



Defence Research and  
Development Canada

Recherche et développement  
pour la défense Canada



# Measures of performance for analysis of tactical communication simulations conducted using a low bandwidth test bed

*J.-D. Caron  
L. Stemate  
DRDC Valcartier*

**Defence R&D Canada – Valcartier**

Technical Memorandum

DRDC Valcartier TM 2006-791

February 2007

Canada



# **Measures of performance for analysis of tactical communication simulations conducted using a low bandwidth test bed**

J.-D. Caron  
L. Stemate  
DRDC Valcartier

**Defence R&D Canada – Valcartier**

Technical Memorandum

DRDC Valcartier TM 2006-791

February 2007

Authors

---

Jean-Denis Caron

Luminita Stemate

Approved by

---

Dean Haslip

Section Head, Land Operational Research

Approved for release by

---

Christian Carrier

Chief Scientist

© Her Majesty the Queen as represented by the Minister of National Defence, 2007

© Sa majesté la reine, représentée par le ministre de la Défense nationale, 2007

## Abstract

---

This memorandum summarizes the three-fold Operational Research (OR) support provided to the Technology Demonstration Project entitled High Capacity Technical Communications Network (HCTCN), by the Defence Research and Development Canada Valcartier OR team. First, refinements were proposed to three basic measures of performances implemented in the Low Bandwidth Test Bed simulation environment developed for this project: location fidelity, currency and latency. Second chronologically, but the main effort of this study was to develop a tool that implements these refined measures. This tool, fully documented in this document, including a user's guide, represents one of the key requirements for the successful analysis and delivery of the results obtained through experimentation by the HCTCN project team. Third, the memorandum also addresses the question of the number of times each scenario needs to be executed within the test bed in order to ensure statistically valid results.

## Résumé

---

Ce mémorandum résume le soutien en trois volets que l'équipe de Recherche opérationnelle de Recherche et développement pour la défense Canada - Valcartier a apporté au projet de démonstration de technologies intitulé: Réseaux de communications tactiques à grande capacité, aussi connu sous le nom de HCTCN. Premièrement, des améliorations ont été proposées pour trois mesures de performance de base implantées dans un environnement de simulation spécialement développé pour ce projet: fidélité de la position, temps entre les mises à jour et délai de transmission. Deuxième chronologiquement mais l'effort principal de cette étude a été le développement d'un outil qui intègre ces mesures améliorées. Cet outil, documenté de façon complète dans ce document, incluant un guide de l'utilisateur, représente un des éléments clé pour l'analyse et la livraison avec succès des résultats obtenus par expérimentation par l'équipe du projet HCTCN. Troisièmement, la question du nombre de fois qu'un scénario doit être exécuté avec le banc d'essai pour que les résultats soient statistiquement valides est également abordée dans ce document.

This page intentionally left blank.

## Executive summary

---

On the tactical battlefield, the means of communication between vehicles or dismounted soldiers is generally a line-of-sight radio operating in the VHF or UHF band. Sharing information in digital form on an all-informed basis is highly desirable to avoid a single point of failure and to ensure continuity of operations. Under this exchange model, nodes (vehicles, dismounted soldiers, etc.) try to maintain exact copies of key information in each other's database via asynchronous data replication. When communication channels have low and variable throughput and unreliable connectivity, maintaining full synchronization is next to impossible. Initial experience with transmitting tactical data in digital form over Army radios has highlighted the inadequate data capacity of the wireless channels.

In April 2000, a technology demonstration project, entitled "High Capacity Technical Communications Network (HCTCN) Low Bandwidth Test Bed (LBTB)", was initiated. The objective of the project was to demonstrate the potential for selected technologies in wireless communications networks and information management to increase the limited capacity of tactical communications systems to support command and control requirements. The project featured three areas of investigation: tactical networking, high capacity radio bearers and information management. For the information management segment of the project (for which DRDC Valcartier had responsibility) a research test bed concept called the Low Bandwidth Test Bed (LBTB) was conceived by project staff at DRDC Valcartier and implemented through contract, to evaluate battlefield information management strategies applied in a low bandwidth tactical wireless communications environment. As part of the technical specifications for this contract, the capacity to support three basic measures of performance was defined for inclusion in the LBTB. These were:

- Currency – Time elapsed since the last update of a given piece of information in a node's tactical database;
- Latency – Time taken to transfer a data replication (messages and information of any type) between two nodes; and
- Data Consistency – Comparison of data values in designated table row(s) or in column(s) in a table row, in two or more different databases to determine whether data values are equal.

During the course of the contract, the data consistency measurement was dropped due to the technical complexity of its implementation. A fourth measurement, location fidelity, was added to the specifications after the contract had begun:

- Location Fidelity – Comparison between the real location of a node and its perceived location by other tactical nodes.

The members of HCTCN requested Operational Research (OR) support from the DRDC Valcartier OR group in autumn 2004. The OR support included the following:

- An analysis regarding the number of times a scenario should be executed with the LBTB to ensure statistically valid results;
- An analysis of the existing Measures of Performance (MOPs) implemented in the LBTB, and the proposal for refinements to the existing MOPs to better suit the detailed analysis requirements of the HCTCN TDP. These refinements included several possible interpretations of the Location Fidelity measure as well as the consideration of more aspects such as transmission priority and replication message type; and
- Development of a software tool to implement data analysis protocols based on the refined MOPs.

In LBTB, as a scenario is processed, a large quantity of information about the events, the changes to the database, the network performance and a set of pre-defined measurements are kept in the log database for future analysis. Not all of this is relevant for the location fidelity, latency and currency measurements. The authors developed Structured Query Language (SQL) scripts to retrieve the information required to calculate these MOPs. The relevant information is copied to Excel files that are used by a Visual Basic for Applications Microsoft Excel spreadsheet tool, built by the authors, to compute the measures.

The statistical analysis conducted showed significant variability in the results caused by the randomness in LBTB. The independent replications method was used to determine, for each measure, the expected relative error as a function of the number of times a scenario is executed with the LBTB. It was shown that for location fidelity and latency, because of the important variability, a large number of runs (at least 30) should be executed in order to provide statistically valid results. For the currency measure, the number is approximately 20 runs.

The project sponsors also asