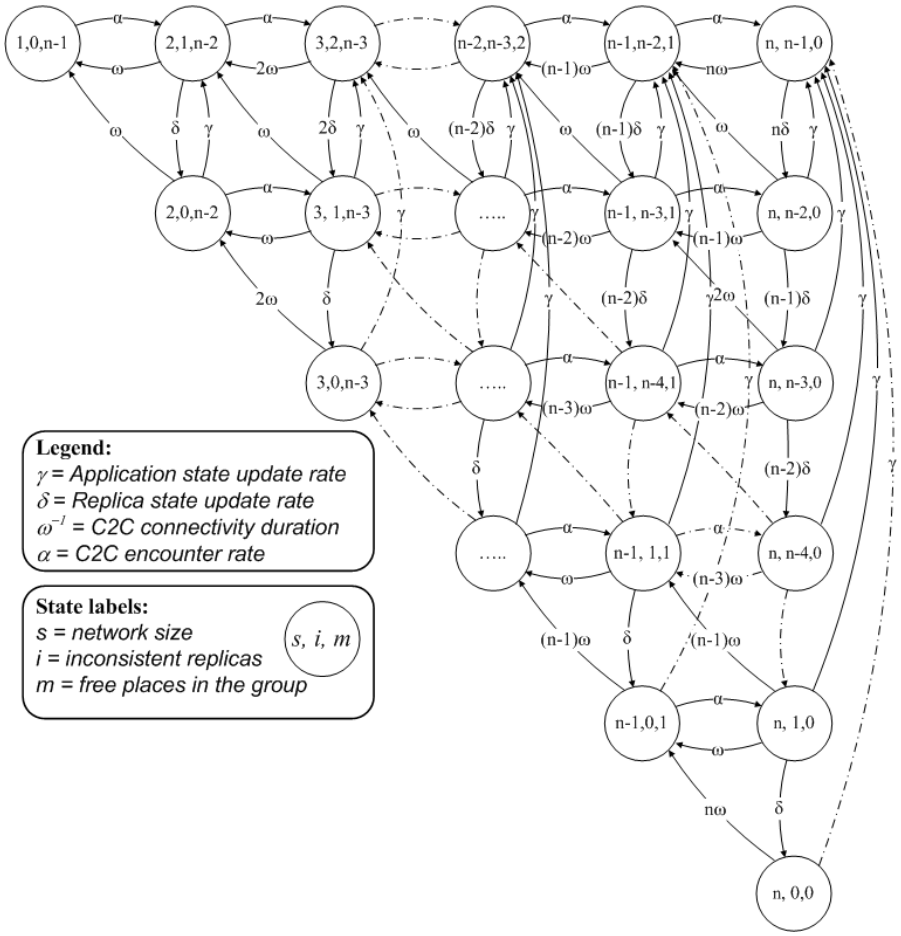

Figure 3.4: Markov model of the SAN model when *n=k*

In Figure 3.4, we distinguish three main sets of states:

- The states labelled "*i=0*" denote those where all vehicles of a group at certain time are consistent.
- The states labelled "*s=n, m=0*" represented in the last column denote those where the system has no more free places and there are "*s-i*" vehicles which are consistent.
- The state labelled "*n, 0, 0*" denotes that the system is saturated with the maximum number of vehicles and all are consistent.

It is noteworthy that the SAN model of Figure 3.2 is generic and can be used to automatically generate the Markov chain associated with any network size *n*. As an example, Figure 3.5 shows the Markov chain corresponding to the case *n=k*=3.

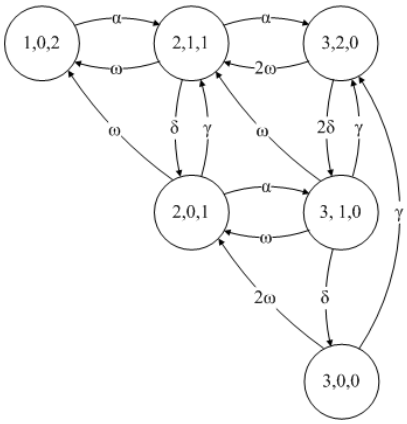

Figure 3.5: Markov model of the SAN model for *n*=*k*=3

3.3.2.2 State graph when *n*>*k*

Figure 3.6 shows an example with *n*=6 and *k*=3.

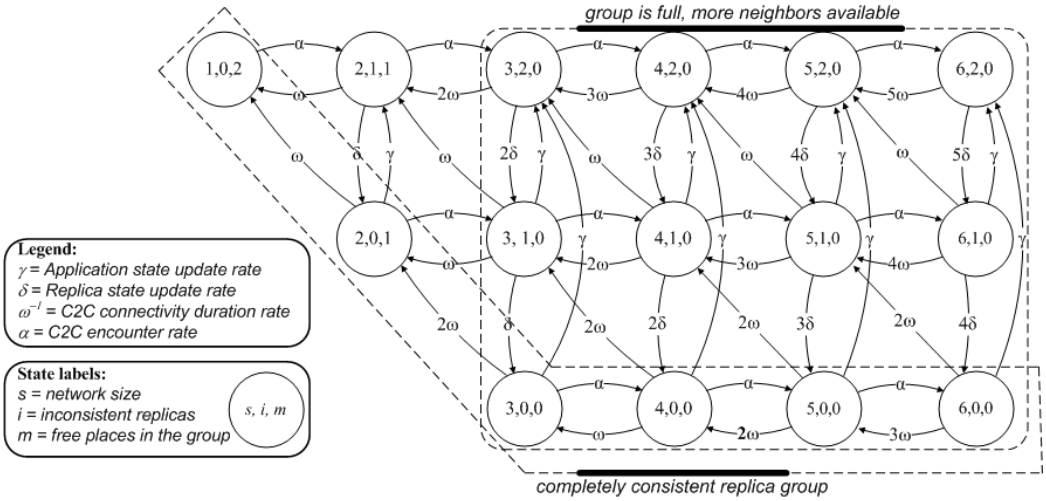

Figure 3.6: Markov model with network size *n*=6 and replica group *k*=3

As in the previous case, each state is labelled by the triplet (*s, i, m*) where *s* denotes the current size of the network, *i* the number of inconsistent replicas and *m* the number of free places in the group, *i.e.*, the number of network members that can still be accepted to reach the group size *k.*

Two state subsets are highlighted in the figure. The first one denoted as "the group being full, more neighbours are available" reflects the dynamics of a fixed maximum size group, being maintained in a variable size connected vehicular *ad-hoc* network. In this subset of the model, vehicles from the known neighbourhood of the replica group are used as replacements when a

vehicle leaves the group. Note that the maximum number of reachable eligible vehicles in the network when the group is full is given by (*n*-*k*).

The second subsets of states highlighted in the figure correspond to the case where all the group replica are consistent. The assessment of the replica consistency can be made by quantifying the probability for the system to be in such states.

3.3.3 Model processing

As already mentioned in the beginning of Section 3.3.2, traditional Markov chain analytical processing techniques can be used when all times activities are exponentially distributed. When the C2C encounter process is described by a *Pareto* distribution, the SAN model can be processed using the Discrete-event simulation techniques included in the Möbius tool.

Three main quantitative measures are of interest in the context of this case study:

1) The probability that all the group replicas' states are fully consistent, denoted as *P(cons)*.

2) The probability that the application is available to the users, *i.e.*, at least one group replica is available and can be accessed by the users, taking into account crash failures that may affect the vehicles on which the service replicas are running. This measure corresponds to the *application availability* and is denoted as A_{app}.

3) The probability that the users can access at least one group replica and the corresponding replica has a consistent state. This measure denoted as *service availability* $A_{service}$, combines both service correctness through the probability of having fully consistent replicas and application availability.

The first measure corresponds to the probability that the system is in one of the states where the place IC of the Figure 3.2 SAN model is not marked (*m*(IC)=0). Concerning the second measure, we consider the simple case where the service is assumed to be unavailable when none of the replicas in the current group is available due to crash failures. It is assumed that the crash failures affecting the vehicles on which the replica are executed fail independently of each other and that the availability of vehicle is known and given by A_S. More complex models for the availability assessment will be presented in the Chapters 4 and 5.

In the following, we present some analytical formula characterizing the above two measures when all the timed activities are exponentially distributed. Numerical and sensitivity analysis results corresponding to the exponentially and *Pareto* distributed C2C encounters are discussed in Section 3.4.

3.3.3.1 Replica consistency

Let us consider the Markov chain generated from the SAN model of Figure 3.2 for the case $n \geq k$.

Considering steady state measures, the probability *P(cons)* that all existing group replicas are fully consistent (*m*(IC)=0) can be obtained analytically [Kemeny & Snell 1960]. The corresponding expression is given by Equation (1).

$$P(\mathbf{cons}) = \frac{1}{\sum_{j=0}^{n}\left(\frac{\alpha^{j}\omega^{(n-j)}}{j!}\right)} \sum_{i=1}^{n} \frac{(\alpha\delta)^{(n-i+1)}\omega^{(i-1)}}{\prod_{k=1}^{n-i+1}(k(\omega+\delta)+\gamma)} \tag{1}$$

3.3.3.2 Application availability

To evaluate the application availability under the assumptions discussed in the beginning of this section, one needs to take into account the dynamic evolution of the number of replicas in the group (*i.e.*, the size of the replica set) and the availability of each vehicle running a replica denoted as A_S.

The evolution of the replica set size is actually fully determined by the rates of arrivals and departures from the group (α and ω respectively). As this is a birth-death process, the probability of a certain replica size is equivalent to the queue-length probability of an M/M/n/n queuing system, which can be found in any standard queuing theory book [Bolch *et al.* 2006]. Let us denote by $P(i, \omega, \alpha)$ the probability that the size of the replica set is *i*. The analytical expression of $P(i, \omega, \alpha)$ is given by Equation (2).

$$P(i,\omega,\alpha) = \frac{\alpha^{i}}{i!\omega^{i}\sum_{j=0}^{i}\frac{1}{j!}\left(\frac{\alpha}{\omega}\right)^{j}} = \frac{1}{\sum_{j=0}^{i} j!C_{j}^{i}\left(\frac{\omega}{\alpha}\right)^{j}} \tag{2}$$

Hereby, $C_j^i = \frac{i!}{j!.(i-j)!}$ denotes the binomial coefficients. Note that although Equation (2) can be derived directly from a simple M/M/n/n model, this is not the case for the more complex 2-dimensional Markov chain structure that leads to the derivation of Equation (1).

In Equation (3), the application availability is calculated given that a specific user perceives a specific server with the availability A_S.

$$A_{app} = \sum_{i=1}^{k} P(i,\omega,\alpha).(1-(1-A_S)^{i}) \tag{3}$$

This equation states that the application is available as long as there is at least one group replica available in the group.

3.3.3.3 Service availability

The service availability as perceived by the users, $A_{service}$, combines the application availability and the probability that the local replica invoked by the user has a local state that is consistent with the application state. This measure can be computed by equation (5).

$$A_{service} = \sum_{j=1}^{k} P(m(CO) = j).(1-(1-A_s)^j \quad (4)$$

In Equation (5), $P(m(C0)=j)$ is the probability associated to the states of the SAN model of Figure 3.2 such that the place CO marking holds j tokens ($1 \leq j \leq k$), *i.e.*, the replica set is composed of j consistent replicas. This probability is different from *P(cons)* which corresponds to the case where all the replicas are fully consistent.

3.4 Results

This section presents some numerical results for sensitivity analysis for the probabilities and dependability measures derived in the previous section. We first discuss the values assigned to the parameters of the models and then present some results illustrating the behavior of the probabilities characterizing the inconsistency among group replica, the probability of reaching the maximum group size, the application availability, and finally the service availability. Also, the sensitivity results take into account the impact on the results of the C2C encounters distribution (*exponential* vs. *Pareto*).

3.4.1 Parameter values

Naturally there are constraints on the rates used in the model if the resulting numbers should show results of realistic scenarios. For instance the rate δ at which a replica updates its local state to reflect the latest change of the application state can be estimated by considering the inverse of the average end-to-end communication time between two replicas in the group. There is a theoretical minimum to this delay which is the time it takes to transfer one update message to a link-layer neighbor. Assuming that one update message can be sent within 0.1 *seconds* under good conditions (one hop, with MAC delay) the corresponding update rate will be around $\delta = 10/sec$. To analyze the impact of this parameter on the results, two different values of δ are considered: 10 and 1 per *second*. The order of magnitude of the application state change rate γ depends on the considered application. Two different values of this parameter have been considered for the results presented in this section: 0.1 and 1 per *sec*, corresponding to an average time between two consecutive changes of 10*sec* and 1*sec* respectively.

As regards, the rate α of encountering new vehicles that are able to join the group, we have also considered different values corresponding to different traffic situations, based on the analyses carried out in Chapter 2. The minimum rate is zero meaning not meeting any vehicles at all on the road. An example of value for this rate is 0.3/*sec* which is approximately equivalent to vehicles in a highway with a density ρ = 1*vehicle*/100*m*, with an average speed of 105*km*/*h*. This value could

be much higher when considering multi-hop communication (see Chapter 2). Various values for ω, the rate at which a vehicle leaves the network can also be considered to reflect different behaviors of the participating nodes.

Rather than focusing on the absolute values of α and ω, one can analyze the results by considering the relative ratio α/ω. Higher values of this ratio correspond to environments where the probability of meeting a new vehicle is higher than the probability that a participating vehicle will leave the network. In a freeway scenario where cars join and leave the road via on and off-ramps and cars travel with a mean speed of 105*km*/*h* excluding trucks the rate of vehicles joining the group is quite low because the cars travel with approximately the same speed. Assuming 18 vehicles joining the group per minute gives a join rate of 0.3/*sec*. With an assumption that each vehicle stays in the group on average for 2 to 4 *seconds*, the actual leave rate is about ω =0.6/*sec* (see chapter 2 for more details); accordingly, the ratio α/ω will be around 0.5. Higher values can be obtained for other scenarios.

To summarize, the settings of the parameters α, ω, δ, and γ that represent the range of values considered in our sensitivity analyses is as presented in Table 3.2.

Table 3.2: The range values (/*sec*) for each parameter

Parameter	α	ω	γ	δ
Range	[0.05; 10]	[0.1; 10]	[0.1; 1]	[1; 10]

3.4.2 Sensitivity analyses

This section presents some sensitivity analyses results illustrating the dynamic evolution of the replica group size, the probability of having fully consistent replica and the service availability.

3.4.2.1 Evolution of the replica group size

Figure 3.7 plots the evolution of the probability $P(k, \alpha, \omega)$ to reach the maximum group size ($n=k$) when the C2C encounters are *exponentially* distributed and *Pareto* distributed. As expected, it is shown that the probability to reach the maximum group size decreases as n increases. The probability plotted in Figure 3.7 is only affected by the values of α and ω that characterize the dynamic evolution of the network topology: the higher the ratio α/ω, the closer the probability of reaching the maximum group size gets the value 1. Also, it can be seen that this probability is very sensitive to the value of α/ω when this ratio is relatively low (less than 10^2 in the example setting).

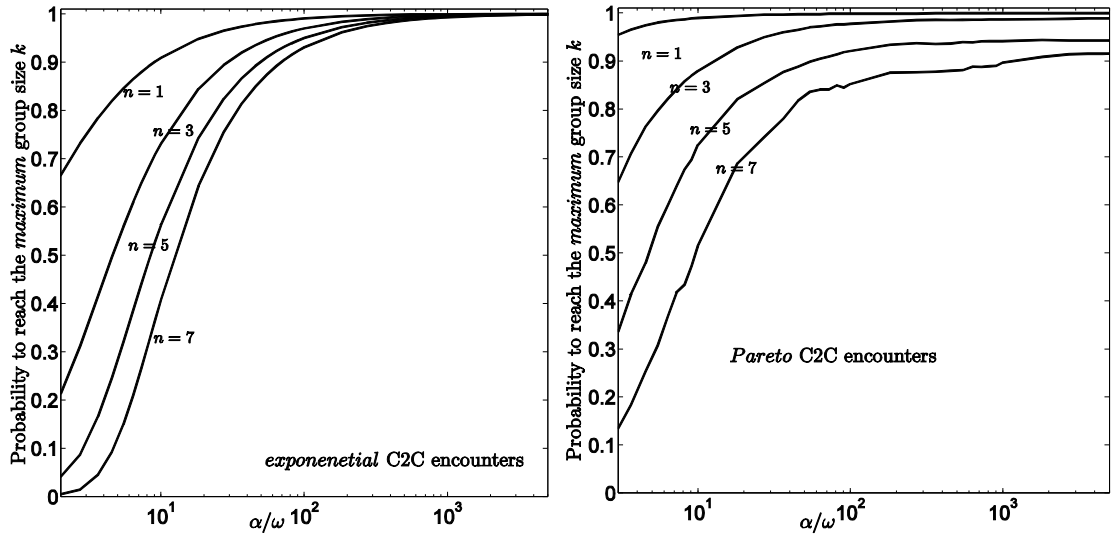

Figure 3.7: Probability *P(k, α, ω)* of reaching the maximum group size *n=k*

The comparative analysis of the encounters distribution impact on the probability $P(k, \alpha, \omega)$ is better reflected in Figure 3.8 that plots the results for *exponential* and *Pareto* C2C encounters when $n=k=3$. Let us remind that based on Chapter 2 results, these distributions are representative of two different mobility scenarios corresponding to freeways (*exponential*) and urban cities (*Pareto*). We can observe two different trends depending on the value of the ratio α/ω: when this ratio is lower than 10^2, the maximum group size can be reached with a higher probability than the *Pareto* distribution. The trend is inverted in the environments for which the ratio α/ω is higher, but the difference is low in this case.

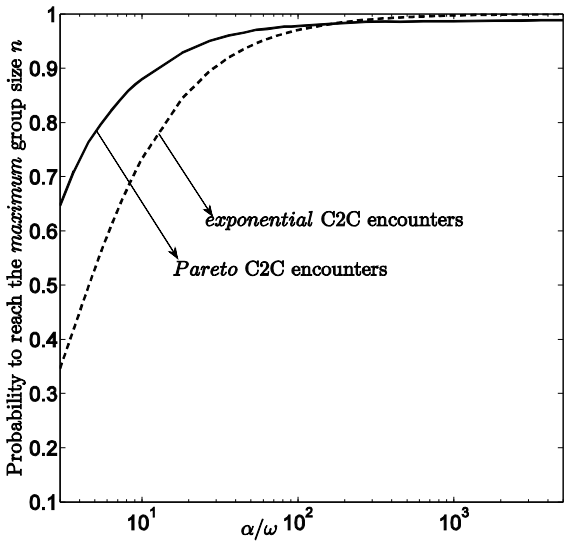

Figure 3.8: Probability of presence of maximum group size *n=k=3*

3.4.2.2 Replica Consistency

Figure 3.9 shows that the probability of a fully consistent replica set converges rather fast to a limit value for increasing α/ω. This limit value depends strongly on the application state change, γ, and the rate δ for a replica to update its state and reflect the application state change. This is intuitive too, since timely updates are needed to achieve consistent replica groups.

Figure 3.9 shows the combined impact of the parameters γ and δ on the probability *P(cons)*. It can be seen that a decrease of δ of one order of magnitude (from 10 to 1), could lead to a degradation of the consistency probability in the order of 4 to 9 times, depending on the considered values for the ratio α/ω and the value of γ. Also, for α/ω values higher than 10, the higher the ratio δ/γ, the higher the consistency probability for both mobility models corresponding to *exponential* and *Pareto* C2C encounters. However, the comparative analysis of the results corresponding to the *exponential* and *Pareto* distributed C2C encounters shows a significant impact of the mobility model. Let us consider for example the case γ = 1/sec: the probability *P(cons)* when δ equals to 1/*sec* is 3.6 times higher for *Pareto* encounters compared to the *exponential* case. Also, increasing the replica state update rate δ from 1/*sec* to 10/*sec* leads to a 3.56 times improvement of *P*(*cons*) for *exponential* encounters compared to 1.3 times for *Pareto* encounters.

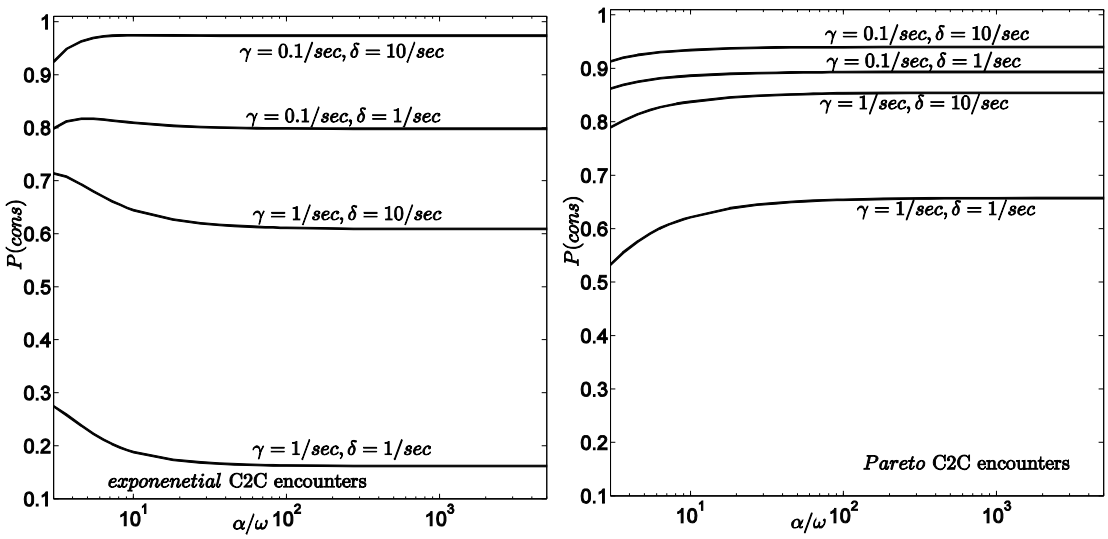

Figure 3.9: Probability of having a fully consistent group for *n*=*k*=6

By setting the application state change rate γ=1/*sec* and the replica state update rate δ=10/*sec*, Figure 3.10 shows the probability of having a fully consistent group for $n=k$ and $n \geq k$, considering the case of *exponential* C2C encounters. We can notice that this probability is very sensitive to the number of replica *k*, more than to the network size *n*, *e.g.*, the $P(cons)_{n=8,k=8}$ is approximately equal to the $P(cons)_{n=6,k=6}$.

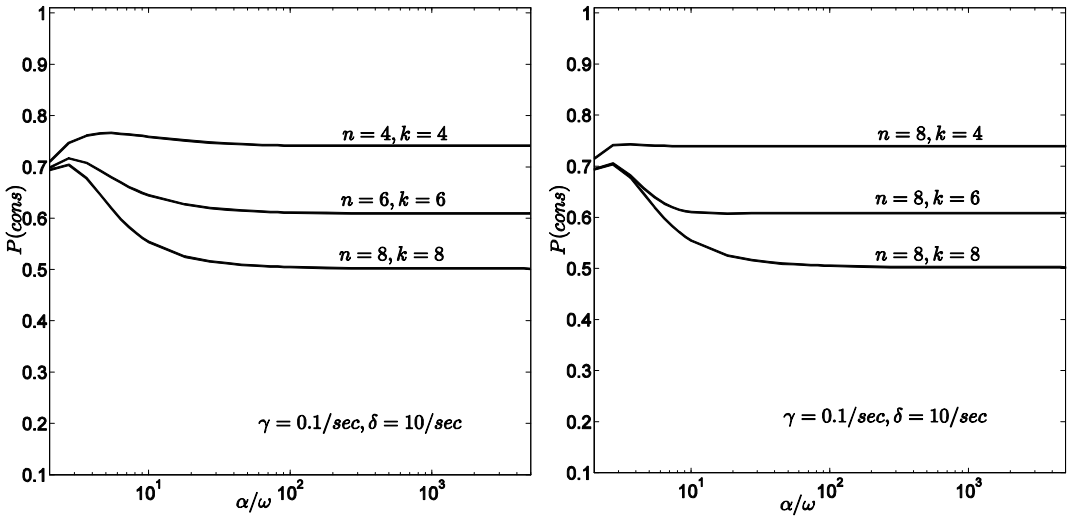

Figure 3.10: Probability of having a fully consistent group: *exponential* C2C encounters

Considering the behavior of the consistency probability as a function of the replica group size, Figure 3.11 shows that for traffic situations corresponding to small values of the ratio α/ω (*e.g.*, in highway systems) the probability of achieving a full replica group is reduced significantly when the desired group size grows. For the desired group size k=8 in a network with n=8 vehicles, the ratio α/ω must be larger than 3 to get a probability higher than 80% of achieving a fully consistent replica group. Such a probability increases when considering smaller sets of replicas (*e.g.*, k=6).

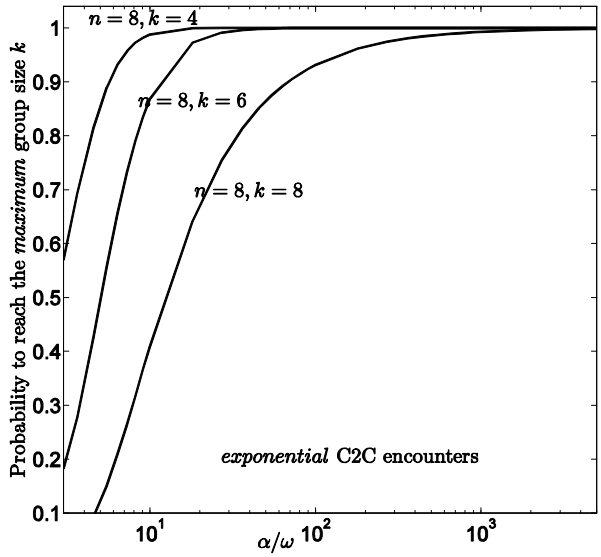

Figure 3.11: Probability of reaching different maximum group size *k* with *n*=8.

More generally, it can be noticed that the probability *P(cons)* remains high (about 80%) even when the ratio γ/δ is as low as 10. Higher replica consistency levels should be observed when the application state change rate γ is considerably lower than the rate at which the group replicas update their state to reflect such changes. Typically, higher replicas update rates will be observed when the *ad-hoc* network interconnecting the group replica is fast and not congested.

3.4.2.3 Application and service availability

This section presents the sensitivity analyses related to the assessment of application and service availability as defined in Section 3.3.3, considering firstly the case of mobility scenarios with *exponentially* distributed C2C encounters. The analysis of the C2C encounters distribution is (*exponential* vs. *pareto*) is discussed at the end of this section.

The application and service availability measures A_{app} and $A_{service}$ are computed based on Equations (3) and (5), assuming that each vehicle holding a group replica has an availability A_s=0.9.

Figure 3.12 plots A_{app} and $A_{service}$ for different group replica sizes. In order to illustrate the necessary tradeoffs between the application availability and the service correctness measured through replica consistency, this figure also plots the evolution of *P(cons)*, the probability that the group replicas are fully consistent.

Two different settings of α, ω, γ, and δ are considered. In the first setting, the parameters are as follows: α=1/*sec*, ω=0.1/*sec*, γ=0.1/*sec*, and δ=1/*sec*. In the second setting, the C2C encounter rate is one order of magnitude lower (α=0.1/*sec*), with the same values for the other parameters.

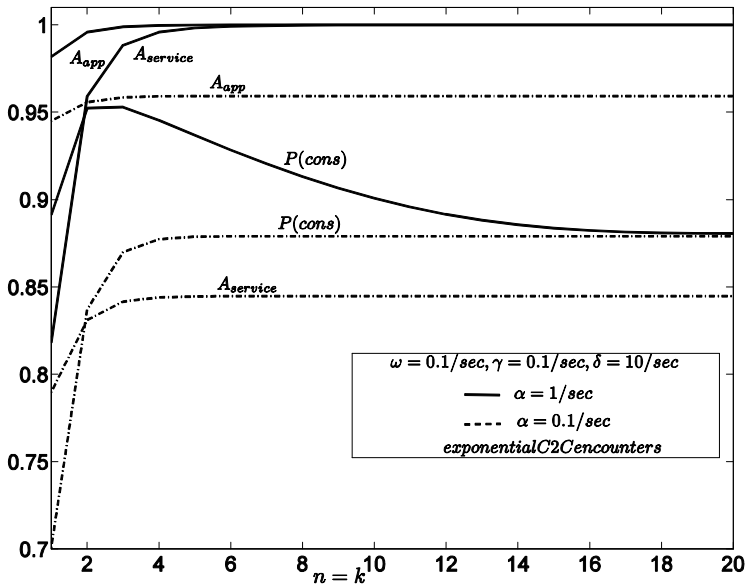

Figure 3.12: Service and Application availability vs. group consistency for different group sizes

Let us remind that the availability of the service as perceived by the users, defined by $A_{service}$, results from the combined effect of the application availability A_{app} and the probability that the selected replica is consistent (see Equation 4). Figure 3.12 shows two different trends for the two considered C2C encounter rates. For mobility scenarios characterized by a low C2C encounter rate α=0.1/*sec*, parameter settings, the service availability is of the same order of magnitude of the application availability. Also, it can be observed that the service availability reaches a stable behaviour very quickly for small values of *n*, around 4. As expected, higher C2C encounter rates lead to higher values for the service availability.

Indeed, the analysis of the trends illustrated by Figure 3.12 and also by Figure 3.13 showing the evolution of $A_{service}$, A_{app} and *P(cons)* can help the designers to make necessary tradeoffs and select the most appropriate network and group size for the corresponding mobility scenarios. For the parameters numerical values considered in Figure 3.12, the group size providing the best consistency while providing a high level of service availability is around 3 or 4.

The results discussed above correspond to the case when the C2C encounters are exponentially distributed. As illustrated in Figure 3.13 similar trends are also observed for *Pareto* distributed C2C encounters. However, for the parameters considered in this figure, lower service availability levels are observed in the *Pareto* case.

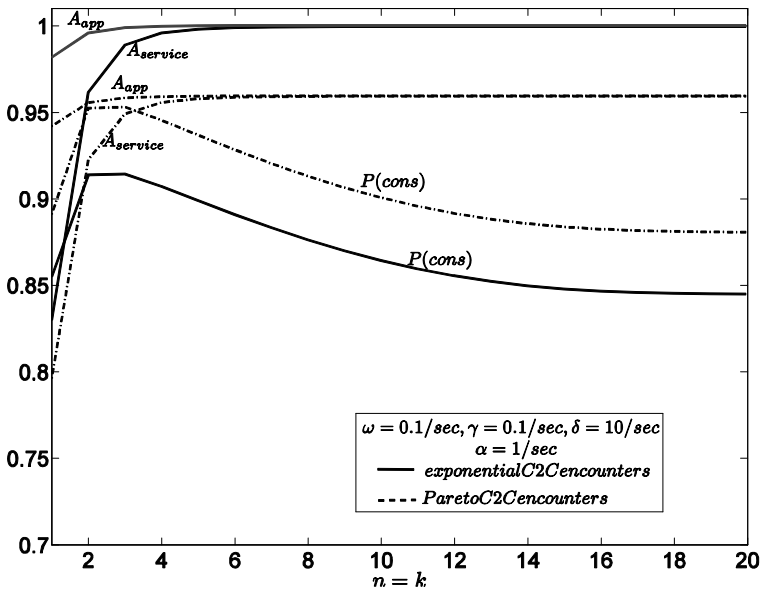

Figure 3.13: Service and Application availability vs. group consistency for different group sizes: *Pareto* vs. *exponential* C2C encounters

3.5 Conclusion and outlook

In this chapter, we have presented an analytical modeling study based on Stochastic Activity Network and Markov chains that allowed us to analyze the behavior of a replication middleware service in *ad-hoc* based dynamic environments, considering quality of service and dependability related metrics. The main measures concern the probability of having fully consistent replicas, as well as application and service availability. In particular, we carried out several sensitivity analysis studies to see how replica consistency is affected in a broad range of scenarios. We can see that the probability to achieve a full group decreases significantly for group sizes with four or more replicas per group. Even though the group will not grow to its full size, the probability of having fully consistent replicas is above 80% in case that the update rate is more than ten times faster than the application state change rate. Remind that the application state change rate reflects the time between two change events and the replica state update rate reflects the time it takes to send an update message to a group replica. With a replica group size of three, the probability of reaching a full replica group is almost 50% higher when α is 100 times bigger than ω.

Moreover we have shown that as service availability increases with a higher number of replicas, the service correctness starts to decrease if the group size goes beyond 4 servers.

The model presented in this chapter assumes perfect replica selection and no signaling overhead when exchanging servers in the replica groups. Thus the trends shown in the results section in this chapter are optimistic with respect to the level of inconsistency. Another limitation of the model is that it does not consider events like network congestion. In a sense network failures and congestion events can be considered as an increase in the ω parameter (leading to shorter duration during which a vehicle can be connected and remains in the group). The exact amount of signaling overhead involved in keeping the participating server vehicles and the user vehicles up to date on the current group member list is left for future studies.

Chapter Four: Availability Modeling of a Virtual Black Box for Automotive Systems

"We all live every day in virtual environments, defined by our ideas."
Michael Crichton, American, Author Quotes

The case study addressed in the previous chapter takes into account the vehicles mobility and C2C encounters in the ad-hoc domain only. The second case study studied in this chapter which concerns the Virtual Black-Box application (VBB), investigates the impact from the dependability point of view of both C2C and C2I encounters. This application consists in storing on a dedicated server at the fixed infrastructure, historical information about the vehicles that can be retrieved in case of a problem. To protect the data against potential losses before an access to the fixed infrastructure is available, the data can be replicated and temporarily stored on neighboring encountered vehicles using wireless communication technologies, before being permanently saved on the server. The aim of this chapter is to analyse and evaluate the availability of the historical data (recorded in the virtual black-box, VBB) taking into account possible data replication strategies, and various mobility scenarios, using Stochastic Activity Networks.

4.1 Introduction

The use of wireless communication technologies to support safety and dependability critical services (hazard warning, safety and traffic management, etc.) poses significant challenges from the dependability point of view. For example, automated highways or platooning applications could constitute an efficient solution to the increasing traffic congestion in urban areas [Hallé & Chaib-draa 2005, Hamouda *et al.* 2008, Hamouda *et al.* 2009a, Hamouda *et al.* 2010a, Hamouda *et al.* 2010b]. On the other side, such solutions may induce safety problems (due for example to the loss of communications for vehicles coordination or to malicious threats) that should be dealt with as early as the design phase. Recording important historical information about the state of the vehicles to be retrieved should provide valuable support that could be used to improve the whole system, or even by the insurance companies, in case of problems.

We concentrate on important information related to a vehicle and its environment, in a manner that is similar to the black box of an aircraft. However, instead of a robust and expensive hardware based black box, the data storage is based on a more cost effective software solution. Periodically, historical data items about the vehicle's state (speed and movements of the Vehicle, actuation of brakes, direction indicator, light, throttle position(s), etc.) are recorded and can be replayed in the event of an accident [Killijian *et al.* 2009]. Typically, the data recorded just before the accident is critical and is very valuable for accident investigation. Information collected on a relatively long period of time may be useful to trace back possible progressive degradation within the vehicle and provide feedback to the driver. Other information may support the vehicle constructor to improve the vehicle design. In this chapter we put emphasis on the first category of data. *i.e.*, data related to accident investigation.

Due to the limited storage facilities within the vehicle, and to the fact that in case of an accident the data stored on the vehicle can be lost, the virtual black box is resident on the fixed

infrastructure (that is usually available for other purposes such as coordination of the traffic or for the communications between the vehicles [Hallé & Chaib-draa 2005]).

To increase the availability of the virtual black box (VBB), the data related to the concerned vehicle, referred to as the *Vehicle*, is replicated temporarily on vehicles encountered by the Vehicle, referred to as participant vehicles (or *participants*). Various data replication algorithms can be used to protect the data against accidental and malicious threats, before they are permanently stored on the fixed infrastructure. The efficiency of such solutions depends on the data replication strategy itself and on various environmental factors, such as the rates at which connections occur with the fixed infrastructure and between the vehicles in the *ad-hoc* domain, and the reliability of the communications.

This chapter presents a dependability model based on Stochastic Activity Networks to compare different replication strategies and to understand how different mobility scenarios can affect the data availability in the context of the virtual black box application. The effect of mobility is studied by considering different distribution laws to describe the vehicle-to-vehicle encountering process based on the results of Chapter 2.

The remaining of this chapter is structured as follows. Section 4.2 outlines the main characteristics of the VBB application and the data replication strategies analyzed in this study. Section 4.3 presents the dependability model and the quantitative measures assessed. Section 4.4 discusses the main results and finally Section 4.5 summarizes the main conclusions.

4.2 The Virtual Black Box application

The Vehicle periodically and almost continuously collects data items in the form of records. A record gathers information related to the Vehicle speed and movements, actuation of brakes, direction indicator, light, throttle position(s), etc. Typically, the size of one record is about 5 *kilo* bytes.

From a practical point of view, the last successive data records generated during the last 15 to 30 *seconds* before the accident form a full set of data allowing tracing back the accident. Let z be the number of records during this period of time. As the data is updated very frequently, the loss of a small number of data records, among this full set of data, may not affect significantly the accuracy of the collected information. Particularly, when we need to understand what had happened just before the occurrence of an accident, it could be sufficient to analyse only r records among the last z records generated before the accident.

Primarily, the records are stored in the Vehicle itself, before being delivered to the fixed infrastructure as soon as a service access point is encountered to complement and update the VBB. Between two updates, the most recent information that is critical in case of an accident is only in the Vehicle itself, and there is a high probability that it will be lost due to the accident. The idea is to take advantage of surrounding vehicles encountered to replicate safely the data before being transmitted to the fixed infrastructure. The temporary data can be transmitted to the fixed infrastructure either by the Vehicle or by the participant vehicles.

Replication of the data records may be handled either by creating full copies of the records (we refer to this as replication by duplication) or by more sophisticated mechanisms using a replication by fragmentation strategy to protect the data records against accidental as well as malicious threats, in the same way as in [Courtès *et al*. 2006].

The replication by fragmentation strategy is based on the erasure coding techniques that are well suited to ensure data availability and confidentiality in the presence of permanent failures [Weatherspoon & Kubiatowicz 2002, Lin *et al*. 2004]. Let us consider a data record collected by the Vehicle at a given time that must be saved at the fixed infrastructure via some participant vehicles. An erasure coding algorithm with parameters n and k ($n \geq k$), produces n fragments of the original record that are scattered among surrounding vehicles. An optimal erasure code allows (n–k) failures (or erasures) to be tolerated (beside that of the primary replica), *i.e.*, k fragments are necessary and sufficient to recover the original data record [Xu *et al*. 1999].

Wherever a participant vehicle gains access to the fixed infrastructure, it transfers all the data fragments that are replicated on its storage facilities. In particular, in case of an accident, it is expected that all the fragments on the participant vehicles will eventually be delivered to the fixed infrastructure to be used for the analysis.

We consider that every encounter between vehicles offers a storage opportunity. Specially, every vehicle encountered is considered to be a participant vehicle that unconditionally accepts storage requests from the Vehicle. The Vehicle unconditionally sends one data fragment to each vehicle encountered. Note that scenarios in which not all encounters offer storage opportunities (e.g., with vehicles refusing to cooperate) can be simply modeled by introducing a participant/encounter ratio as an additional parameter. Replication by duplication of the data corresponds to the case n=k=1.

To sum up, the VBB information that is on the fixed infrastructure is composed of the records related to the state of the Vehicle. The last z successive records are analysed in case of an accident (or at least r records among this full set). A record is produced by the Vehicle. It is either delivered directly to the fixed infrastructure by the Vehicle itself, or it is divided into n fragments that are scattered among surrounding vehicles, to be delivered to the fixed infrastructure. A record can be reconstituted based on at least k fragments in the fixed infrastructure. If less than k fragments reach the fixed infrastructure, the record is considered as lost.

The objective of the dependability model presented in this chapter is to quantify the availability of the VBB application and to analyse the impact of the parameters z and r, and the parameters n, k of the erasure coding algorithm, under different mobility scenarios. The data availability of the VBB application clearly depends on the mobility of the vehicles and their connectivity dynamics, in particular, the rate at which the vehicles meet and the rate at which they meet an access point of the fixed infrastructure. These characteristics depend on the type of the environment (a highway, an urban area, a geographic area with a low density of cars, etc.). Based on the results presented in Chapter 2, we will consider two types of distributions for characterizing C2C encounters (*Pareto* or *Exponential*) and the *Exponential* distribution for modeling C2I encounters.

4.3 Availability modeling

In this section, we present *i*) the measure that will be evaluated to analyse the availability of the VBB application, and *ii*) an availability model based on Stochastic Activity Networks (SAN) allowing the evaluation of this measure, considering the scenario described in Section 4.2.

We analyze the unavailability of the VBB application via the evaluation of the probability of data loss, *i.e.*, the asymptotic probability, noted *UA*, of reaching a state where more than *r* data records among *z* records generated during certain time interval are lost before being delivered to the fixed infrastructure.

Stochastic Activity Networks (SAN) are well suited to evaluate the data availability of the VBB application taking into account the considerations mentioned above. This formalism and the associated Möbius tool provide compositional operators that are useful to master the complexity of the models, both at model construction and model processing phase. In particular, the system model can be built by the composition of atomic models using Join and Replicate operators (see Chapter 1).

In the remaining of this section, we first present an overview of the system model, and then we give more details on the various sub-models.

4.3.1 System model overview

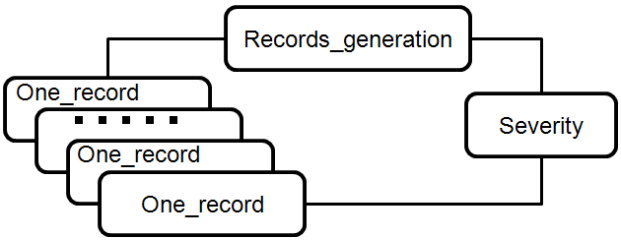

Figure 4.1: Model structure

Figure 4.1 shows the overall structure of the model describing the VBB application. The model includes *z* replicas of the One_record sub model that are composed with two other sub models: Records_generation and Severity.

The One_record sub model describes the behavior of a record as resulting from its data loss modes and the associated data replication strategy presented in Section 4.2. The Severity sub model describes the impact of multiple data losses of data records in the system. The sub model Record_generation is used to model and manage the records generated by the Vehicle in the absence of data loss, to initialize the other sub models and to synchronize their evolution according to the whole system evolution.

4.3.2 The submodels description

Considering the discussion in Section 4.2, our study is based on the following assumptions:

- Failures affecting the vehicles leading to the loss of a data record (at the Vehicle or at the participant vehicles) follow an *exponential* distribution with rate λ.
- Vehicle-to-Vehicle encounters are described by: *i*) an *exponential* Distribution with rate α, or *ii*) a *Pareto* Distribution with shape parameter p and scale s.
- Vehicle-to-Infrastructure encounters follow an exponential distribution with rate β.

One_record: The SAN sub model shown in Figure 4.2 describes the replication and scattering process for an (n, k) erasure code for one record created by the Vehicle. The model focuses on the vehicular *ad-hoc* part of the VBB application, purposefully ignoring issues related to the implementation of the fixed infrastructure functionalities. Thus, one fragment of one disseminated record is considered "*safe*" (*i.e.*, it cannot be lost) whenever either its creator "the Vehicle" or a participant vehicle storing it is able to access the fixed infrastructure. In other words, the server of the fixed infrastructure of the VBB application is assumed to be very reliable (*i.e.*, it has enough redundancy and fault tolerance mechanisms to ensure a very low failure rate that can be neglected). Finally, we assume that when a participant vehicle fails before reaching the fixed infrastructure, all the fragments it holds are lost.

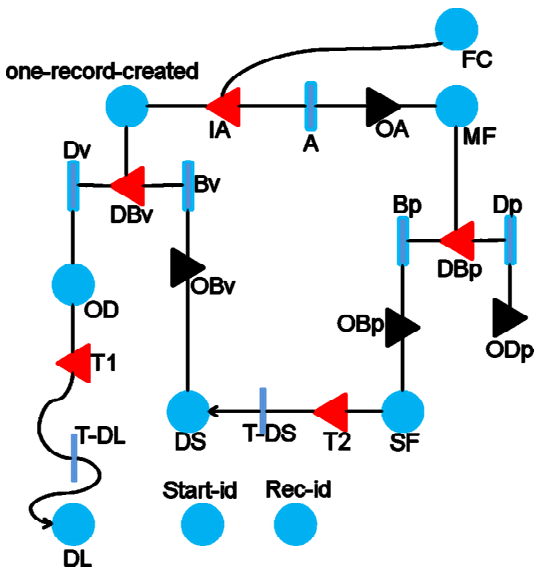

Figure 4.2: One_record SAN model

Thus, with (n, k) erasure coding, one record is definitely lost if and only if the record on the Vehicle is lost and less than k fragments of the data record are available on the participants and at the fixed infrastructure.

The One_record sub model consists of three main processes represented by timed activities:

- A: the C2C encounter process; depending on the considered mobility scenario, this process is represented by an *exponential* distribution with constant rate α or by a *Pareto* distribution with parameters *p* and *s*.

- B: the C2I encounter, represented by (Bv, Bp) with rates β. Bv is for the Vehicle and Bp is for the participants.

- C: the data loss process, at the Vehicle (Dv) and the participants side (Dp); both have *exponential* distributions with constant rate λ.

The sub model in Figure 4.2 is divided into two interacting subnets.

The subnet on the left describes the evolution of a data record at the Vehicle side: either it is lost (with rate λ "activity Dv"), or it reaches the fixed infrastructure (with rate β "activity Bv"). Places one_record_created and OD denote situations where the data record on the Vehicle is "*available*" or "is *lost*", respectively. When a record is created, the place one_record_created will be marked with one token (see Records_generation submodel).

The subnet on the right describes: *i*) the data encoding process with an erasure code (*n*,*k*) leading to the creation of "fragments" (place MF) on participant vehicles as they are encountered (activity 'A'), and *ii*) the process leading to the storage of the fragments (place SF) at the fixed infrastructure (rate β "activity Bp"), or its loss caused by the failure of the participant vehicle (rate λ "activity Dp"). At the top of the right-hand side subnet is place FC whose initial marking denotes the number of fragments to create (*n*). The transition rates associated with the loss of a fragment or its storage on the fixed infrastructure are weighted by the marking of place MF, *i.e.*, the number of fragments that can enable the corresponding transitions. The timed activity 'A' will be fired when the place one_record_created is marked and still there are fragments to distribute to participants (*i.e.*, place FC marked). This is managed by the predicates in the input and output gates, IA and OA, respectively. The firing of activity 'A' decrements the marking of place FC by one and increments the marking of place MF by one, and this is without affecting the marking of place one_record_created again.

Two places with associated immediate activities (T_DS, T_DL) are used in the sub model to identify when the data record is safely stored in the fixed infrastructure (place DS), or is definitely lost (place DL), respectively. The "data *safe*" state is reached (*i.e.*, DS is marked) when the original data record from the Vehicle or at least *k* fragments from the participants reach the fixed infrastructure. The "data *loss*" state is reached (*i.e.*, DL is marked) when the original data record from the Vehicle is lost, and less than *k* fragments are available on the participants and at the fixed infrastructure side. This condition is represented by a predicate associated with the input gate (T1). Finally, the predicates and the functions associated with the output gates (OBv, ODp, OBp) and the input gates (DBv, DBp, T2) manage the sub model by applying "*liveliness predicate*", true *if and only if* $m(DS) = m(DL) = 0$: as soon as either DS or DL contains a token, no activity can be fired in this submodel.

Place Rec_id saves the ID of each data record generated in the system. Place start_id is used for the initialization of the sub model (*i.e.*, a new One_record is generated). When the place DL is marked for an One_record submodel and still the system generating the records is available, another record will be generated by adding one token to the place Start_id.

Severity: This sub model presented in Figure 4.3 describes how the combination of individual record losses could bring the system to total loss situations as described in Section 4.2.

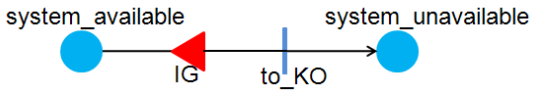

Figure 4.3: Severity SAN model

The system_available place is initially marked by one token. The predicates and functions associated with the input gate IG in this sub model describe the impact on the global availability of multiple losses affecting several records. When the instantaneous activity to_KO is fired, the system_unavailable becomes marked indicating that the VBB application has reached an unavailable state and has less than r safe records.

Records_generation: This submodel, presented in Figure 4.4, is used to define the initial configuration of the records and to initialize the One_record submodels associated with each participant vehicle. For the Vehicle, up to *z* records are generated during certain time interval. Thus the system model is composed of *z* replicas of the One_record submodel.

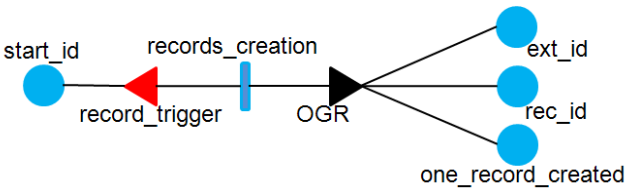

Figure 4.4: Records_generation SAN model

The Record_generation submodel contains four places; all of them have initially zero token except Start_id that has one token. Places (Start_id, Rec_id, and one_record_created) are shared with the corresponding One_record submodel replicas included in the configuration of the VBB. The place ext_id is a global place shared by all sub models, to act as a counter. Initially *z* replicas are created. Each time the timed activity records_creation is fired; a new record is generated and assigned a Rec_id. Also place one_record_created is marked to initialize the One_record submodel associated with this record. The ID assigned to the record is stored in the place Rec_id. When a new record is created, rec_id gets the value stored in Ext_id, which in turn is incremented by one. Moreover, the predicates in the input gate record_trigger are used for managing the *z* records generation, and for stopping all generation when the system arrives to the unavailable state (*i.e.*, system_unavailable is marked in Severity submodel).

4.3.3 The SAN complete model

The SAN model resulting from the composition of the SAN submodels presented in Figure 4.2, 4.3 and 4.4, using joining "Join" and replication "Rep" composition operators, is illustrated in Figure 4.5.

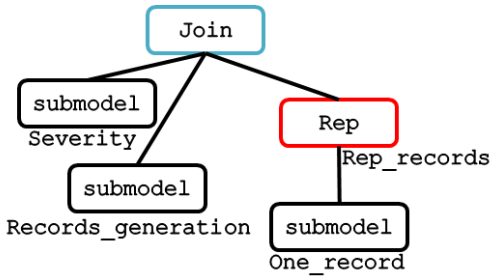

Figure 4.5: SAN complete model

4.4 Results and Sensitivity Analysis

In this section, we illustrate the type of results obtained from the processing of the SAN model of Figure 4.5, and show the sensitivity analyses with respect to various parameters affecting the VBB data availability.

The quantitative measure *UA* defined in Section 4.3 corresponds to the probability of having a token in the place system_unavailable of Figure 4.3. The analyses focus on the impact on *UA* of the parameters defined in Table 3.1.

Table 4.1: Parameters

α	The vehicle-to-vehicle encounter rate
β	The vehicle-to-infrastructure encounter rate
c	The connectivity ratio = α/β (the rates at which vehicles meet relative to the rate at which connection to the fixed-infrastructure is possible)
λ	Data loss rate, for the Vehicle and the participants
n, k	Parameters of the erasure code
r, z	The accuracy required of the historical information to analyse what happened when an accident occurs

As regards the rate at which vehicles can connect to the fixed infrastructure, we consider an average nominal value derived from the analyses of chapter 2, β=150/*h* that corresponds to situations where an access point is available each about 30 *seconds*. This is the case for example of a vehicle moving at an average speed of 100 *km*/*h* on a highway with an access point each 1.5 *km*. The consideration of other scenarios with a higher or a lower density of access points is taken into account through the analysis of different values of the connectivity ratio.

A question is raised concerning the number of records *r* needed among the *z* last generated records to reconstitute an accident. For example, if there is an accident, we will be interested to

know the state of the Vehicle for the last 15 *seconds*. If a vehicle is moving with an average speed of 80 to 130 *km*/*h* and the Vehicle is generating a record each 2 to 5 *seconds*; we can look forward to at least 3 or 4 records among 5 generated records during the last 15 *seconds* before the accident. In other words, we can tolerate the loss of two records, in the worst case, during the last 15 *seconds* before the accident. Different values of *r* and *z* will be analysed to highlight the impact of these parameters on the VBB data availability.

In the remaining of this section, we will first concentrate on the unavailability of one data record. Then, we will address the unavailability of the VBB application.

4.4.1 Unavailability of one data record

The unavailability of one data record generated by the Vehicle corresponds to the probability of having a token in place DL of Figure 4.2.

We first consider the case of replication by duplication and then we compare the two replication strategies (by duplication and by fragmentation).

4.4.1.1 Replication by duplication

Figures 4.6 and 4.7 show the unavailability of a single data record as a function of the failure rate λ, in the case of the replication by duplication strategy (n=k=1) when the vehicle encounters are described by either an *exponential* or a *Pareto* distribution. Also, these figures present the results corresponding to the case when replication in the *ad-hoc* domain is not used n=k=0; (*i.e.*, the data record is stored on the Vehicle until it is delivered to the fixed infrastructure when an access point is reached). This will highlight the potential benefit of using encounters in the *ad-hoc* domain to improve the data availability.

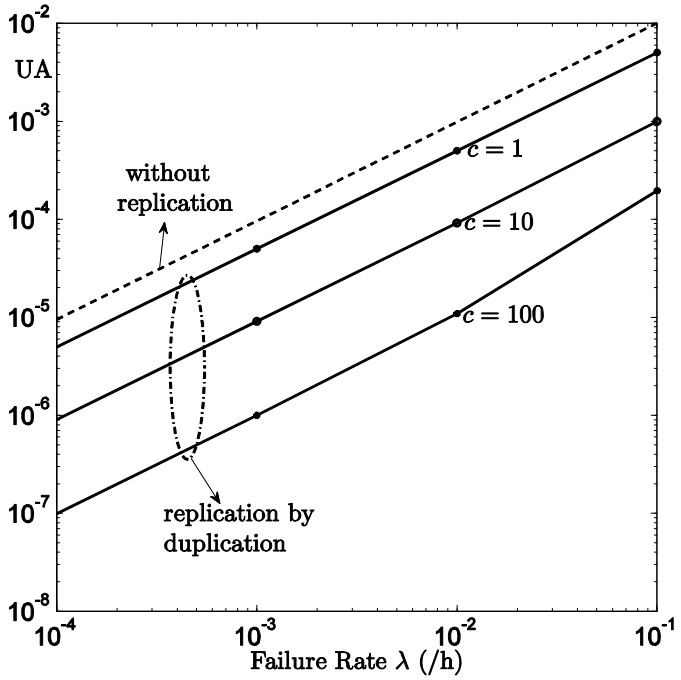

Figure 4.6: Data record unavailability for *exponential* encounters

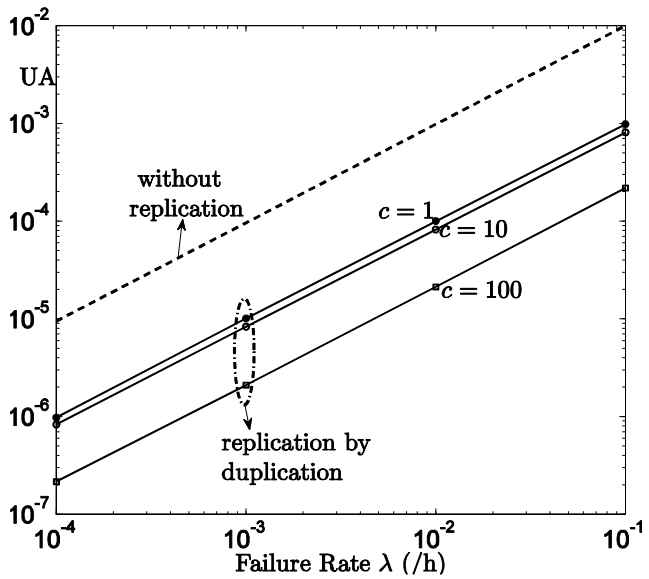

Figure 4.7: Data record unavailability for *Pareto* encounters

The results are shown for three different values of the *ad-hoc*-to-fixed infrastructure connectivity ratio *c* (1, 10, 100) corresponding to different environments and mobility scenarios. The parameters assigned to the *Pareto* distribution (*p* and *s*) are such that the mean time between

encounters is equal to $1/\alpha$, (the mean time between encounters of the *exponential* distribution to which it is compared).

It can be observed that in both figures, *UA* is almost proportional to the failure rate λ. The level of unavailability depends on the value of the *ad-hoc*-to-fixed infrastructure connectivity ratio *c* characterizing the corresponding mobility scenario. For a given failure rate λ, *UA* is inversely proportional to *c*. However, the relative variation of the unavailability when increasing *c* is different for *exponential* and *Pareto* encounters. Considering first the case of exponentially distributed encounters (Figure 4.6), for $\lambda=10^{-3}/h$, the unavailability corresponding to c=10 is 9 times higher compared to c=100, whereas the variation is about 5.5 times only when increasing *c* from 1 to 10.

The situation is slightly different for *Pareto* distributed encounters. Indeed, the relative variation of the data unavailability as a function of *c* remains low:

- About 1.2, when comparing mobility scenarios corresponding to c=1 and c=10, and
- About 3.9, when comparing mobility scenarios corresponding to c=10 and c=100.

The comparative analysis of the impact of the *exponential* and the *Pareto* distributions on the data unavailability is more visible in Figure 4.8 which plots together the results corresponding to these cases, for c=1 and c=100.

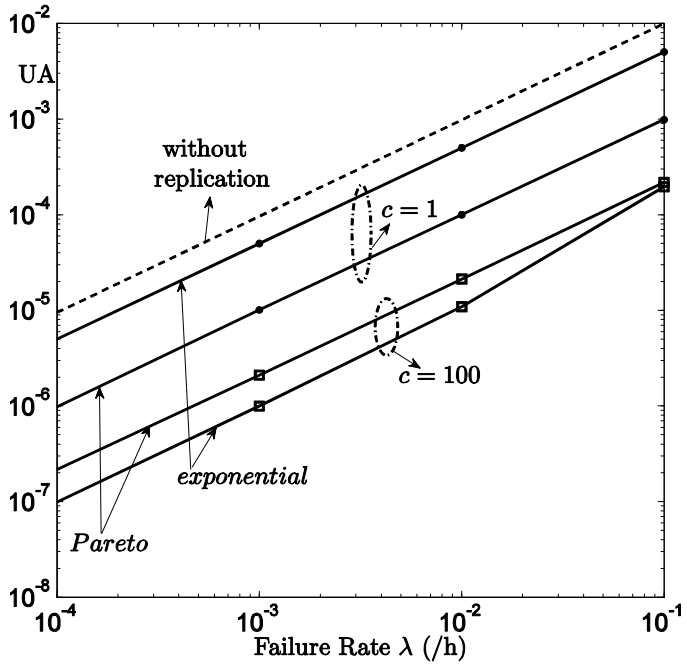

Figure 4.8: Data record unavailability: *Pareto* and *exponential* encounters

Another interesting observation that can be derived from Figures 4.6, 4.7 and 4.8 concerns the analysis of the potential gain brought by the use of replication compared to scenarios where replication is not used. It can be seen that the gain mainly depends on the characteristics of the considered environment, especially, the connectivity ratio *c*: the higher is this ratio, the better is the availability gain obtained. Also, we can observe that the maximum gain, for the parameters considered in our study, is obtained in mobility scenarios with *Pareto* distributed encounters for *c*=1, and in scenarios with *exponential* distribution encounters for *c*=100.

4.4.1.2 Impact of the replication strategy

Figures 4.9 and 4.10 highlight the impact of the replication strategy on the unavailability *UA* when considering a single record and *exponential* and *Pareto* distributed encounters, respectively.

Each figure compares the unavailability associated to the replication by duplication strategy (denoted as *UA*(1,1)) to the unavailability associated to the replication by fragmentation strategy using an erasure code (*n*, *k*) (denoted as *UA*(*n*,*k*)). The comparison is made through the ratio *UA*(1,1)/*UA*(*n*,*k*).

Similar results are obtained for the *exponential* and *Pareto* encounter scenarios. It can be noticed that, in both cases, the unavailability decreases as *n* and *k* increase. Also, in both cases, it appears that the potential gain yielded by the use of an erasure code compared to replication by duplication is rather modest (less than 2 times if we consider the most favorable case). Nevertheless, it should be mentioned that a noticeable advantage of erasure codes over the replication by duplication is the fact that they are generally used to ensure security related properties too, in particular confidentiality. The higher is *k* the better is the confidentiality. However, the assessment of these properties is not the focus of this chapter.

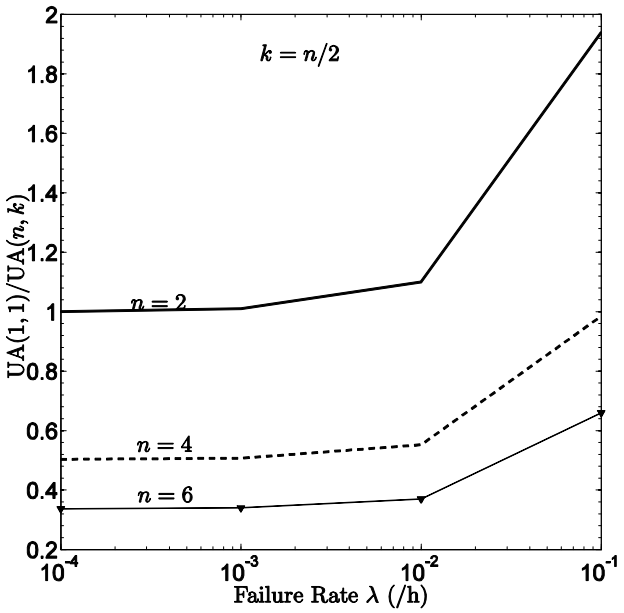

Figure 4.9: Impact of the replication strategy: one record, *exponential* encounters, *c*=100

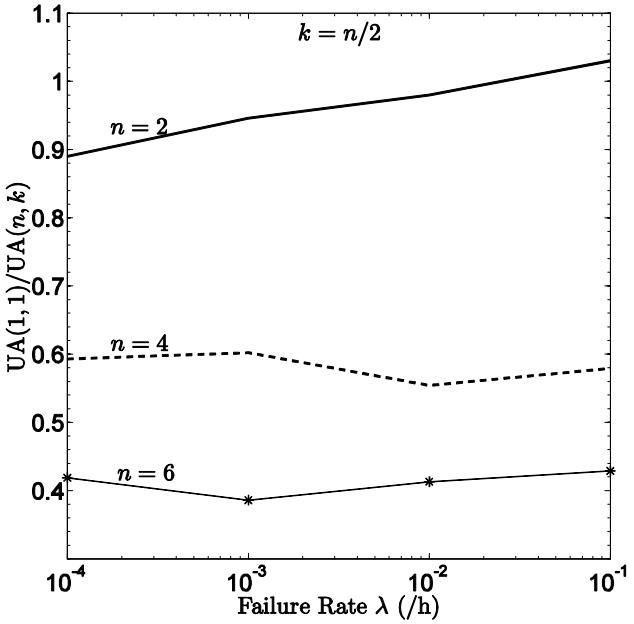

Figure 4.10: Impact of the replication strategy: One record, *Pareto* encounters, *c*=100

4.4.2 VBB unavailability

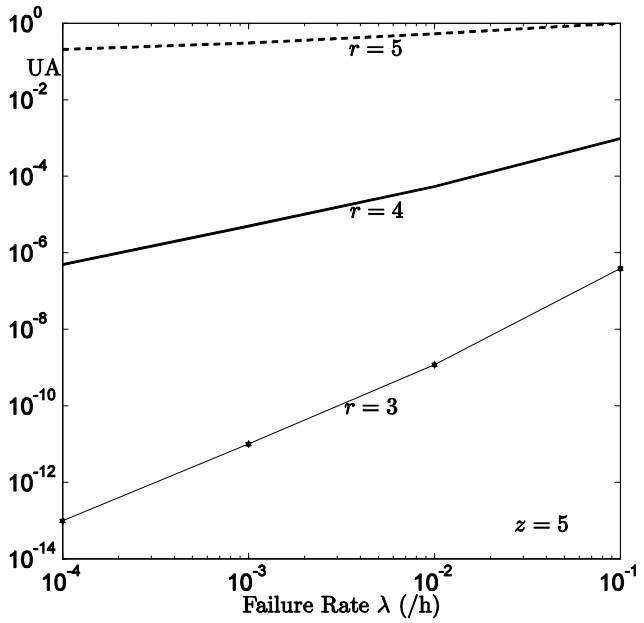

Figure 4.11: VBB unavailability: replication by duplication, *exponential* encounters, *c*=100

In this section, we evaluate the unavailability of the VBB application considering the case of multiple records such that *r* among the last *z* generated records are needed to analyse what happened if an accident has occurred. We consider the case of replication by duplication, and mobility scenarios characterized by *exponential* encounters. Similar trends have been observed for the case of *Pareto* distributed encounters.

Figure 4.11 shows the VBB unavailability considering the case *z*=5 with different values for *r*. It shows that when *z* is fixed the unavailability varies linearly with and is highly sensitive to *r*. In addition, the proportionality factor decreases as λ increases. As an example, the increase of r from 3 to 4 yields a five orders of magnitude decrease of the unavailability for $\lambda = 10^{-3}/h$ and four orders of magnitude for $\lambda = 10^{-2}/h$.

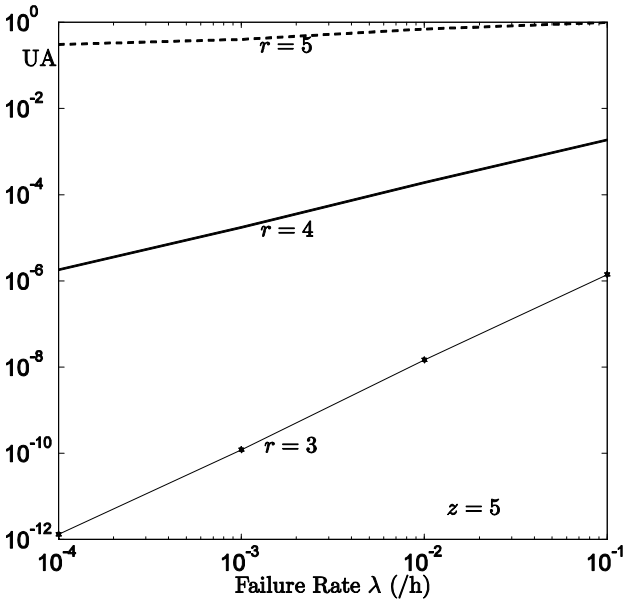

Figure 4.12: VBB unavailability: replication by duplication, *Pareto* encounters, *c*=100

The impact of the variation of z on the VBB unavailability is illustrated in Figure 4.13. The results concern three different pairs (*z*, *r*). It can be noticed that the higher is *z* the better is the unavailability and the difference is significant (e.g., the unavailability corresponding to (*z*=7, *r*=5) is 7 times higher than the unavailability with (*z*=5, *r*=3) when $\lambda = 10^{-3}/h$).

The last result highlighted in Figure 4.14 concerns the comparison of the scenario where replication is not used with the replication by duplication strategy, through the ratio of the unavailabilities corresponding to these cases denoted as *UA*(0,0)/*UA*(1,1). For the sake of comparison, we present together the results corresponding to *exponential* and *Pareto* encounters for two connectivity ratios: *c*=10 and *c*=100. It is noteworthy that the results are almost similar for the different values (*z*=7, *r*=5), (*z*=5, *r*=3), and (*z*=3, *r*=1), considered in Figure 4.13.

It can be seen that the potential gain due to replication is significant especially in the environments where the failure rate λ is low (e.g., $10^{-4}/h$). The maximum gain (about 10^4) is

obtained in the case of *exponential* encounters for *c*=100. The comparative analysis of the *exponential* and *Pareto* results shows that the distribution of encounters can have a significant impact in this case.

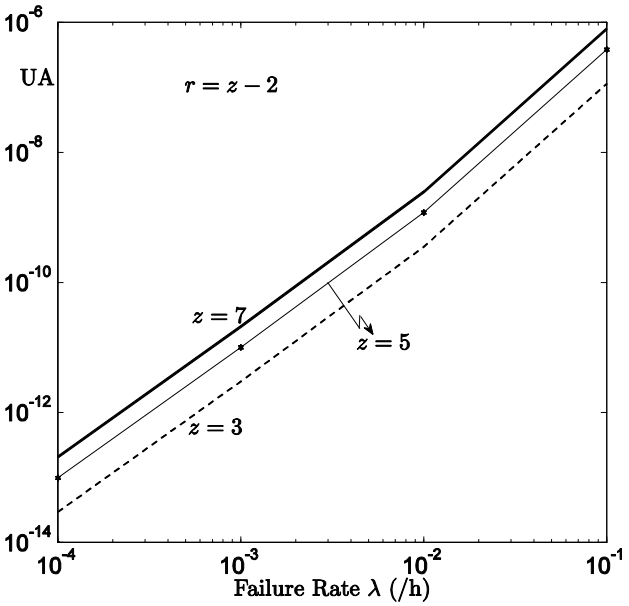

Figure 4.13: Impact of z on VBB unavailability: simple replication, *exponential* encounters, *c*=100

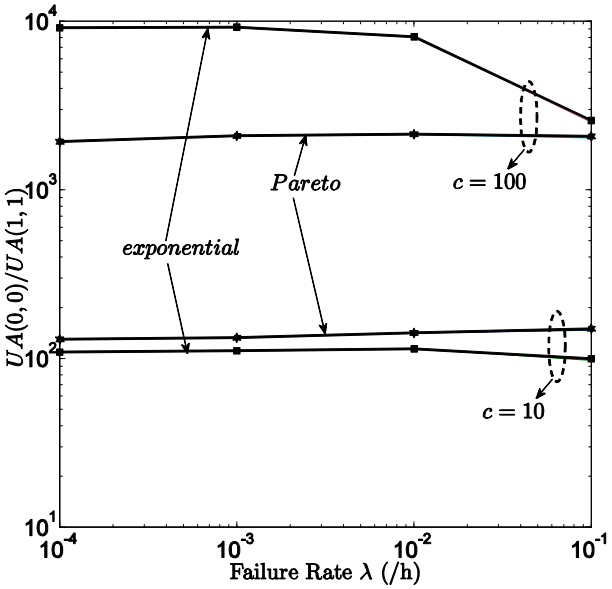

Figure 4.14: Replication by duplication vs. no replication: *c*=100

4.5 Conclusion

This chapter addressed the dependability modeling and evaluation of an innovative application in automotive systems that are aimed at providing a virtual black-box for storing historical information about the state of the vehicles on a dedicated server at the fixed infrastructure. To protect the data records against potential losses before they are delivered to the server, they are replicated and disseminated to encountered vehicles with which they can communicate using wireless technologies. Each of these vehicles delivers its data to the server when they can connect to the fixed infrastructure.

This chapter focused on the availability evaluation of the historical data recorded in the black-box using Stochastic Activity Networks, considering mainly the data recorded just before an accident. The proposed model allowed us to compare different data replication strategies under various mobility scenarios and environmental factors.

This chapter includes significant extensions of the previous work related to the dependability evaluation of a cooperative backup service reported in [Courtès 2007]. The latter, based on generalized stochastic Petri nets, assumed an exponential distribution for the vehicle encounter process, and addressed the case of a single data record only. Here, we extended the original model to analyse the availability of the virtual black-box application taking into account multiple data records and considering more general mobility scenarios. For example, we included in the analysis situations where encounters are described by a *Pareto* distribution as observed in Chapter 2, to assess the impact of the distribution of encounters on the behavior of the application under different replication strategies. For this purpose, we used Stochastic Activity Networks.

The results obtained in this study show that the VBB application may exhibit different levels of availability when comparing mobility scenarios with *Pareto* encounters and scenarios with *exponentially* distributed encounters, depending on the value of the *ad-hoc*-to-fixed infrastructure connectivity ratio in the corresponding environment. Another noteworthy result concerns the availability gain yielded by the data replication in the *ad-hoc* domain compared to the case where such a replication is not used. The unavailability can be decreased thanks to replication by a factor up to the connectivity ratio (100 in our study) when considering a single data record. This result confirms the conclusion obtained in [Courtès 2007] and shows that it is also valid in mobility environments where encounters are described by a *Pareto* distribution. Moreover, when considering the availability at the VBB application, our study shows that the gain brought by replication can be significantly higher (e.g., 10^4 for *exponential* encounters and c=100).

The analyses presented in this chapter can be refined by considering safety-related measures, taking into account the distribution and the rate of occurrence of accidents. Also, the model can be extended to analyse the dependability of the historical information recorded by the VBB application, over a long period of time and not only considering the data recorded just before an accident.

Chapter Five: Automated Highway Systems Safety

"*The most terrible job in warfare is to be a second lieutenant leading a platoon when you are on the battlefield"*
Dwight D. Eisenhower, 34th President of the USA

This chapter addresses the safety modeling and evaluation of Automated Highway Systems that are based on the use of platoons of vehicles driven by automated agents. This case study investigates some critical issues that have not been covered in the previous Chapters 3 and 4. In particular, we analyze the safety impact of the strategy used to coordinate the vehicles' operations, inside each platoon and between platoons, when vehicles enter or exit the highway, or when maneuvers are carried out to recover from failures affecting the vehicles or their communication. Some of the results presented here are described in [Hamouda *et al.* 2008, Hamouda *et al.* 2009a].

5.1 Introduction

Traffic congestion is increasingly growing especially in urban areas. One of the solutions for this problem is automated traffic. Many research programs have been carried out or are currently underway to implement Automated Highway Systems (AHS), based on automatically controlled platoons of vehicles. The objective of the AHS is to reduce congestion and increase the throughput and safety of the highway by adding automation to the vehicles and the roadside without having to build new roads. The investigated techniques are aimed at providing guidance for vehicles to improve the traffic flow and the highway safety by reducing accidents, while reducing fuel consumption and pollution. In this context, several studies have been dedicated to collaborative driving systems, based on coordinated vehicles on highways equipped with the necessary infrastructure (see e. g., [Fenton *et al.* 1976, Fenton & Mayhan 1991, Varaiya 1993, Benz *et al.* 1996, Godbole *et al.* 1996, Furukawa 2000, Tsuji *et al.* 2001, Masayasu *et al.* 2003, Burggraf *et al.* 2007, El Masri *et al.* 2010]). They were particularly devoted to the design of control architectures for automatic driving and their verification, and to performance evaluation in terms of capacity and traffic flow [Burggraf *et al.* 2007]. To the best of our knowledge the safety modeling and quantitative evaluation of such systems have been seldom addressed. This problem is challenging in the domain of automated highway systems implemented on ad-hoc networks.

In this chapter, we address safety of AHS based on platooning applications implemented in a mobile context with ad-hoc networks. A platoon is a series of coordinated vehicles that are moving in the same direction on a highway [Godbole *et al.* 1996]. The vehicles are driven by more or less automated agents, interacting in a multi-agent environment [Hallé & Chaib-draa 2005]. Switching to manual driving is possible under specific circumstances.

Our work aims at developing evaluation approaches and models that make it possible to analyze the AHS safety taking into account several phenomena, such as accidental fault occurrences, success and failures of the recovery maneuvers, and vehicles coordination strategies. The developed models are aimed at providing support to the designers for the analysis of possible solutions of AHS, based on safety evaluation.

We consider as a case study the architectures developed in the context of the PATH project (Partners for Advanced Transit and Highways [Miller 1997]). These architectures implement automatic recovery maneuvers to ensure the platoons' safety in the presence of different types of failures affecting the vehicles and their environment. To this end, they require coordination between the vehicles in the platoon and with neighboring platoons. Various coordination strategies can be considered. The models presented in this chapter are aimed at analyzing and quantifying the impact of such strategies and other characteristics of AHS on the traffic safety.

To cope with the complexity of the studied system, we have developed models based on Stochastic Activity Networks, to evaluate the impact of vehicle failures as well as maneuvers failure and success, on the Automated Highway Systems safety. Replicated submodels associated with each vehicle, describing the corresponding failure modes and recovery maneuvers and their severity, are composed with submodels characterizing the configuration of the platoons and their dynamic evolution. Numerical results are presented to highlight the impact of the coordination strategy and other dependability related parameters.

The chapter is organized as follows. Section 5.2 presents the automated highway system considered, together with its failure mode analysis. Section5.3 presents the proposed safety modeling approach and its associated SAN model. Section 5.4 summarizes the results obtained and discusses their impact on the design of platooning applications. Finally, Section 5.5 concludes and summarizes the main lessons learned.

5.2 System description

Each platoon is composed of a *leader* that is the first car of the platoon and a set of followers. A platoon that contains one vehicle is called *free agent*. Figure 4.1 shows three platoons: *p1* with three vehicles, a leader and two followers, *p2* is a neighboring platoon, and *p3* is an example of free agent. The intra-platoon distance (Δx) ranges usually between one to three meters. The inter-platoon distance between two platoons (Δp) in the same lane varies between thirty and sixty meters.

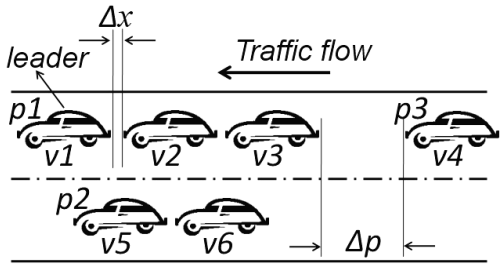

Figure 5.1: Context of a platooning application

Moreover, the gaps between the platoons can be chosen large enough to guarantee no collision. Large gaps serve the additional purpose of attenuating disturbances (such as rapid decelerations) and preventing them from propagating far upstream.

The PATH research program has defined hierarchical control architectures for platooning applications. The platoons use lateral and longitudinal positioning controllers (magnetic equipments) to allow the vehicles to safely follow each other. The vehicles are coordinated by means of communications, based among other things on information from the magnetic equipments. Several maneuvers have been defined to allow the system to be in safe conditions in the absence and in the presence of failures (*fail-safe* mode).

The main maneuvers consist in splitting a platoon, merging platoons, or making a vehicle exit or enter the platoon. In case of a failure affecting a vehicle in the platoon, the maneuvers allow the vehicle to leave its platoon without any hazard, for the purpose of continuously running the platoon without any problem. Before starting a maneuver, the faulty vehicle communicates with its platoon's leader (that initializes the *coordination* of the *maneuvers*). According to the failure mode, some maneuvers may require a communication between the leaders of neighboring platoons in addition to communications with adjacent vehicles [Lygeros *et al.* 1996]. If the faulty vehicle is the leader, specific maneuvers must be applied to allow the platoon vehicles to select a new leader.

We briefly present background information on the PATH architecture that is needed to understand our safety models. We mainly focus on the failure modes considered and the recovery maneuvers used to ensure AHS safety, taking into account different strategies for intra-platoon and inter-platoon coordination.

5.2.1 PATH Architecture

Figure 5.2 presents one of architectures proposed by the PATH project in [Godbole *et al.* 1996, Varaiya 1993]. Five hierarchical layers are distinguished. The top two layers of the hierarchy provide centralized control for vehicles on entire sections of the highway. Together they form the "*roadside controller*". The top layer, called the *network layer*, is responsible for the traffic flow on the entire highway system. Its task is to prevent congestion and maximize throughput by dynamic routing of traffic along the interconnected network of highway. The second layer called the *link layer* coordinates the operation of sections, or links, of the highway. Its primary concern is to maximize throughput while maintaining safe conditions of operation. It decides which lanes the vehicles should travel in to get to their destination as fast as possible, the speeds at which they should travel and the size of the platoons they should form. It also monitors incidents on the highway and diverts traffic in order to minimize the impact using aggregate statistical information about the traffic in a section. [Rao & Varaiya 1994] provides details of the link layer design.

The remaining layers are responsible for decentralized control of individual vehicles. The task of the *coordination layer* is to coordinate the operation of platoons with their neighbours, using the main maneuvers mentioned in the beginning of this section. The coordination layer controller receives the link layer commands and translates them into specific maneuvers that the platoons can carry out by sending them to the *regulation layer*. The last task is to translate the coordination layer maneuvers into commands to throttle, steering and brake input for the actuators on the vehicle. For this purpose it utilizes a number of continuous time feedback control laws [Godbole & Lygeros 1994] that use the sensor readings to calculate the actuator inputs required for a particular maneuver. The bottom *physical layer* is not part of the control hierarchy. It contains the

actual plant (the vehicles with their sensors, actuators and communication equipment and the highway topology).

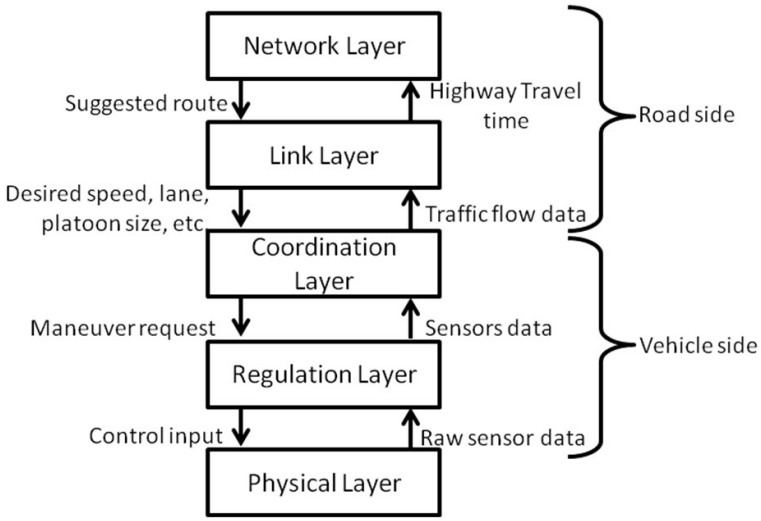

Figure 5.2: PATH hierarchy architecture

5.2.2 Failure modes and recovery maneuvers

Several failure modes, with various severity levels, can affect the vehicles involved in platoons and their safety [Lygeros *et al.* 1996], [Godbole *et al.* 1996, Lygeros *et al.* 1996, Lygeros *et al.* March 2000]. The classification of [Lygeros *et al.* 1996] considers various failure conditions affecting the vehicles or the road side infrastructure components. In the case study presented in this chapter, we only consider the vehicle faults (i.e., *vehicle side*) assuming that the road side controlling infrastructure is reliable.

Depending on the failure severity, various maneuvers can be considered to ensure the safety. Some maneuvers may need to stop the faulty vehicle or help it to exit safely from the highway as soon as possible with the assistance of adjacent vehicles[5]. In fact, the selection of a maneuver depends on three factors:

i) the type of the faults;

ii) the capabilities of the faulty vehicle;

iii) the capabilities of the adjacent vehicles.

[5] *Adjacent vehicles*: refer to the vehicles providing assistance to the faulty vehicle, for example to help it to get out of the highway.

In the following, we first present the failure modes that might affect a single vehicle, their severity and the associated maneuvers. Then, we discuss the case of failures affecting multiple vehicles. Finally, we present the catastrophic situations that could lead the automated highway system to an unsafe state.

5.2.2.1 Single vehicle failures

Six potential failure modes have been identified, presented in Table 5.1. This table shows for each failure mode, some examples of cause leading to the failure mode, the severity class, and the maneuver that ensures the safe continuity of service despite the presence of failures. It does not include minor failure modes that do not warrant specific control actions as the safety of the vehicle is not threatened.

The severity classes associated with the failure modes are ranked by decreasing order. *Class A* (including *A1*, *A2* and *A3*) is the highest, gathering the most critical failures that need to stop the vehicle on the highway. In this case, the vehicle cannot continue moving on the highway safely and has either already come to a stop or it should be commanded to do so and wait to be towed away. Three maneuvers are defined for this purpose:

- *Gentle Stop* (*GS):* the faulty vehicle uses its brakes smoothly to stop;
- *Crash Stop* (*CS)*: the faulty vehicles uses maximum emergency braking;
- *Aided Stop* (*AS)*: the faulty vehicle is stopped by the vehicle immediately ahead.

Specific control laws are then used to ease congestion, divert traffic away from the incident, assist emergency vehicles, and get the queued vehicles out.

The severity classes (*B* and *C*) include the failure modes that can be recovered by allowing the faulty vehicle to get out of the highway without stopping the traffic. For *Class B* failures (including *B1* and *B2*), the vehicle may continue moving but has lost some essential capability and it must therefore exit the highway as soon as possible. Moreover, it needs the assistance of its neighbors to do so. Two maneuvers are defined for this purpose:

- *Take Immediate Exit-Escorted* (*TIE-E):* the faulty vehicle leaves the system as part of a two vehicle platoon in which the faulty vehicle is the follower;
- *Take Immediate Exit* (*TIE*): where the faulty vehicle gets out of the highway as soon as possible alone as a free-agent platoon.
- When a *Class C* failure occurs, the vehicle needs no assistance to exit. Typically the vehicle is fully functional but should leave the system soon to avoid further problems and hazards (in case a second failure mode occurs for example). In this case, the following maneuver is applied:
- *Take Immediate Exit-Normal* (*TIE-N):* the faulty vehicle uses normal lane change and split protocols of [Hsu *et al.* 1994] instead of emergency lane change.

Table 5.1: Failure modes and associated maneuvers

Failure mode	Examples of cause	Severity class	Associated Maneuver
FM1	No brake or throttle control	A3	Aided Stop (AS)
FM2	Inability to detect vehicles in adjacent lanes; No steering control	A2	Crash Stop (CS)
FM3	Inter-vehicle communication failure; No sensing of distance and velocity of car ahead (long and short range)	A1	Gentle Stop (GS)
FM4	No control of transmission / selection of gear; No lateral sensing of vehicles	B2	Take Immediate Exit-Escorted (TIE-E)
FM5	Reduced steering capability	B1	Take Immediate Exit (TIE)
FM6	Single failure in a redundant sensor set	C	Take Immediate Exit-Normal (TIE-N)

Details about the atomic maneuvers composing each of the six maneuvers presented in Table 5.2 and the *inter-vehicle* coordination required to implement them, are presented in [Lygeros 1995]. A summary definition of these atomic maneuvers is given in Table 5.2.

Table 5.2: Atomic maneuvers

Atomic maneuver	Description
Forced Split (FS)	Used by a faulty vehicle to become a free agent.
Emergency Lane Change (ELC)	Used by a free agent with reduced capabilities. The faulty vehicle requests the leader of the platoon in the adjacent lane to create a gap so that the faulty vehicle can change lane.
Front Dock (FD)	Initiated to join with the vehicle in front. The initiating vehicle requests the leader of the preceding platoon for a front dock. This leader orders the last vehicle in its platoon to front dock with the initiator. This vehicle will then decelerate to close the gap between itself and the initiator. In the end, the initiator becomes the first follower of the new platoon. The leader of this new platoon joins with the trailing platoon by decelerating.
Atomic Gentle Stop (AGS)	Initiated by a follower with brakes off failure. The faulty vehicle decelerates and asks the leader for assistance to stop on the highway. The vehicle immediately ahead of the faulty vehicle applies gentle braking and lets the faulty vehicle collide with it from behind. Then, it uses its brakes to bring the combined mass of both vehicles to a stop.
Atomic Crash Stop (ACS)	Used by a faulty vehicle that is ordered to stop and can do so by using its brakes. The fault is not severe enough to require maximum emergency braking. The vehicle will use gentle braking to minimize the disturbances to following vehicles.
Atomic Aided Stop (AAS)	Similar to Gentle Stop except the severity of the fault requires the faulty vehicle to apply maximum emergency braking

It is noteworthy that the severity class also determines the priority of the corresponding maneuver. This is important when multiple failure modes occur. The priorities within each class are as follows: Within *Class A, A3* has the highest priority and *A2* has a higher priority than *A1*. In *Class B, B1* and *B2* have equal priority. In case of occurrence of multiple failure modes in the same vehicle, the maneuver with the highest priority is applied.

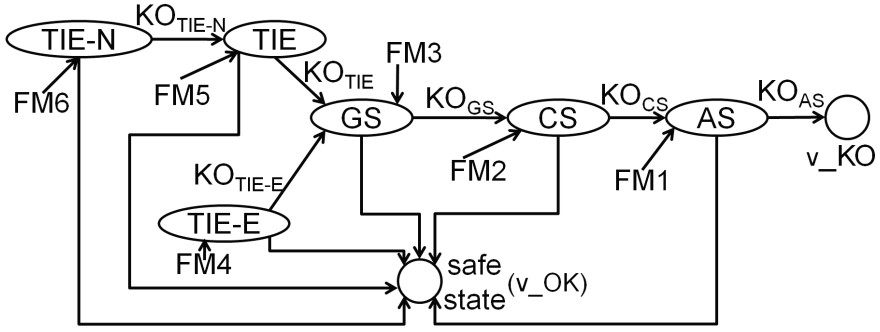

Figure 5.3: failure modes, maneuvers, safety impact

The successive failure of maneuvers may eventually lead to a state where no maneuvers are available to recover the faulty situation. This is illustrated by the state machine in Figure 5.3, where v_KO identifies such a state. The transitions correspond to the occurrence of failure modes, or to the results of maneuver executions that might succeed (transitions to the safe state, v_OK) or fail (KO transitions). Whether the state v_KO corresponds to an unsafe state for the AHS or not, depends on the state of the adjacent vehicles (this is discussed in section 5.2.2.3).

5.2.2.2 Multiple vehicles failures

When nearly simultaneous failures affect multiple vehicles, in particular adjacent vehicles, in the same platoon or in neighboring platoons, the maneuver with the highest priority is applied. The success of a maneuver depends on many factors, for example, the state of faulty vehicles in the platoon, the capability of the *adjacent vehicles* needed to assist the faulty vehicle to realize the maneuver (particularly the leaders concerned by the maneuver), and the traffic flow.

As an example, let us assume that a vehicle *v1* is faulty and has to perform the *TIE* maneuver. If another vehicle is already performing a maneuver with a higher priority, the maneuver requested by *v1* will be refused. Hence, *v1* will ask for another maneuver of a higher priority until the requested maneuver is accepted. Similarly, when a maneuver fails, the system evolves towards a more degraded failure mode and one of its associated maneuvers must be attempted to put the system in a safe state.

5.2.2.3 Impact of failures on the AHS safety

The scenarios described in Figure 5.3 concern a single vehicle. Catastrophic situations leading the system to an unsafe state result from the occurrence of simultaneous failures affecting multiple adjacent vehicles in a small neighborhood in space and in time.

Based on the analysis presented in [Lygeros 1995], we summarize in Table 5.3 three catastrophic situations that would lead the AHS to an unsafe state, taking into account the number of failures affecting different adjacent vehicles and their severity.

Table 5.3: Catastrophic situations

Situation	Description
ST_1	At least two *Class A* failures
ST_2	At least one *Class A* failure AND { (two *Class B* failures) OR (one *Class B* AND one *Class C* failures) OR (three *Class C* failures) }
ST_3	At least four failures whose severities correspond to *Class B* or *Class C*

5.2.3 Vehicles coordination

Platooning applications require coordination between the vehicles in the platoon (*intra-platoon*) and with neighboring platoons (*inter-platoon*). A vehicle is involved in the coordination process when it creates a platoon, it enters an existing platoon, or it leaves a platoon to switch to manual driving. Various communication models (*centralized* and *decentralized*) have been proposed in [Hallé 2005] for the *inter-* and *intra-platoon* coordination, based on the PATH architecture. They are briefly summarized hereafter.

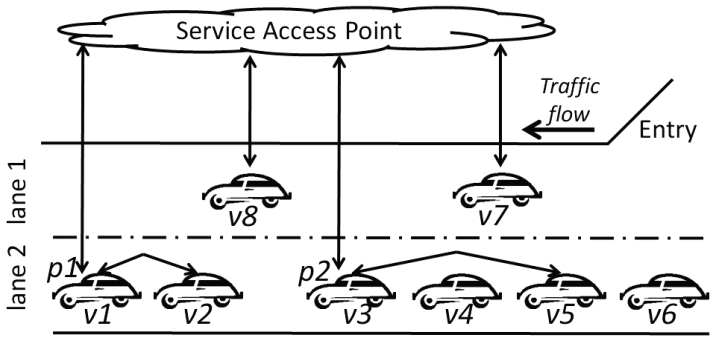

Figure 5.4: Centralized inter-platoon coordination

5.2.3.1 Inter-platoon coordination

Communications between platoons can be achieved only through the leaders, and the coordination can be centralized or decentralized.

In the *centralized* model the coordination between the leaders of neighboring platoons is performed through a centralized *Service Access Point* (SAP) that is on the road-side. The coordination between different maneuvers is achieved at the level of the SAP. Figure **5.4** presents an example considering an AHS composed of two lanes with two platoons, *p1* followed by *p2* on *lane2*, and two free-agents, *v7* and *v8*, on *lane1*. Let us assume that i) *v7* and *v8*, which just entered the highway, decide to join respectively platoons *p2* and *p1*, and ii) simultaneously and independently, vehicles *v2* and *v5*, belonging respectively to platoons *p1* and *p2*, are coordinating maneuvers to exit the AHS after passing through *lane1*. The SAP determines the priorities between the maneuvers involving the four concerned vehicles and communicates its decision to the leaders of the platoons including the concerned vehicles. The decision would be to assign the highest priority to the maneuvers requested by *v7* and *v8*, because it is important to release *lane1* as quickly as possible, so that *v2* and *v5* can leave the highway.

In the case of *decentralized inter-platoon coordination*, the decision is made by the leaders of the concerned platoons. The information related to the state of all vehicles is stored in an onboard system that contains a knowledge base of the neighborhood traffic. This coordination strategy has an impact on the implementation of some atomic maneuvers. Compared to the centralized strategy, it involves fewer vehicles in the accomplishment of some maneuvers. Let us consider as an example the case of a faulty vehicle that needs to perform a *Take Immediate Exit-Escorted*

(*TIE-E*) maneuver with the support of a neighboring platoon. If the *inter-platoon coordination* is centralized, the implementation of this maneuver involves: 1) all the vehicles in front of the faulty vehicle (including the leader) and the vehicle just behind it, and 2) the leader of the neighboring platoon. However, in the *decentralized inter-platoon coordination* strategy, only the leaders of the two platoons and the vehicles just in front and behind the faulty vehicle contribute to the maneuver. More details are provided in [Godbole *et al.* 1996, Hallé & Chaib-draa 2005].

5.2.3.2 Intra-platoon coordination

In the *centralized intra-platoon coordination* model the coordination of operation and maneuvers involving the vehicles of a platoon is centred on one vehicle: the *leader*. For example, during a split maneuver that is initiated to allow the safe exit of a faulty vehicle, three vehicles are involved: the leader, the splitter, and the vehicle following the splitter (if it exists). The faulty vehicle should announce the need to initiate this maneuver to its platoon's leader. The leader then calculates the distance and the speed to be respected by the vehicles that are involved in the maneuver, and orders the involved vehicles to change them accordingly.

In the case where the *intra-platoon coordination is decentralized*, each platoon member has knowledge of the platoon formation and can react independently, by communicating directly with other vehicles. The leader is informed of changes as it is the representative of the platoon for inter-platoon coordination. This is similar to the platoon coordination presented in [Sakaguchi *et al*. 2000], except that our model does not involve all the platoon members in the execution of a manoeuvre [Hallé 2005].

5.2.3.3 Coordination strategies

The combination of the centralized and the decentralized models for inter and intra-platoon coordination leads to the four possible strategies that are summarized in Table 5.4. The dependability models presented in the following section are aimed at analyzing the impact of such strategies on the safety of the AHS.

Table 5.4: Coordination strategies considered

Strategy	Inter-platoon model	Intra-platoon model
DD	Decentralized	Decentralized
DC	Decentralized	Centralized
CD	Centralized	Decentralized
CC	Centralized	Centralized

5.3 Safety modelling

We consider a two lane AHS with one platoon in each lane. Vehicles in each platoon can change from one platoon to the other one freely. Each platoon contains up to *N* vehicles. We model this system, taking into account the six failure modes and the associated maneuvers presented in Table 5.1, the catastrophic situations of Table 5.3 and the four coordination strategies of Table 5.4.

The measure evaluated corresponds to the probability that the modeled AHS is in one of the catastrophic situations described in Table 5.3, as a function of time (*t*). This measure is referred to as system unsafety, and is denoted by $\bar{S}(t)$.

As discussed in Section 5.2, several factors need to be considered when analyzing the impact of failures on the safety of an AHS. In particular, the success or failure of a recovery maneuver depends on the state of the adjacent vehicles contributing to the maneuver. Thus, the models should also describe some characteristics of the configuration of the platoons as well as their dynamic evolution.

In the following, we present an overview of the SAN models that are aimed at evaluating the system unsafety taking into account the considerations mentioned above. We first present an overview of the whole system model, and then we describe the submodels composing the whole model.

5.3.1 Overview of the system model

Figure 5.5 shows the overall structure of the model describing the AHS composed of two lanes. The model includes *2N* replicas of the One_vehicle sub model that are composed with three other submodels: Configuration, Dynamicity, and Severity.

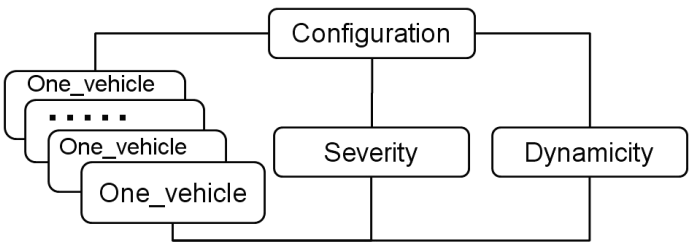

Figure 5.5: Model structure

The One_vehicle submodel describes the behaviour of a vehicle as resulting from its failure modes and the maneuvers presented in Table 5.1. The Severity submodel describes the impact of multiple failures affecting several vehicles. The sub model Dynamicity is used to model the dynamics of the system in the absence of failures, resulting from *join* and *leave* events that correspond to vehicles entering or getting out of the highway. The Configuration submodel initializes the other submodels and synchronizes their evolution according to the whole system evolution.

5.3.2 Presentation of the sub models

In the following, we detail each submodel.

5.3.2.1 One_vehicle

The SAN submodel shown in Figure 5.6 describing the vehicle behaviour models the failure modes of the vehicle and associated maneuvers, presented in Table 5.1. The model consists of six interconnected elementary SANs. Each elementary SAN models the occurrence of a failure mode for a given class of severity and the associated maneuver. An elementary SAN consists of: i) two places (CCi, SMi), ii) two input gates (fi, IGi), iii) two output gates (OGi, fmi), and iv) two timed activities (Li, *maneuver*). This model is replicated *2N* times (i.e., one model for each vehicle).

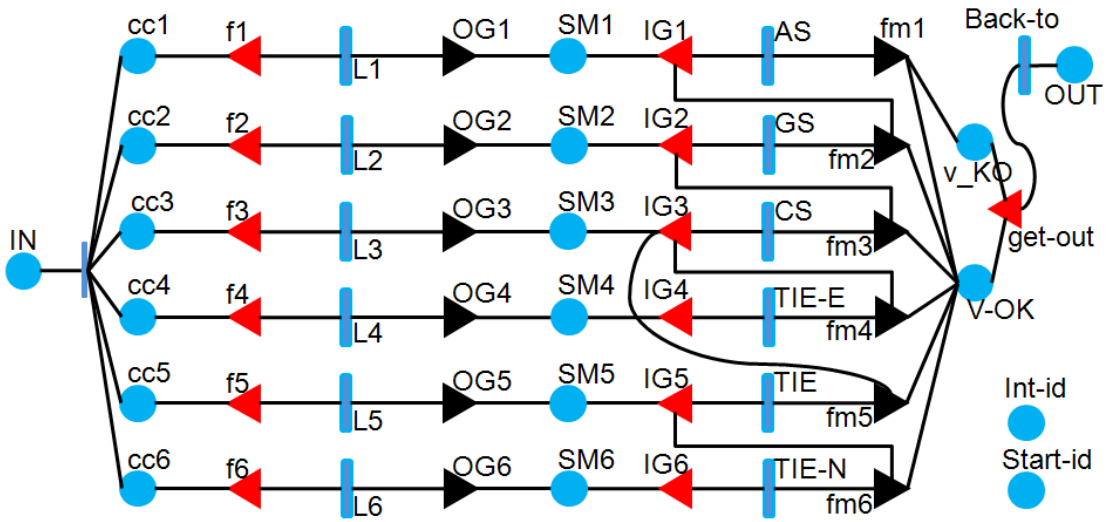

Figure 5.6: One_vehicle SAN model

Places CCi are local to each sub model. Each place CCi will have one token when a vehicle enters the platoon (i.e., place IN is marked). Place Int_id saves the ID of each vehicle in the system. Place Start_id is used for the initialization of the submodel.

Place CCi identifies the initial state from which the failure mode described by timed activity Li with firing rate λ_i could be fired. The occurrence of the failure mode activates the associated maneuver (place SMi is marked). The selection of the appropriate *maneuver* (TIE-N, TIE, TIE-E, CS, GS, or AS) depends on its priority compared to other maneuvers that might be already active, and on the state of the adjacent vehicles contributing to the maneuver. The predicates and the functions associated with the input gates IGi and the output gates fmi manage the priority of maneuvers as defined in Table 5.1 and check the marking of places SMi of the adjacent vehicles, according to the coordination strategy presented in Table 5.4. When a higher priority maneuver is activated, all lower priority maneuvers associated with the same vehicles are inhibited. The execution times of the maneuvers are described by exponentially distributed timed activities with firing rates ($\gamma_{TIE\text{-}N}$, γ_{TIE}, $\gamma_{TIE\text{-}E}$, γ_{CS}, γ_{GS}, and γ_{AS}).

If the maneuver succeeds, place v_OK is marked to indicate that the vehicle gets out of the highway safely. The maneuver failure leads the vehicle to start the next higher priority maneuver, as explained in Section 5.2.2. Eventually, if the maneuver in highest priority AS fails, v_KO is marked, and the vehicle becomes a free agent (this is not represented in the model because it will constitute a third platoon). The two existing platoons continue their way without this vehicle.

When a vehicle gets out of the platoon by reaching one of the places v_OK or v_KO, another vehicle could join the system. This is modeled through the timed activity Back_to and the marking of place OUT (see also Figure 5.8).

5.3.2.2 Severity

This submodel presented in Figure 5.7 describes the combination of failure modes affecting multiple vehicles that lead the system to an unsafe sate. Each time a failure mode Li is fired in an One_vehicle submodel, the marking of the place indicating the corresponding severity class is incremented (class_A, class_B, class_C). These extended places are shared by all the submodels. Each of them is modeled as an array listing the ongoing maneuvers with the number of failure modes of the corresponding severity class that are active during the execution of the maneuver.

The predicates and functions associated with the input gate KO_allocation and the output gate OG_KO in Figure 5.7 describe the impact on the global safety of multiple failures affecting several vehicles, as presented in Table 5.3. When the instantaneous activity to_KO is fired, the place KO_total becomes marked indicating that the system has reached an unsafe state.

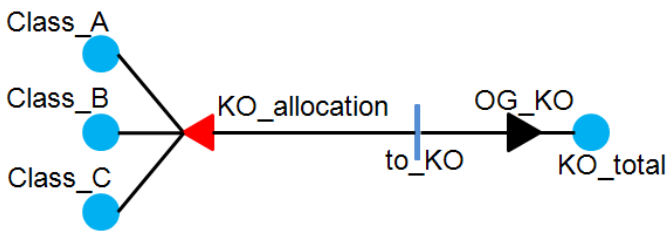

Figure 5.7: Severity SAN model

5.3.2.3 Dynamicity

The SAN submodel is given in Figure 5.8. There are four places (IN, platoon1, platoon2, and OUT). The two places platoon1 and platoon2 are shared between all submodels. They are extended places represented as an array of length *N*, each of them modeling one platoon. All these places have initially zero token.

When IN is marked, the instantaneous activity JP is fired indicating that a vehicle has joined a platoon. Two cases are associated with this activity corresponding to the selection of platoon1 or platoon2, each with probability 50%. However, other probability values could be considered to better fit the considered traffic situation.

There are five timed activities (leave1, leave2, ch1, ch2, and Join). The three activities (leave1, leave2, and Join) implement the voluntary *join* and *leave* of vehicles (i.e., in absence of failures). The other two activities (ch1 and ch2) model the time spent by a vehicle for splitting from a platoon and joining the other one.

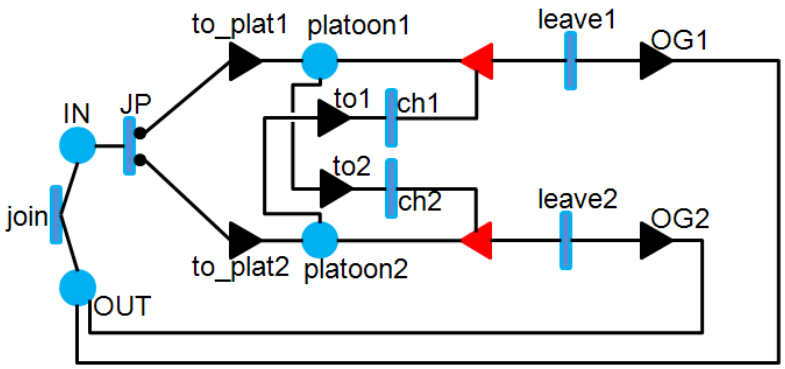

Figure 5.8: Dynamicity SAN model

When a vehicle leaves a platoon, the OUT place will be marked, thus another vehicle could join the highway. All other input and output gates are used for managing the vehicles positions after each leave and join event. In addition, each time a vehicle joins a platoon; it occupies the last position of the platoon.

5.3.2.4 Configuration

This submodel, presented in Figure 5.9, is used to define the initial configuration of the platoons and to initialize the One_vehicle submodels associated with each vehicle included in the platoons. Each platoon can contain up to *N* vehicles. Thus the system model is composed of *2N* replicas of the One_vehicle submodel.

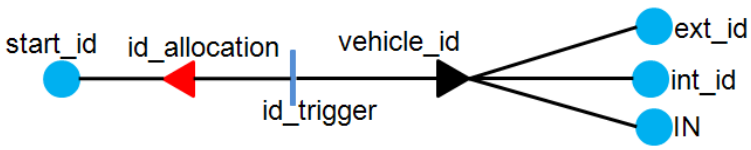

Figure 5.9: Configuration SAN model

The Configuration submodel contains four places; all of them have initially zero token except Start_id which has one token. Places (Start_id, Int_id, and IN) are shared with the corresponding One_vehicle submodel replicas included in the configuration of the AHS. Initially *2N* replicas are created, *N* vehicles for each platoon. The place ext_id is a global place shared by all sub models, to act as a counter. Each time the instantaneous activity id_trigger is fired; a new vehicle is included in the system and assigned a vehicle_id. Also place IN is marked to initialize: i) the One_vehicle submodel associated with this vehicle, and ii) the Dynamicity submodel that will associate the vehicle to a given platoon. The ID assigned to the vehicle is stored in the place

Int_id. When a new vehicle joins the system, Int_id gets the value stored in Ext_id, which in turn is incremented by one.

5.3.2.5 SAN system composed model

The system SAN model resulting from the composition of the SAN submodels presented in the previous sub-sections, using "Join" and "Rep" composition operators, is illustrated in Figure 5.10.

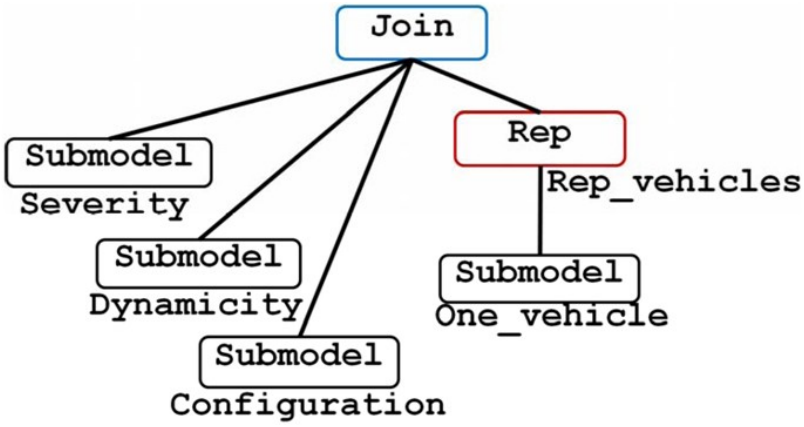

Figure 5.10: SAN composed model

5.4 Results and sensitivity analyses

We illustrate the type of results obtained from the processing of the SAN model presented in Section 5.3, and show sensitivity analyses with respect to various parameters impacting the AHS safety.

The unsafety measure $\bar{S}(t)$ defined in Section 5.3 corresponds to the probability to have a token in the place KO_total of Figure 5.7. The analyses focus on the impact on $\bar{S}(t)$ of the failure rates associated with the failure modes, the maximum number of vehicles per platoon, the trip duration, and the AHS coordination strategies.

5.4.1 Assumptions and numerical values of the parameters

We assume that all the processes represented by timed activities in the SAN models have exponential distributions (i.e., have constant occurrence rates).

Let λ be the smallest failure rate. To facilitate sensitivity analyses, the values of the failure rates λ_i associated with the six failure modes FMi identified in Table 5.1 are expressed in terms of λ. In this case study considering the contribution of all sources of failures that can lead to the considered failure mode, we have used the following values:

$\lambda_6 = 4\lambda$; $\lambda_5 = 3\lambda$; $\lambda_4 = 2\lambda$; $\lambda_3 = 2\lambda$; $\lambda_2 = 2\lambda$; $\lambda_1 = \lambda$.

The values of execution rates associated with the *maneuvers* ($\gamma_{TIE\text{-}N}$, γ_{TIE}, $\gamma_{TIE\text{-}E}$, γ_{CS}, γ_{GS}, γ_{AS}) range from 15/hr and 30/hr (maneuver durations between 4 and 2 minutes).

We suppose that the two highway lanes start initially with *N* vehicles in each platoon (platoon1 and platoon2). At any time each vehicle can change from its platoon to the other one, with constant change rates (respectively ch1 and ch2 for platoon1 and platoon2 as shown in the Dynamicity submodel of Figure 5.8). We consider the same numerical values for the two change rates equal to 6/hr.

The numerical values used are inspired from real life similar situations. However, these values can be easily modified.

Each vehicle in platoon2 leaving the highway should pass through platoon1 and stay 3 to 4 minutes in platoon1, before getting out from the highway.

The results presented in the following subsections have been obtained, using the simulator provided by the Möbius tool. Each point of the graphs has been computed as a mean of at least 10000 simulation batches, converging within 95% probability in a 0.1 relative interval. Actually, the total number of simulation batches mainly depends on the value of the failure rate considered.

5.4.2 Failure rate and number of vehicles impact

We first show in Figure 5.11 the impact of *N*, the maximum number of vehicles per platoon on $\bar{S}(t)$, for trip durations varying from 2 to 10 hours.

This figure shows that:

- For a given value *N*, the probability of reaching the unsafe state increases by one order of magnitude when the trip duration increases from 2 to 10 hours.
- For a given trip duration, increasing *N* leads to a significant increase of $\bar{S}(t)$. For example, when *n* is increased from 8 to 12, the unsafety is one order of magnitude higher, for a 10 hours trip duration.

For a failure rate equal to 10^{-5}/hr, the level of unsafety remains low when *N* is less than 10. Higher values of *N* lead to a more degraded safety especially when considering long trip durations.

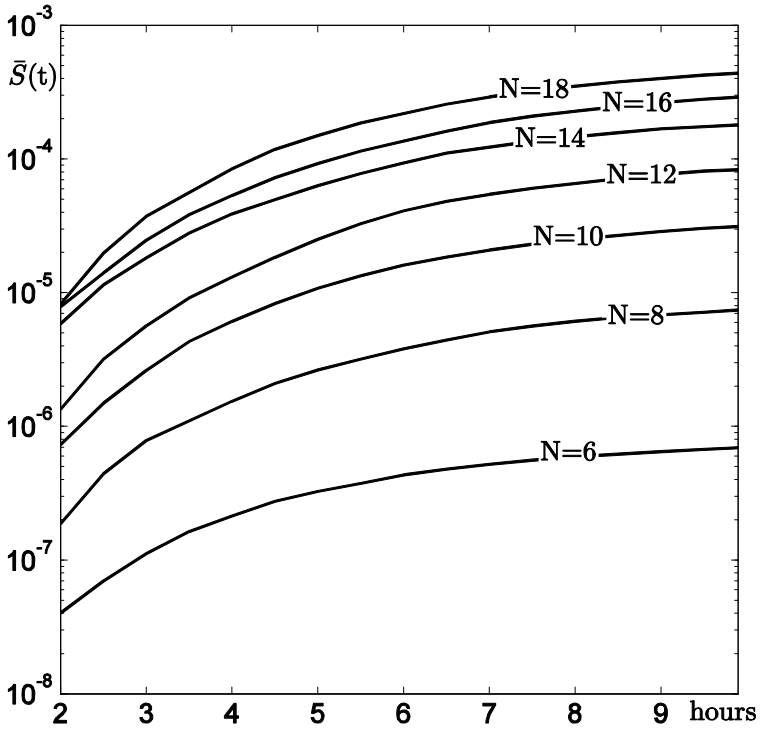

Figure 5.11: $\bar{S}(t)$ for different platoon lengths, *N* (*join* rate=12/hr, and *leave* rate=4/hr)

The impact of the failure rate is illustrated in Figure 5.12 considering three values for λ. We notice that the probability of reaching an unsafe state is very sensitive to the value of the failure rate. For example, increasing the failure rate from 10^{-6}/hr to 10^{-5}/hr, leads to an increase of unsafety of about 175 times, for a trip duration of 6 hours. The variation of system unsafety is lower (about 40 times) when increasing the failure rate from 10^{-5}/hr to 10^{-4}/hr for the same trip duration. Additionally, it can be noticed that the sensitivity of $\bar{S}(t)$ to the trip duration is higher for lower values of the failure rate λ. For $\lambda = 10^{-4}$/hr the steady state is reached very quickly.

For a 2 hours trip duration, the unsafety is almost 10^{-12}/hr for $\lambda = 10^{-6}$/hr. This is why the corresponding result is not plotted.

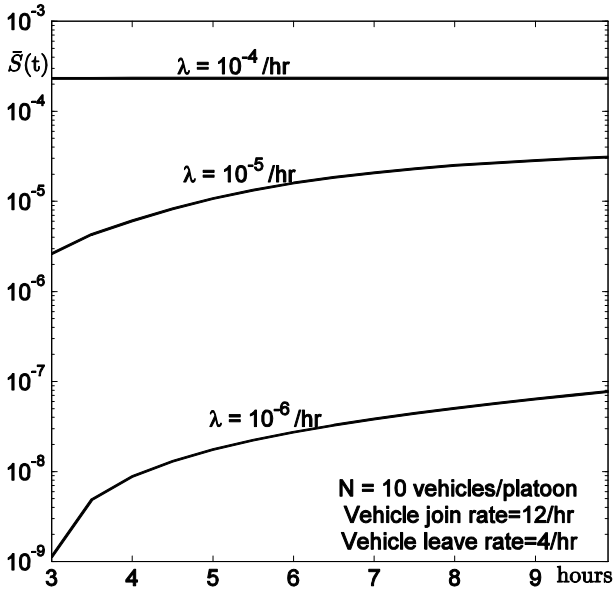

Figure 5.12: $\bar{S}(t)$ **for various failure rates,** λ

Figure 5.13 shows the impact of the failure rate on system unsafety when the maximum number of vehicles per platoon, *N*, increases from 10 to 18, considering 6 hour trip duration. We can see that the failure rate has more impact for smaller number of vehicles per platoon.

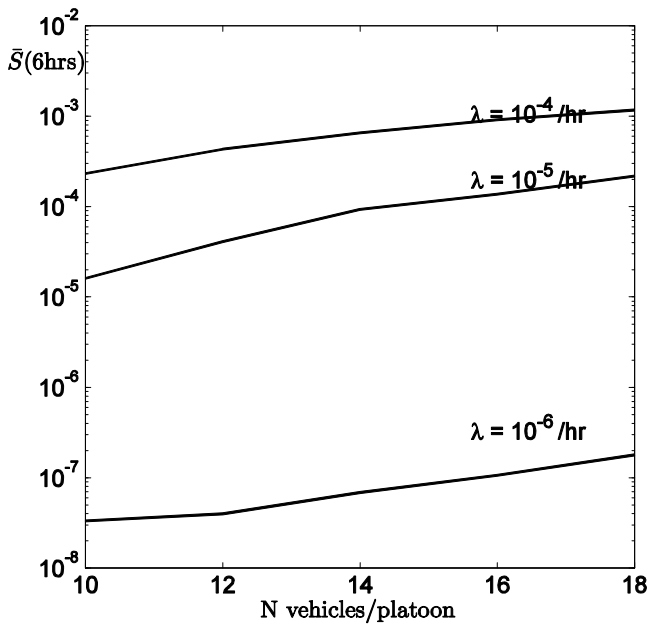

Figure 5.13: $\bar{S}(t)$ **at *t*=6 hrs versus *N* for various** λ
(*join* rate=12/hr, and *leave* rate=4/hr)

5.4.3 Influence of *leave* and *join* rates

Actually, the system unsafety should depend on the number of vehicles in each platoon that might be affected by failures. The number of vehicles depends on the frequency at which vehicles join and leave the platoon. In order to have a better understanding of the combined influence of the *join* and *leave* rates, we analyze the evolution of system unsafety as a function of the load of the system $\rho = \frac{join\_rate}{leave\_rate}$.

The results are plotted in Figure 5.14 for different values for the *join* and *leave* rates. It is interesting to see that similar trends are observed for all the curves, with the highest unsafety observed for the highest join rate.

Comparison of the results corresponding to different values of ρ and a fixed value of the leave rate shows that the highest value ρ leads to the highest level of unsafety. However, the results are of the same order of magnitude.

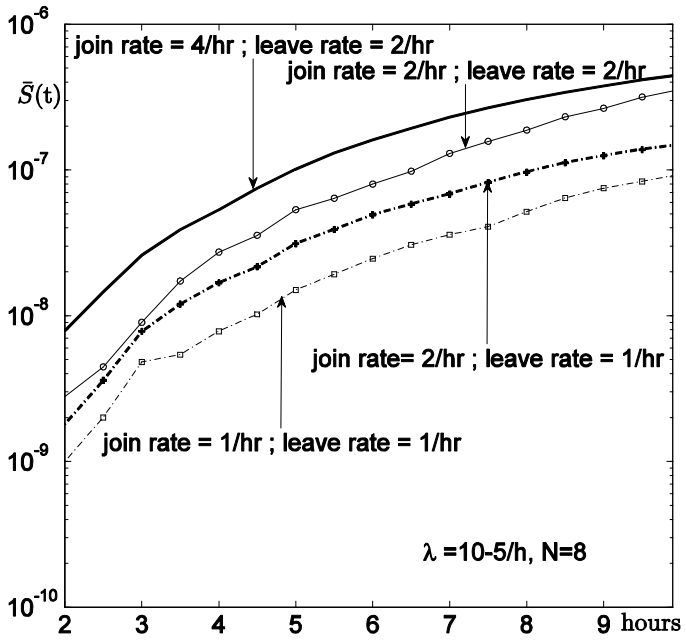

Figure 5.14: $\bar{S}(t)$ versus trip duration, for various *join* and *leave* rates

5.4.4 Influence of coordination strategy

All the results presented in Sections 5.4.2 and 5.4.3, correspond to the case of a *decentralized inter-* and *intra-platoon coordination* strategy (DD). Figure 5.15 compares the unsafety for the four strategies presented in Table 5.4: DD (*Decentralized inter-* and *intra-platoon*) DC (*Decentralized inter-platoon* and *Centralized intra-platoon*), CD (*Centralized inter-platoon* and *Decentralized intra-platoon*), and CC (*Centralized inter-* and *intra-platoon*). We can see that the *inter-platoon* strategy has more impact than the *intra-platoon*, with a higher safety observed for the *decentralized inter-platoon* strategy. This is due to the fact that more vehicles are involved in the *centralized inter-platoon coordination* (see Section 5.2.3.1).

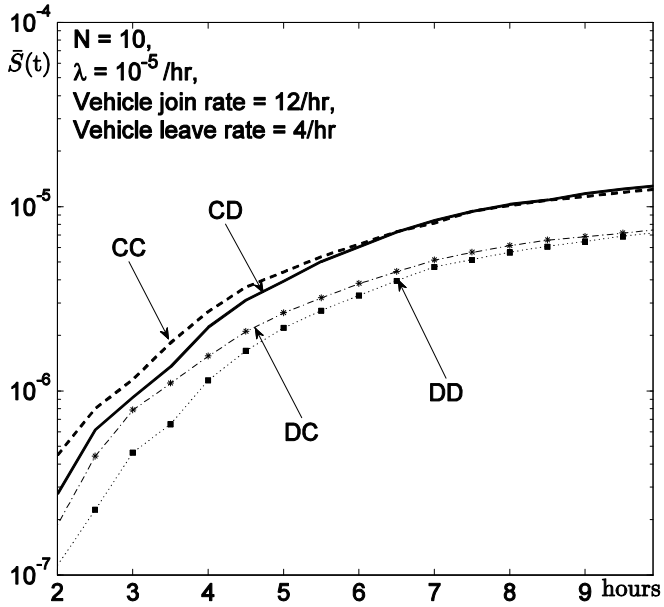

Figure 5.15: $\bar{S}(t)$ and coordination strategy

The impact of the coordination strategy is low even for higher values of *N*. This is shown in Figure 5.16 where the system unsafety at *t = 6hrs* is plotted for different values of *N*.

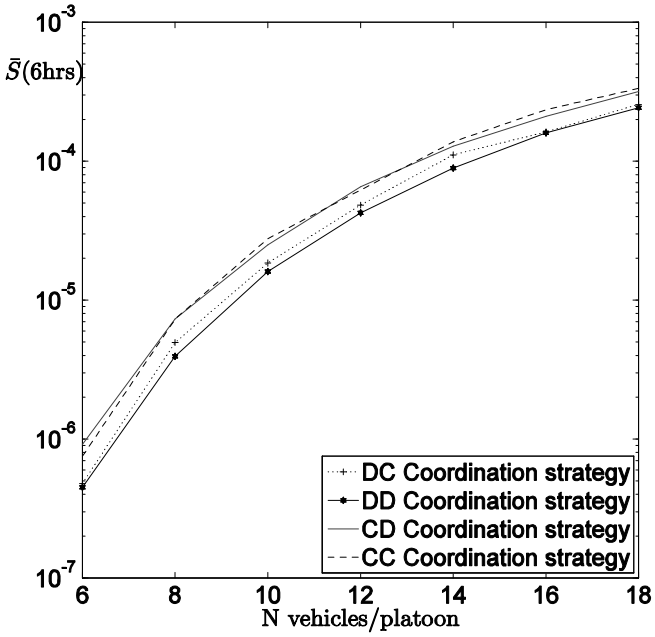

Figure 5.16: $\bar{S}(t)$ at *t*=6hrs versus platoon length *N*, λ=10^{-5}/hr, *join* rate=12/hr, *leave* rate=4/hr

5.5 Conclusion

In this chapter, we have addressed the safety modeling and evaluation of an automated highway system. The models take into account the failure modes affecting vehicles, their severity level, and their associated recovery maneuvers. The modeling approach has been designed to master the complexity of the models taking into account the dynamic evolution of the highway system. The proposed models are based on stochastic activity networks. The system model is elaborated based on submodels characterizing the vehicles' behaviour resulting from failures and recovery maneuvers, that are then replicated and composed with other submodels describing the system configuration and its dynamic evolution as the result of vehicles joining and leaving the highway.

To illustrate the feasibility of the approach and the kind of results that can be achieved, we considered the case of a highway composed of two platoons. We performed sensitivity studies to analyze the impact of several parameters on the safety of an automated highway system: the failure rates associated with failure modes affecting vehicles, the maximum number of vehicles per platoon, and different coordination strategies. In particular, the analyses we made have allowed us to quantify and perform a comparative analysis of the level of safety that can be expected with the system studied for different configurations and parameters ranges.

The work and results presented in this chapter can be considered as a preliminary step in addressing the safety evaluation of automated highway systems. Nevertheless, the results already allowed us to quantify safety and to perform a comparative analysis with some preliminary indication with the following factors: 1) the optimal size of platoons; 2) the maximum trip duration; 3) the most suitable coordination strategy of the platoons that lead to better safety. The

first and third factors could be set by the user. Future work is needed to analyze how to control these factors in an operational context to optimize safety. For the parameters considered in our study, the size of the platoons should not exceed 10 which are consistent with the numbers considered in experimental tests, as reported in [Miller 1997] for example.

The models presented for this case study can be extended to analyze highways composed of a larger number of platoons, considering more complex scenarios. Also, further work is planned to evaluate other collaborative driving systems using e.g., the concept of teamwork for platoon formations [Hallé & Chaib-draa 2005].

Conclusions and Perspectives

"*People do not like to think. If one thinks, one must reach conclusions. Conclusions are not always pleasant.*"
Helen Keller, US blind & deaf educator, 1880-1968

"*A conclusion is the place where you got tired of thinking.*"
Arthur Bloch, American writer, 1948

Dependability of mobile-based services has become a hot topic in the automotive domain in the last recent years. As long as *ad-hoc* wireless networks, Bluetooth, radio communication and other embedded technology will enable our vehicles to sense traffic and stop accidents before they happen, the use of such wireless and mobile technologies to share and produce data for entertainment as well as to improve traffic safety will become widespread. Ultimately, users expect such services to achieve acceptable levels of dependability.

This dissertation focused on the development of modelling approaches aimed at the dependability evaluation of vehicular applications using a combination of Car-to-Car and Car-to-Infrastructure wireless and mobile communications. To the best of our knowledge, this problem has been seldom addressed and there is a lack of practical examples illustrating how dependability evaluation in this context can be carried out. Three case studies are presented namely, a replication service in the *ad-hoc* domain, a virtual black-box application, and a platooning application. They allowed us to illustrate various challenges that are typical of vehicular applications using C2C and C2I communications and the modelling of dependability characteristics taking into account different mobility scenarios and design parameters. We mainly took into account the impact of accidental threats on the delivered services. Several dependability properties and quantitative measures have been addressed depending on the case study under investigation, with a particular focus on *availability* and *safety*.

Contributions

The main contributions of the work presented in this dissertation are summarized in the following:

1) We presented a holistic evaluation approach that has been developed in the context of the HIDENETS European project to evaluate an integrated way: i) the failure modes affecting the vehicles, their severity level, and the associated recovery mechanisms, and ii) the dynamics of the systems resulting from their mobility. This approach has been illustrated providing examples of interaction between mobility models, simulators and analytical models. Two of three case studies presented are examples of the cross fertilization among different methods, since we feed system dependability models with parameter distributions derived through mobility models simulations.

2) We analyzed the distributions characterizing C2C and C2I connectivity under different mobility scenarios including freeways and urban traffic environments. Various techniques have been investigated including simulation, analytical connectivity models and statistical processing of real life mobility traces. We have focused on the analysis of the rate at which vehicles meet, the duration that such vehicles stay in connectivity range, and the rate at which a vehicle meets an access point to the fixed infrastructure. Such analysis concluded that the exponential distribution and the Pareto distribution are adequate for modelling the time

between C2C encounters in freeways and urban mobility environments, respectively. These distributions have been used as an input for the dependability modelling of the case studies investigated in the thesis.

3) We proposed a dependability model for the three different case studies based on Stochastic Activity Networks, following a similar approach for the dependability evaluation of each of these case studies. The modelling approach has been designed to: 1) master complexity by taking advantage of the compositional operators offered by the SAN formalism, and 2) to analyze and assess the impact of mobility. For two of these cases studies (the replication service and the virtual black box), models are developed taking into account the information derived from the connectivity analysis of mobility scenarios. For the automated highway system, the mobility of the vehicles was analyzed directly in the dependability model in a more abstract way by considering the evolution of the configuration of the platoons resulting from vehicles leaving or joining a platoon.

4) We focused on specific dependability attributes depending on the characteristics of each case study. For the replication service, we characterized the service *availability* and the data *consistency* among the replica group. For the virtual black box application, we analysed and evaluated the *availability* of the historical data recorded in the virtual black-box. For the platooning application, we addressed the assessment of traffic *safety*.

5) The sensitivity analyses allowed us to identify scenarios providing the most beneficial schemes from the dependability point of view, considering the influence of different design and environmental parameters:

    a. For the replication service and the virtual black box applications, we have shown that the geographic mobility scenarios represented by the distribution of the time between C2C encounters (exponential or Pareto distribution) significantly impacts the dependability related properties.

    b. For the replication service application, we analyzed the impact of the rate at which the application state changes on consistency and availability related properties.

    c. For the virtual black-box, a comparison of replication and duplication strategies based on erasure codes and simple replication techniques was carried out, considering the case of a single data record and multiple data records.

    d. For the platooning application, the dependability model takes into account the configuration of the platoons and their dynamic evolution and describes different failure modes affecting the vehicles and associated recovery manoeuvres considering various possible coordination strategies between the vehicles, inside each platoon and between platoons.

In addition, the sensitivity analysis results allowed us to illustrate how the designers can draw some practical conclusions from the models when considering the service *unavailability*/*unsafety* under different mobility scenarios. For example, in the case of the replication service, the replica group size and the number of replicas could be chosen according to the *availability* level to be achieved. Similar analyzes can be done in the context of the virtual black box application to chose the appropriate data replication strategy and parameters associated with an erasure coding

algorithm. As regards the platooning application, the results presented in the dissertation allowed us to illustrate the impact on safety of the coordination strategy for different trip durations and platoon sizes.

Finally, it is worthwhile to note that the modelling approach illustrated through the three different case studies has a general scope and applicability and can be followed in other application areas using wireless communication technologies and mobile ad-hoc networks.

Upcoming research work

The contributions of the work presented in this dissertation towards the dependability modelling and evaluation of mobile-based applications could be extended in several ways. In the following, we provide insight into possible research directions that would complement our work.

Let us first note that the models presented in this dissertation can be extended to analyze more complex use cases and mobility scenarios. For example, the models can be extended and refined to take into account failures occurring at the networking side and the user/driver side to be able to provide more comprehensive *availability*/*safety* estimations. A more detailed case study is also needed to illustrate the different levels of the holistic evaluation approach presented in Chapter 1: i.e., the user, application, architecture and communication levels. In this context, the combination of different types of formalisms and evaluation techniques including analytical models, simulation and experimental measurements might be needed in order to provide detailed evaluations of end-to-end scenarios taking into account different decomposition and abstraction levels of the system.

Another challenging problem that needs future research concerns the development of mobility models that are representative of realistic mobility scenarios, for vehicular applications and also for other application areas. The sensitivity analysis results presented in the dissertation have shown that the distribution of the time between C2C encounters can have a significant impact on dependability. Ongoing research initiatives aiming at the development of micro-simulations of vehicular traffic scenarios or the collection of real-life traces in different mobility environments should provide interesting inputs for developing such mobility models.

One important aspect to be further investigated in the future concerns the development of an evaluation platform integrating different tools and simulators that can be combined and chained automatically. The objective would be to define well specified interfaces and transformations in such way that the output of a given tool can be used automatically as an input to other evaluation tools providing quantitative measures characterizing the behaviour of the systems from a different perspective or at a different abstraction level. A preliminary evaluation workflow going in this direction was proposed in the context of the Hidenets project.

Finally, in this dissertation, we analyzed some dependability attributes taking into account the impact of accidental threats only. The impact of malicious threats should also be investigated in future work in order to provide more realistic estimation of dependability attributes and to help the designers in making optimal tradeoffs between availability, safety and security related properties based on quantitative measures.

Appendix A: Background on SAN Modeling and Möbius

This appendix presents the main principles behind the Stochastic Activity Networks (SAN) formalism and the associated Möbius tool.

A.1. Möbius overview

Möbius is a software tool for modeling the behavior of *complex systems*. The first step in the model construction process is to generate a SAN model (cf. summary in next section or more details in [Sanders & Meyer 2001]). The most basic model in the framework is called an *atomic model*, and is made up of *state variables* and *actions*: state variables hold state information about a model, while actions provide the mechanism for changing model states- so- called activities in SAN.

If the model being constructed is intended to be part of a larger model, then the next step is to compose it with other models (i.e., atomic or composed models) to form a larger model. This is sometimes used as a convenient technique to make the model modular and easier to construct. Although a composed model is a single model with its own state space, it is not a "flat" model. It is hierarchically built from submodels, which largely preserve their formalism-specific characteristics so that the composed model does not destroy the structural properties of the submodels.

After a composed model is created, the next step is to specify some measures of interest on the model using some *reward* specification formalism: the Möbius tool captures this pattern by having a separate model type, called reward models, that augments composed models with reward variables.

The next step is typically to apply some *solver* to compute a solution to the reward model: a solver is mechanism that calculates the solution to reward variables a solver. The computed solution to a reward variable is called a *result*: since the reward variable is a random variable, the result is expressed as some characteristic of a random variable (this may be, for example, the mean, variance, or distribution of the reward variable).

A.2. Atomic SAN models

This section contains a brief description of the SAN primitive objects: *places*, *activities*, *input gates*, and *output gates*. These objects and their usage are illustrated in each case study modelling section.

Places represent the state of the modeled system; they are represented graphically as circles. Each place contains a number of tokens, that represents the marking of the place. Note that the tokens in a place are homogeneous, in that only the number of tokens in a place is known; there is no identification of different kinds of tokens within a place.

Activities represent actions in the modeled system that take some specified amount of time to complete. There are two types of activities: timed and instantaneous. Timed activities have a duration that impacts the performance of the modeled system (such as a communication delay or

the time associated with a retransmission timer). They are represented graphically as thick vertical lines. Activity time distribution functions can be generally distributed random variables, where each distribution can depend on the marking of the network. Instantaneous activities represent actions that complete immediately when enabled in the system. They are represented as thin vertical lines.

Case probabilities, represented graphically as circles on the right side of an activity, model uncertainty associated with the completion of an activity; each case stands for a possible outcome, such as a routing choice in a network, or a failure mode in a faulty system. Each activity has a probability distribution, called the "case distribution", associated with its case; this distribution can depend on the marking of the network at the moment of completion of an activity. If no circles are shown on an activity, one case is assumed with a probability of one. Each activity has also a reactivation functions. This function gives marking dependent conditions under which an activity is reactivated. Reactivation of an activated activity means that the activity is aborted and that a new activity time is immediately obtained from the activity time distribution.

Input gates control the enabling of activities and define the marking changes that will occur when an activity completes. Input gates are represented graphically as triangles; an arc is connected the controlled activity, other arcs are connected to the places upon which the gate depends, also called input places. Each input gate is defined with an enabling predicate and function. The enabling predicate is a Boolean function that controls whether the connected activity is enabled; it can be any function of the markings of the input places. The input gates function defines the marking changes that occur when the activity completes. If a place is directly connected to an activity with an arc, it is equivalent to an input gate with a predicate that enables the activity whenever the place has more than zero tokens along with a function that decrements the marking of the place whenever the activity fires.

Output gates define the marking changes that will occur when activities complete. The only difference between input and output gates is that an output gate is associated with a single case of the activity. An output gate is represented graphically as a triangle with its flat side connected to an activity or a case of an activity; on the other side of the triangle is a set of arcs to the places affected by the marking changes. An output gate is defined only with a function: the function defines the marking changes that occur when the activity completes. There is also a default scenario for output gates. If an activity is directly connected to a place, it is equivalent to an activation in which an output gate has a function that increments the marking of the place whenever the activity is fired.

A.3. Composed models

The Möbius framework allows the construction of composed models from previously defined models, which allows the modeler to adopt a hierarchical approach to modeling, by constructing submodels as meaningful units and then placing them together in a well-defined manner to construct a model of a system.

Model composition is accomplished by the state-sharing approach, which links submodels together by identifying sets of state variables. For example, it is possible to compose two SAN models by holding a particular place in common. That allows for interaction between the submodels, since both can read from and write to the identified state variable. This form of state-

sharing is known as equivalence sharing, since both submodels have the same relationship to the shared state variable.

The composed model formalism used by Möbius for SAN models is *Replication*/*Join*: this formalism permits to define a composed model in the form of a tree, in which each leaf node is a predefined atomic or composed model, and each non-leaf node is classified as either a Join node or a Replicate node. A *Join* is used to compose two or more submodels using equivalence sharing. A *Replicate* is used to construct a model consisting of a number of identical copies of its single child. Each child node of a Replicate or Join node can be a Replicate, a Join, or a single atomic or composed model. Replicate is used to construct a model consisting of a number of identical (indistinguishable) copies of its single child.

A.4. Reward models

Reward models build upon atomic and composed models, equipping them with the specification of a performance measures. Möbius implements a reward model called a *performance variable*, which allows for the specification of a measure on one or both of the following:

- The states of the model, giving a rate reward *performance variable*.
- Action completions, giving an impulse reward *performance variable*.

A rate reward is a function of the state of the system at an instant of time. An impulse reward is a function of the state of the system and the identity of an action that completes, and is evaluated when that particular action completes; a *performance variable* can be specified to be measured at an instant of time, to be measured in steady state, to be accumulated over a period of time, or to be time-averaged over a period of time. Once the rate and impulse rewards are defined, the desired statistics on the measure must be specified. Möbius includes solving for the mean, variance, or distribution of the measure, or for the probability that the measure will fall within a specified range.

A.5. Solver

Möbius supports two classes' solution techniques: discrete event simulation and state-based, analytical/numerical techniques. Any model specified using Möbius may be solved using simulation, whilst only models having delays that are all exponentially distributed, or having no more than one concurrently enabled deterministic delay, may be solved using a variety of analytic techniques applied to a generated state space.

Möbius simulation supports two modes of discrete event simulation: transient and steady-state. In the transient mode, the simulator uses the independent replication technique to obtain statistical information about the specified reward variables. In the steady-state mode, the simulator uses batch means with deletion of an initial transient period to solve for steady-state, instant-of-time variables. Estimates available during simulation include mean, variance, interval, and distributions. Confidence intervals are computed for all estimates.

References

[Artimy *et al.* 2004] M. M. Artimy, W. Robertson and W. J. Phillips, "*Connectivity in inter-vehicle ad-hoc networks*", In Electrical and Computer Engineering, 2004. Canadian Conference, pp.293-298, 2004.

[Avizienis *et al.* 2004] A. Avizienis, J.-C. Laprie, B. Randell and C. Landwehr, "*Basic concepts and taxonomy of dependable and secure computing*", In IEEE Transactions on Dependable and Secure Computing, vol. 1, pp.11-33, 2004.

[Bai & Krishnamachari 2009] F. Bai and B. Krishnamachari, "*Spatio-temporal variations of vehicle traffic in VANETs: facts and implications*", Proceedings of the sixth ACM international workshop on VehiculAr InterNETworking, International Conference on Mobile Computing and Networking, Beijing, China, pp.43-52, 2009.

[Bai *et al.* 2003] F. Bai, N. Sadagopan and A. Helmy, "*The IMPORTANT framework for analyzing the Impact of Mobility on Performance Of RouTing protocols for Ad-hoc NeTworks*,", IEEE Information Communication Communications Conference (INFOCOM 2003), IEEE CS Press, San Francisco, CA, USA, vol. 2, pp.825-835, 2003.

[Balakrishnan & Trivedi 1995] M. Balakrishnan and K. S. Trivedi, "*Component-Wise Decomposition for an Efficent Reliability Solution of Complex Models of System with Repairable Components*", International Symposium on Failt-Tolerant Computing (FTCS), pp.259-268, 1995.

[Balbo *et al.* 1988] G. Balbo, S. Bruell and S. Ghanta, "*Combining Queuing Networks and GSPNs for the Solution of Complex Models of System Behavior*", IEEE Transaction on Computers, vol. 37, pp.1251-1268, 1988.

[Basile *et al.* 2002] C. Basile, M.-O. Killijian and D. Powell, "*A Survey of Dependability Issues in Mobile Wireless Networks*", LAAS-CNRS, report N°. 02637 2002.

[Baskett *et al.* 1975] F. Baskett, K. Chandy, R. Muntz and F. Palacios, "*Open, Closed, and Mixed Networks of Queues with Different Classes of Customers*", Journal of ACM, vol. 22 (2) 1975.

[Benz *et al.* 1996] T. Benz, A. Braun, R. Krause, Pochmuller, W. H. Schulz, M. Schulze and J. Sonntag, "*CHAUFFEUR - TR 1009 User, safety and operational requirements*", Project Deliverable D3.1.1 1996.

[Béounes *et al.* 1993] C. Béounes, M. Aguéra, J. Arlat, S. Bachmann, C. Bordeau, E. Doucet, K. Kanoun, J.-C. Laprie, S. Metge, J. de Souza, D. Powell and P. Spiesser, "*SURF-2: A Program for Dependability Evaluation of Complex Hardware and Software Systems*", IEEE Computer Society Press, 23th IEEE International Symposium on Fault-Tolerant Computing, pp.668-673, 1993.

[Berson *et al.* 1991] S. Berson, E. de Souza e Silva and R. Muntz, "*A Methodology for the Specification and Generation of Markov Models*", Numerical Solution of Markov Chains, pp.11-36, 1991.

[Betous-Almeida & Kanoun 2004] C. Betous-Almeida and K. Kanoun, "*Construction and Stepwise Refinement of Dependability Models*", Performance Evaluation, vol. 1 (56), pp.277-306, 2004.

[Blywis *et al.* 2009] B. Blywis, F. Juraschek, M. Günes and C. Graff, "*MoNoTrac: A Mobile Node Trace Generator*", Tech. report at sTechnische Universität Hamburg-Harburg, p.3, 2009.

[Bobbio & Trivedi 1986] A. Bobbio and K. S. Trivedi, "*An Aggregation Technique for the Transient Analysis of Stiff Markov Chains*", IEEE Transaction on Computers, vol. 9 (35), pp.803-814, 1986.

[Bolch *et al.* 2006] G. Bolch, S. Greiner, d.-H. Meer and K. S. Trivedi, *Queueing Networks and Markov Chains*, Print ISBN: 9780471565253, Wiley-InterScience, 2006.

[Bondavalli *et al.* 2010] A. Bondavalli, P. Lollini and L. Montecchi, "*QoS Perceived by Users of Ubiquitous UMTS: Compositional Models and Thorough Analysis*", Journal of Software, Special issue on Selected Papers of the 6th IFIP Workshop on Software Technologies for Future Embedded and Ubiquitous Systems, vol. 1, 2010.

[Bondavalli *et al.* 2000] A. Bondavalli, I. Mura, S. Chiaradonna, R. Filippini, S. Poli and F. Sandrini, "*DEEM: a Tool for the Dependability Modeling and Evaluation of Multiple Phased Systems*", The 32nd Annual IEEE/IFIP international conference on Dependable Systems and Networks (DSN-02), pp.231-236, 2000.

[Bondavalli *et al.* 1999] A. Bondavalli, I. Mura and K. S. Trivedi, "*Dependability Modeling and Sensitivity Analysis of Scheduled Maintenance Systems*", European Dependable Computing Conference EDCC-3, vol. 1667, pp.7-23, 1999.

[Bozinovski *et al.* 2004] M. Bozinovski, H.-P. Schwefel and R. Prasad, "*Algorithm for Controlling Transaction Consistency in SIP Session Control Systems*", IEE Electronics Letters, vol. 40 (3), pp.209-211, February 2004 2004.

[Burgess et al. 2006] J. Burgess, B. Gallagher, D. Jensen and B. N. Levine, "*MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks*", INFOCOM 2006. 25th IEEE International Conference on Computer Communications, Barcelona, Gatalunya, Spain, pp.1-11, 2006.

[Burggraf *et al.* 2007] F. Burggraf, W. N. Carey, P. Johnson and K. B. Woods, *Intelligent Transportation Systems and Vehicle-Highway Automation*, Published by Transportation Research Board of the National Academies, ISSN: 0361-1981, 2007.

[Buzen 1973] J. P. Buzen, "*Computational Algorithms for Closed Queuing Networks with Exponential Servers*", Communications ACM, vol. 16 (9) 1973.

[Bychkovsky *et al.* 2006] V. Bychkovsky, B. Hull, A. Miu, M. Balakrishnan and S. Madden, "*A measurement study of vehicular internet access using in situ Wi-Fi networks*", Proceedings of the

12th Annual International Conference on Mobile Computing and Networking, MOBICOM 2006, Los Angeles, CA, USA, pp.50-61, 2006.

[C2CC 2007] C2CC, "*List of Applications (BMW, DaimlerChrysler, Volkswagen*", Project in Germany 2007.

[C2CCC 2007] C2CCC, "*Car 2 car communication consortium manifesto, overview of the c2c-cc system*", http://www.car-to-car.org/fileadmin/downloads/C2C-CC_manifesto_v1.1.pdf, version 1.1 2007.

[CAMP 2003] CAMP, "*Vehicle Safety Communications Project, Task 3: Identify Intelligent Vehicle Safety Applications Enabled by DSRC*", Interim Report 2003.

[Camp *et al.* 2002] T. Camp, J. Boleng and V. Davies, "*A survey of mobility models for ad-hoc network research*", in Wireless Communications and Mobile Computing (WCMC), vol. 2 (5), p.483–502, 2002.

[Casimiro Costa *et al.* 2007] A. Casimiro Costa, A. Bondavalli, M. Calha, M. Clemesten, A. Daidone, M. Dixit, Z. Ègel, L. Falai, F. Di Giandomenico, A. F. Hansen, G. Huszerl, A. Kövi, M.-O. Killijian, T. Lippmann, Y. Liu, E. Matthiesen Moller, H. Moniz, A. Nickelsen, J. Nielsen, T. Renier, M. Roy, J. Rufino, S. H.-P. and E. Svinnset, "*Resilient Architecture*", EU FP6-IST Project HIDENETS, Deliverable D2.1.2 2007.

[Chaintreau *et al.* 2005] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass and J. Scott, "*Pocket Switched Networks: Real-world mobility and its consequences for opportunistic forwarding*", Technical Report, UCAM-CL-TR-617, University of Cambridge, UK, p.26, 2005.

[Chaintreau *et al.* 2007] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass and J. Scott, "*Impact of Human Mobility on Opportunistic Forwarding Algorithms*", IEEE Trans.on Mobile Computing, vol. 6 (6), pp.606-620, 2007.

[Chen *et al.* 2005] I.-R. Chen, G. Baoshan, S. George and C. Sheng-Tzong, "*On failure recoverability of client-server applications in mobile wireless environments*", in Reliability, IEEE Transactions, vol. 54, pp.115-122, 2005.

[Chen *et al.* 2001] Z. D. Chen, H. Kung and D. Vlah, "*Ad-hoc relay wireless networks over moving vehicles on highways*", In MobiHoc01: Proceedings of the 2nd ACM international symposium on Mobile ad-hoc networking & computing, New York, NY, USA. ACM Press. (New York, NY, USA. ACM Press.), pp.247-250, 2001.

[Cheng & Robertazzi 1989] Y. Cheng and T. G. Robertazzi, "*Critical connectivity phenomena in multihop radio models*", IEEE Transactions on Communications, vol. 37 (7), pp.770-777, 1989.

[Chlamtac *et al.* 2003] I. Chlamtac, M. Conti and J. J.-N. Liu, "*Mobile Ad-Hoc Networking: Imperatives and Challenges*", Ad Hoc Networks, vol.1 (1), pp.13-64, 2003.

[Choi *et al.* 1994] H. Choi, V. Kulkami and K. S. Trivedi, "*Markov Regenerative Stochastic Petri Nets*", Performance Evaluation, vol. 20, pp.337-356, 1994.

[Ciardo *et al.* 1990] G. Ciardo, A. R. Marie, B. Sericola and K. S. Trivedi, "*Performability Analysis Using Semi-Markov Reward Processes*", IEEE TRANSACTIONS ON COMPUTERS, vol. 39 (10), pp.1251-1264, 1990.

[Ciardo *et al.* 1989] G. Ciardo, J. Muppala and K. S. Trivedi, "*SPNP: Stochastic Petri Nets Package*", IEEE Computer Society Press, International Workshop on Petri Nets and Performance Models, pp.142-151, 1989.

[Ciciani & Grassi 1987] B. Ciciani and V. Grassi, "*Performability Evaluation of Fault-Tolerance Satellite Systems*", IEEE Transactions on Communications, vol. 35, pp.403-409, 1987.

[CORSIM] *CORSIM*, "http://mctrans.ce.ufl.edu/featured/TSIS/Version5/corsim.htm".

[Courtès 2007] L. Courtès, *Cooperative Data Backup for Mobile Devices*, École doctorale Systèmes, École doctorale Systèmes, spécialité Systèmes Informatiques, Institut National Polytechnique de Toulouse, 2007.

[Courtès *et al.* 2006] L. Courtès, M.-O. Killijian and D. Powell, "*Storage Tradeoffs in a Collaborative Backup Service for Mobile Devices*", Proc. of the 6th European Dependable Computing Conf., IEEE CS Press, pp.129-38, 2006.

[Cox & Miller 1965] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes*, Chapman and Hall Ltd., 1965.

[Daly *et al.* 2000] D. Daly, D. D. Deavours, J. M. Doyle, P. G. Webster and W. H. Sanders, "*Möbius: An extensible tool for perfomance and dependability modeling*", In 11th International Conference, TOOLS 2000, Lecture Notes in Computer Science, pp.332-336, Schaumnurg, IL B.R. Haverkort, H. C. Bohnenkamp, and C. U. Smith (Eds.), 2000.

[E2213 2003] E2213, "*Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle Systems : 5 GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY)*", ASTM E2213-03 2003.

[El Masri *et al.* 2010] M. El Masri, S. Abdellatif and G. Juanole, "*On Resource management in heterogeneous wireless access networks - application to automated highway systems*", Proceedings of the 10th annual international conference on New Technologies of Distributed Systemss, NOTERE'10, p. 6 2010.

[Enkelmann 2003] W. Enkelmann, "*FleetNet Applications for Inter-Vehicle Communication*", IEEE Intelligent Vehicles Symposium; Colmubus, OH, USA, IV, pp.162-167, 2003.

[Fenton *et al.* 1976] R. E. Fenton, G. Melocik and K. Olson, "*On the Steering of Automated Vehicles: Theory and Experiment*", IEEE Transaction on Automatic Control, vol. AC-21, pp.306-315, 1976.

[Fenton & Mayhan 1991] R. E. Fenton and R. J. Mayhan, "*Automated Highway Studies at the Ohio State University - An Overview*", IEEE Transaction on Vehicular Technology, vol. 40, pp.306-315, 1991.

[Fiore & Härri 2008] M. Fiore and J. Härri, "*The Networking Shape of Vehicular Mobility*", International Symposium on Mobile Ad Hoc Networking & Computing, Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing, Hong Kong, Hong Kong, China, pp.261-272, 2008.

[Fiore *et al.* 2007] M. Fiore, J. Härri, F. Fethi and C. Bonnet, "*Vehicular mobility simulation for VANETs*", Proc. of the 40th IEEE Annual Simulation Symposium (ANSS'07), 2007, Norfolk, USA, pp.301-309, 2007.

[Florin & Natkin 1985] G. Florin and S. Natkin, "*Les réseaux de Petri stochastic*", Technique et Science Informatiques, vol. 4 (1), pp.143-160, 1985.

[Fota *et al.* 1999] N. Fota, M. Kaâniche and K. Kanoun, "*Dependability Evaluation of an Air Traffic Control Computing Systems*", Performance Evaluation, vol.4 (34), pp.553-573, 1999.

[Furukawa 2000] Y. Furukawa, "*Status and Future Direction of Intelligent Drive Assist Technology*", IEEE Intelligent Transportation Systems, pp.113-118, 2000.

[Gass *et al.* 2006] R. Gass, J. Scott and C. Diot, "*Measurements of In-Motion 802.11 Networking*", Seventh IEEE Workshop on Mobile Computing Systems & Applications (WMCSA'06), Semiahmoo Resort, Washington, USA, pp.69-74, 2006.

[Gehring & Fritz 1997] O. Gehring and H. Fritz, "*Practical Results of a Longitudinal Control Concept for Truck Platooning with Vehicle to Vehicle Communication*", Proc. of the 1st IEEE Conference on Intelligent Transportation System (ITSC'97), pp.117-122, 1997.

[Godbole *et al.* 1996] D. N. Godbole, J. Lygeros, E. Singh, A. Deshpande and A. E. Lindsey, *Towards a Fault Tolerant AHS Design Part II: Design and Verification of Communication Protocols*, Institute of Transportation Studies, Paper UCB-ITS-PRR-96-15, 1996.

[Grassi *et al.* 2007] V. Grassi, R. Mirandola and A. Sabetta, "*A model-driven approach to performability analysis of dynamically reconfigurable component-based systems*", Proceedings of the 6th international workshop on Software and performance, pp.103-114, 2007.

[Hadaller *et al.* 2007] D. Hadaller, S. Keshav, T. Brecht and S. Agarwal, "*Vehicular Opportunistic Communication Under the Microscope*", International Conference On Mobile Systems, Applications And Services, Proceedings of the 5th international conference on Mobile systems, applications and services, pp.206-219, 2007.

[Hallé 2005] S. Hallé, *Automated Highway Systems: Platoons of Vehicles Viewed as a Multiagent System*, in Faculté des études supérieures de l'Université Laval, Québec, M.Sc., pp.194, 2005.

[Hallé & Chaib-draa 2005] S. Hallé and R. J. Chaib-draa, "*Collaborative Driving System Using Teamwork for Platoon Formations*", In Applications of Agent Technology in Traffic and Transportation, Whitestein Series in Software Agent Technologies, F. Klügl et al (Eds.), Birkhäuser Verlag, 2005.

[Hamouda 2006] O. Hamouda, Evaluation de la sûreté de fonctionnement d'un système de sauvegarde coopérative pour des dispositifs mobiles (Master de recherche), 2006.

[Hamouda *et al.* 2008] O. Hamouda, M. Kaâniche and K. Kanoun, "*Modélisation et Evaluation de la Sûreté de Fonctionnement d'un Système d'Autoroute Automatisée*", 16ème Congrès de Maîtrise de Risques et Sûreté de Fonctionnement, Communication 6C-2, Avignon- France, p.8, 2008.

[Hamouda *et al.* 2009a] O. Hamouda, M. Kaâniche and K. Kanoun, "*Safety Modeling and Evaluation of Automated Highway Systems*", The 39th Annual IEEE/IFIP international conference on Dependable Systems and Networks (DSN-09), Lisbon, Portugal, pp.73-82, 2009.

[Hamouda *et al.* 2010a] O. Hamouda, M. Kaâniche and K. Kanoun, "*Availability Modelling of a Virtual Black-Box for Automotive Systems*", RISE/EFTS Joint 2nd International Workshop on Software Engineering for REsilieNt systEms (SERENE 2010), pp.52-60, 2010.

[Hamouda *et al.* 2010b] O. Hamouda, M. Kaâniche and K. Kanoun, "*Modélisation et Evaluation de la Sûreté de Fonctionnement de Boîtes Noires Virtuelles pour Systèmes Automobiles*", 17ème Congrès de Maîtrise de Risques et Sûreté de Fonctionnement, Communication 4E-1, La Rochelle-France, p.8, 2010.

[Hamouda *et al.* 2009b] O. Hamouda, M. Kaâniche, E. Matthiesen Moller, J. Gulddahl Rasmussen and H.-P. Schwefel, "*Connectivity Dynamics in Vehicular Freeway Scenarios*", Proceedings of the Second international conference on Global Information Infrastructure Symposium, Hammamet, Tunisia, pp.365-372, 2009.

[Hansen *et al.* 2010] M. B. Hansen, R. L. Olsen and H.-P. Schwefel, "*Probabilistic models for access strategies to dynamic information elements*", Journal of Performance Evaluation, Year of Publication: 2010, vol. 67 (1), pp.43-60, 2010.

[Hansen *et al.* 2008] M. B. Hansen, J. G. Rasmussen and H.-P. Schwefel, "*Connectivity analysis of onedimensional ad-hoc networks*", DMF-2008-06-003 2008.

[Härri *et al.* 2009] J. Härri, F. Filali and C. Bonnet, "*Mobility Models for Vehicular Ad-Hoc Networks: A survey and Taxonomy*", IEEE Communications Surveys & Tutorials, vol. 11 (4) 2009.

[Hartenstein & Laberteaux 2008] H. Hartenstein and K. Laberteaux, "*A Tutorial Survey on Vehicular Ad hoc Networks*", IEEE Communications Magazine, pp.164-171, 2008.

[Hedrick *et al.* 1994] J. K. Hedrick, M. a. Tomizuka and P. Varaiya, "*Control Issues in Automated Highway Systems*", Control Systems Magazine, IEEE, vol. 14 (6), 1994.

[Heidelberge *et al.* 1992] P. Heidelberge, V. Nicola and P. Shahabuddin, "*Simultaneous and Efficient Simulation of Highly Dependable Systems with Different Underlying Distributions*", Winter Simulation Conference, vol. 4, pp.137-164, 1992.

[Helal *et al.* 1996] A. Helal, A. Heddaya and B. Bhargava, "*Replication Techniques in Distributed Systems*", Kluwer Academic Publishers 1996.

[HIDENETS 2006a] HIDENETS, "*Highly dependable ip-based networks and services*, http://www.hidenets.aau.dk/, IST-FP6-STREP-26979", 2006.

[HIDENETS 2006b] HIDENETS (Ed.), Highly Dependable ip-based networks and services, *use case scenarios and preliminary reference model*, Project Deliverable D1.1, IST-FP6-STREP-26979, 2006.

[HIDENETS 2007a] HIDENETS (Ed.), HIghly Dependable ip-based NETworks and Services, Evaluation methodologies, *techniques and tools (final version)*, Project Deliverable D4.1.2, IST-FP6-STREP-26979, 2007.

[HIDENETS 2007b] HIDENETS (Ed.), HIghly Dependable ip-based NETworks and Services, *Resilient architecture (final version)*, Project Deliverable D2.1.2, IST-FP6-STREP-26979, 2007.

[HIDENETS 2007c] HIDENETS (Ed.), HIghly Dependable ip-based NETworks and Services, *Revised Reference Model*, Project Deliverable D1.2, IST-FP6-STREP-26979, 2007.

[HIDENETS 2008] HIDENETS (Ed.), HIghly Dependable ip-based NETworks and Services, *Final evaluation, consolidated results and guidelines*, Project Deliverable D1.3, IST-FP6-STREP-26979, 2008.

[Hong *et al.* 1999] X. Hong, M. Gerla, G. Pei and C.-C. Chiang, "*A Group Mobility Model for Ad-Hoc Wireless Networks*", International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems, Proceedings of the 2nd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems, Washington, USA, pp.53-60, 1999.

[Howard 1971] R. Howard, "*Dynamic Probabilistic Systems: Semi-Markov and Decision Processes*", Wiley, New York 1971.

[Hsu *et al.* 1994] A. Hsu, F. Eskafi, S. Sachs and P. Varaiya, "*Protocol design for an automated highway system*", Discrete Event Dynamic Systems, vol. 2 (1), pp.183-206, 1994.

[Hui 2005] F. Hui, *Experimental Characterization of Communications in Vehicular Ad-Hoc Network*, M.Sc. in Computer Science, Davis, University of California, 2001, 2005.

[Hyytiä *et al.* 2006] E. Hyytiä, P. Lassila and J. Virtamo, "*Spatial Node Distribution of the Random Waypoint Mobility Model with Applications*", IEEE Trans. Mobile Computing, vol. 5 (6), pp.680-694, 2006.

[IEEE_Standard 1999] IEEE_Standard, "*Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*", 1999.

[ISTAG 2010] ISTAG, "*IST Advisory Group: Scenarios for Ambient Intelligence in 2010*", ftp://ftp.cordis.lu/pub/ist/docs/istagscenarios2010.pdf 2010.

[Jackson 1963] J. Jackson, "*Jobshop-Like Queuing Systems*", Management Science, vol.1 (10), pp.131-142, 1963.

[Jetcheva *et al.* 2003] J. Jetcheva, Y.-C. Hu, S. PalChaudhuri, A. K. Saha and D. B. Johnson, "*Design and evaluation of a metropolitan area multitier wireless ad-hoc network architecture*", Proceedings of the 5th IEEE Workshop on Mobile Computing Systems & Applications (WMCSA), Monterey, CA, pp.32-43, 2003.

[Jiang *et al.* 2006] D. Jiang, V. Taliwal, A. Meier, W. Holfelder and R. Herrtwich, "*Design of 5.9 GHz DSRC-based Vehicular Safety Communication*", IEEE Wireless Communications magazine, vol.13 (5) 2006.

[Jones & Randell 2004] C. Jones and B. Randell, "*Dependable Pervasive Systems*", University of Newcastle Research Report CS-TR-839 2004.

[Jr 2009] S. O. Jr, "*IEEE 802.11N: The Road Ahead*", IEEE Computer Society Press., 42 (7) (doi:10.1109/MC.2009.224), pp.13-15, 2009.

[Juanole 2003] G. Juanole, "*Réseaux de Petri Stochastiques*", Ecole d'Eté Temps Réel (ETR'2003), Toulouse, France, 9-12 Septembre, pp.137-154, 2003.

[Kaâniche *et al.* 2003] M. Kaâniche, K. Kanoun and M. Rabah, "*Multilevel Modeling Approach for the Availability Assessment of E-Business Applications*", Software-Practice and Experience, vol. 33, pp.1323-1341, 2003.

[Kaâniche *et al.* 2008] M. Kaâniche, P. Lollini, A. Bondavalli and K. Kanoun, "*Modeling the Resilience of Large and Evolving Systems*", International Journal on Performability Engineering, vol.4 (2), pp.153-168, 2008.

[Kanoun & Borrel 2000] K. Kanoun and M. Borrel, "*Fault-Tolerant System Dependability - Explicit Modeling of Hardware and Software Component-Interactions*", IEEE Transaction on Reliability, vol. 4 (49), pp.363-376, 2000.

[Karagiannis *et al.* 2007] T. Karagiannis, J.-Y. Le Boudec and M. Vojnovic, "*Power Law and Exponential Decay of Inter Contact Times between Mobile Devices*", Proc. of the 13th annual ACM international conf. on Mobile computing and networking 2007.

[Kemeny & Snell 1960] J. G. Kemeny and J. L. Snell, "*Finite Markov Chains*", D. Van Nostrand Co., Inc., Princeton, New Jersy, USA 1960.

[Khadar & Simplot-Ryl 2007] F. Khadar and D. Simplot-Ryl, "*Connectivity and topology control in wireless ad-hoc networks with realistic physical layer*", Conference on Wireless and Mobile Communications (ICWMC'07), IEEE Computer Society Washington, DC, USA (Proceedings of the Third International), p.49, 2007.

[Killijian *et al.* 2004] M.-O. Killijian, D. Powell, M. Banâtre, P. Couderc and Y. Roudier, "*Collaborative Backup for Dependable Mobile Applications*", Proc. of 2nd Int. Workshop on Middleware for Pervasive and Ad-Hoc Computing (Middleware 2004), pp.146-149, 2004.

[Killijian *et al.* 2009] M. O. Killijian, M. Roy, G. Severac and C. Zanon, "*Data backup for mobile nodes: a cooperative middleware and experimentation platform*", Workshop on Architecting Dependable Systems (WADS), supplemental volume of DSN-09, Lisbon (Portugal), p.6, 2009.

[Könning *et al.* 2009] B. Könning, R. Manfred and e. al., "*Final Evaluation, Consolidated Results and Guidelines*", Project Deliverable D1.3, IST-FP6-STREP-26979 2009.

[Kotz & Henderson 2005] D. Kotz and T. Henderson, "*CRAWDAD: A Community Resource for Archiving Wireless Data at Dartmouth*", Pervasive Computing, IEEE, vol. 4 (4), pp.12-14, 2005.

[Kovacs *et al.* 2008] M. Kovacs, P. Lollini, I. Majzik and A. Bondavalli, "*An integrated framework for the dependability evaluation of distributed mobile applications*", RISE/EFTS Joint International Workshop on Software Engineering for REsilieNt systEms (SERENE 2008), pp.29-38, 2008.

[Laprie *et al.* 1995] J.-C. Laprie, M. Kaâniche, K. Kanoun and ... *Guide de la Sûreté de Fonctionnement*, édition Cépaduès, Laboratoire d'Ingénierie de la Sûreté de Fonctionnement, 1995.

[Lee & Gerla 2010] U. Lee and M. Gerla, "*A survey of urban vehicular sensing platforms*", Advances in Wireless and Mobile Networks, vol.54 (4), pp.527-544, 2010.

[Li & Wang 2007] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: A survey", Vehicular Technology Magazine, IEEE, 2 (2), pp.12-22, 2007.

[Lin *et al.* 2004] W. K. Lin, D. M. Chiu and Y. B. Lee, "*Erasure Code Replication Revisited*", Proc. of the 4th P2P, pp.90-97, 2004.

[Lygeros 1995] J. Lygeros, et al., "*Design of an Extended Architecture for Degraded Modes of Operation of IVHS*", in American Control Conference, UCB-ITS'PWP'95'3, pp.3592-3596, 1995.

[Lygeros *et al.* 1996] J. Lygeros, D. N. Godbole and M. Broucke, *Towards a Fault Tolerant AHS Design Part I: Extended Architecture, Institute of Transportation Studies*, PATH Technical Report UCB-ITS-PRR-96-14, 1996.

[Lygeros *et al.* March 2000] J. Lygeros, D. N. Godbole and M. Broucke, "*A Fault Tolerant Control Architecture for Automated Highway Systems*", Control Systems Technology, 8 (2), pp.205-219, March 2000.

[Marie *et al.* 1997] R. A. Marie, B. Plateau, M. Calzarossa and G. Rubino, "*Computer Performance Evaluation: Modelling Techniques and Tools*", 9th International Conference, St. Malo, France, June 3-6, Proceedings Springer 1997.

[Marsan *et al.* 1995] M. A. Marsan, G. Balbo, G. Conte, S. Donatelli and G. Franceschinis, *Modeling with Generalized Stochastic Petri Nets*, 139-141p., John Wiley & Sons Ltd., 1995.

[Marsan & Chiola 1987] M. A. Marsan and G. Chiola, "*On petri nets with deterministic and exponentially distribuited ring times*", Lecture Notes in Computer Science, Springer Verlang, vol. 266, pp.132-145, 1987.

[Martinello 2005] M. Martinello, "*Availability modeling and evaluation of web based services - a pragmatic approach*", PhD. Dissertation, Institut National Polytechnique de Toulouse, LAAS-CNRS Report N°. 05552, 2005.

[Masayasu *et al.* 2003] J. Masayasu, S. Shigeki, U. Ken'ya and M. Hiroshi, "*Design of Lane-Keeping Control with Steering Torque Input*", Transaction of Society of Automotive Engineerings of Japan, vol. 53 (1), pp.163-168, 2003.

[Matthiesen Moller *et al.* 2008] E. Matthiesen Moller, O. Hamouda, M. Kaâniche and S. H.-P., "*Dependability evaluation of a replication service for mobile applications in dynamic ad-hoc networks*", ISAS08, Tokyo, Japan (Lecture Notes in Computer Science 5017 Springer'08), pp.171-186, 2008.

[Matthiesen Moller *et al.* 2007] E. Matthiesen Moller, T. Renier and H.-P. Schwefel, "*A new selection metric for backup group creation in inter-vehicular networks*", In 16th IST Mobile and communications summit 2007.

[Menascé & Almeida 2000] D. A. Menascé and V. A. F. Almeida, "*Scaling for E-Business: Technologies, Models, Performance and Capacity Planning*", Prentice Hall Inc. 2000.

[Meyer 1980] J. F. Meyer, "*On evaluating the performability of degradable computing systems*", IEEE Journal on Selected Areas in Communications, vol. 29 (8), pp.720-731, 1980.

[Meyer 1982] J. F. Meyer, "*Closed-form solutions of performability*", IEEE Transactions on Computing, vol. 7 (31), pp.648-657, 1982.

[Meyer *et al.* 1985] J. F. Meyer, A. Movaghar and W. H. Sanders, "*Stochastic Activity Networks: Structure, Behaviour and Applications*", Proc. International Workshop on Timed Petri Nets, Torino, Italy, pp.106-115, 1985.

[Miller 1997] M. Miller, "*PATH: Societal and Institutional Issues of Automated Highway Systems*", Intellimotion Paper News, vol. 6 (3) 1997.

[Miorandi & Altman 2006] D. Miorandi and E. Altman, "*Connectivity in one-dimensional ad-hoc networks: A queueing theoretical approach*.", Wireless Networks, vol. 12, pp.573-587, 2006.

[Møller & Waagepetersen 2004] J. Møller and R. P. Waagepetersen, "*Statistical Inference and Simulation for Spatial Point Processes*", Chapman & Hall, Boca Raton, Florida 2004.

[Musolesi & Mascolo 2009] M. Musolesi and C. Mascolo, "*Mobility Models for Systems Evaluation. A Survey*", Middleware for Network Excentric and Mobile Applications, Benoit Garbinato, Hugo Miranda and Luis Rodrigues (Eds.), Springer, pp.43-62, 2009.

[Network_Simulator_2] *Network_Simulator_2*, "http://www.isi.edu/nsnam/ns/".

[Nicol *et al.* 2004] D. Nicol, W. H. Sanders and K. S. Trivedi, "*Model-Based Evaluation: From Dependability to Security*", IEEE Transactions on Dependable and Secure Computing, vol. 1 (1), pp.48-65, 2004.

[Nicola 1990] V. Nicola, "*Lumpability of Markov Reward Models*", Technical report, IBM, T.J. Watson Research Center 1990.

[Olesen *et al.* 2006] R. L. Olesen, M. B. Hansen and H.-P. Schwefel, "*Quantitative analysis of access strategies to remote information in network services*", In Global Telecommunications Conference, GLOBECOM - IEEE 2006.

[Popstojanova & Trivedi 2000] K. G. Popstojanova and K. S. Trivedi, "*Stochastic Modeling Formalism for Dependability, Performance, and Performability*", Performance Evaluation: Origins and Directions, pp.403-422, Springer-Verlag, 2000.

[Rabah & Kanoun 2003] M. Rabah and K. Kanoun, "*Perfomability Evaluation of Multipurpose Multiprocessor Systems: the separation of concers*", IEEE Transactions on Computing, vol. 2 (52), pp.223-236, 2003.

[Radimirsch *et al.* 2006] M. Radimirsch, I. De Bruin, A. Casimiro Costa, A. Fosslie-Hansen, G. Huszerl, T. Ingvalsden, M. Kaâniche, M.-O. Killijian, M. Löbbers, E. Matthiesen Moller, M. Reitenspiess, N. Rivière, I. E. Svinnset and H. Waeselynck, "*Use-Case Scenarios and Preliminary Reference Model*", EU FP6-IST Project HIDENETS, Deliverable D1.1 2006.

[Reichardt *et al.* 2002] D. Reichardt, M. Miglietta, L. Moretti, P. Morsink and W. H. Schulz, "*Cartalk 2000— Safe and Comfortable Driving Based upon Inter-Vehicle Communication*", Proceedings of the IEEE Intelligent Vehicle Symposium (IV02) 2002.

[Reiser & Lavenberg 1980] M. Reiser and S. Lavenberg, "*Mean-Value Analysis of Closed Multi-Chain Queuing Networks*", Journal of ACM, vol. 27 (2) 1980.

[Rényi 1964] A. Rényi, "*On Two Mathematical Models of The Traffic on a Divided Highway*", Journal of Applied Probability, vol. 1, pp.311-320, 1964.

[Report_802.11p]Report_802.11p,
"http://grouper.ieee.org/groups/802/11/Reports/tgp_update.htm".

[Roche & Schabes 1997] E. Roche and Y. Schabes, *Finite-State Language Processing*, Bradford Book. MIT Press, Cambridge, Massachusetts, USA, 1997.

[Sadagopan *et al.* 2003] N. Sadagopan, F. Bai, B. Krishnamachari and A. Helmy, "*PATHS: Analysis of PATH Duration Statistics and Their Impact on Reactive MANET Routing Protocols*", MobiHoc'03, Annapolis, Maryland, USA, pp.245-256, 2003.

[SAF 2007] SAF, "*Service Availability Forum, Release 5*", http://www.saforum.org/, 2007.

[SAF 2009] SAF, "*Service Availability Forum*", http://www.saforum.org/, 2009.

[Sakaguchi *et al.* 2000] T. Sakaguchi, A. Uno, S. Kato and S. Tsugawa, "*Cooperative Driving of Automated Vehicles with Inter-Vehicle Communications*", In Proceedings of IEEE Intelligent Vehicles Symposium, pp.516-521, 2000.

[Samar & Wicker 2004] P. Samar and S. B. Wicker, "*On the Behavior of Communication Links of a Node in a Multi-Hop Mobile Environment*", Proceedings of the 5th ACM international symposium on Mobile ad-hoc networking and computing, Roppongi Hills, Tokyo, Japan, pp.145-156, 2004.

[Sanders & Meyer 2001] W. H. Sanders and J. F. Meyer, "*Stochastic activity networks: Formal definitions and concepts*", In Lectures on Formal Methods and Performance Analysis, pp.315-343. Springer Verlag, 2001.

[Seada 2008] K. Seada, "*Insights from a Freeway Car-to-Car Real-World Experiment*", ACM Mobicom Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization, San Francisco, CA, pp.49-56, 2008.

[Sichitiu & Kihl 2008] L. M. Sichitiu and M. Kihl, "*Inter-Vehicle Communication Systems: A Survey*", IEEE Communications Surveys & Tutorials, 2nd Quarter, pp.88-105, 2008.

[Smith *et al.* 1988] R. M. Smith, K. S. Trivedi and A. V. Ramesh, "*Performability Analysis: Measures, an Algorithm, and a Case Study*", IEEE Transactions on Computers, vol. 37 (4), pp.406-417, 1988.

[Spyropoulos *et al.* 2008] T. Spyropoulos, A. Jindal and K. Psounis, "*An Analytical Study of Fundamental Mobility Properties for Encounter-based Protocols*", International Journal of Autonomous and Adaptive Communications Systems, vol. 1 (1), pp.4-40, 2008.

[Sugiura & Dermawan 2005] A. Sugiura and C. Dermawan, "*In traffic jam IVC-RVC system for ITS using Bluetooth*", IEEE Trans. Intelligent Transportation Sys., vol. 6 (3), pp.302-313, 2005.

[Torrent-Moreno 2007] M. Torrent-Moreno, "*Inter-Vehicle Communications : Achieving Safety in a Distributed Wireless Environment : Challenges, Systems and Protocols*", PhD. Dissertation, ISBN: 978-3-86644-175-0, Universitätsverlag Karlsruhe 2007.

[Traces_MMTS] Traces_MMTS, "*These traces are publicly available from : http://www.lst.inf.ethz.ch/research*".

[Trivedi *et al.* 1992] K. S. Trivedi, J. Muppala, S. Woolet and B. Haverkort, "*Composite Performance and Dependability Analysis*", Performance Evaluation, vol. 14, pp.197-215, 1992.

[Tsuji *et al.* 2001] M. Tsuji, R. Shirato, H. Furusho and K. Akutagawa, "*Estimation of Road Configuration and Vehicle Attitude by Lane Detection for Lane Keeping system*", Society of Automotive Engineers, pp.45-51, 2001.

[UDS_White_Paper 2007] UDS_White_Paper, "*Accident Data Recorder - A Contribution to Road Safety*", National Highway Traffic Safety Administration - www.nhtsa.dot.gov 2007.

[Varaiya 1993] P. Varaiya, "*Smart Cars on Smart Roads: Problems of Control*", IEEE Transaction on Automatic Control, vol. 38 (2), pp.195-207, 1993.

[Wang *et al.* 2007] D. Wang, W. Xie and K. S. Trivedi, "*Performability analysis of clustered systems with rejuvenation under varying workload*", Performance Evaluation, vol. 64 (3), pp.247-265, 2007.

[Weatherspoon & Kubiatowicz 2002] H. Weatherspoon and J. Kubiatowicz, "*Erasure-Coding vs. Replication: A Quantitative Comparison*", Revised Papers from the 1st International Workshop on P2P Systems, pp.328-338, Springer-Verlang, 2002.

[Wischhof *et al.* 2005] L. Wischhof, A. Ebner and H. Rohling, "*Information Dissemination in Self-Organizing Intervehicle Networks*", IEEE Trans. Intelligent Transportation Sys., vol. 6 (1), pp.90-101, 2005.

[Xu *et al.* 1999] L. Xu, V. Bohossian, J. Bruck and D. G. Wagner, "*Low Density MDS Codes and Factors of Complete Graphs*", IEEE Transactions on Information Theory, vol. 45 (1), pp.1817-1826, 1999.

[Xu *et al.* 2007] S. Xu, K. L. Blackmore and H. M. Jones, "*An Analysis Framework for Mobility Metrics in mobile Ad-Hoc Networks*", EURASIP Journal on Wireless Communications and Networking, vol. 1, pp.26-42, 2007.

[Yang & Kim 2003] B.-M. Yang and J. Kim, "*Road Traffic Accidents and Policy Interventions in Korea*", Injury Control and Safety Promotion, Swets & Zeitlinger, vol. 10 (2), pp.89-94, 2003.

[Yang *et al.* 2004] X. Yang, J. J.-N. Liu, F. Zhao and N. Vaidya, "*A Vehicle-to-Vehicle Communication Protocol for Cooperative Collision Warning*", Proc. 1st Annual Int'l. Conf. Mobile and Ubiquitous Sys.: Networking and Services, pp.1-4, 2004.

[Yashiro *et al.* 1993] T. Yashiro, T. Kondo, H. Yagome, M. Higuchi and Y. Matsushita, "*A Network Based on Inter-vehicle Communication*", Proc. of Intelligent Vehicles Symp., pp.345-350, 1993.

[Zhang *et al.* 2007] X. Zhang, J. Kurose, B. N. Levine, D. Towsley and H. Zhang, "*Study of a Bus-based Disruption-Tolerant Network: Mobility Modeling and Impact on Routing*", International Conference on Mobile Computing and Networking, Proceedings of the 13th annual ACM international conference on Mobile computing and networking, Montréal, Québec, Canada, pp.195-206, 2007.

Résumé

Cette thèse porte sur le développement de méthodes et de modèles permettant de quantifier la sûreté de fonctionnement d'applications embarquées sur des systèmes mobiles. L'objectif est de fournir des indicateurs pour l'analyse et la sélection des architectures les plus adaptées pour satisfaire les exigences de sûreté de fonctionnement. Nous considérons le cas des applications véhiculaires utilisant des communications inter-véhicules basées sur des réseaux ad-hoc et pouvant avoir accès à des services situés sur des infrastructures fixes.

Nous proposons une approche combinant: 1) des modèles de sûreté de fonctionnement utilisant des réseaux d'activités stochastiques permettant de décrire les modes défaillances et de restauration des systèmes considérés et 2) des simulations et des modèles analytiques de scénarios de mobilité permettant d'estimer des caractéristiques de connectivité de ces systèmes. Cette approche est illustrée sur trois cas d'études dont un système d'autoroute automatisée, et une boîte noire virtuelle basée sur la réplication et la sauvegarde coopérative des données.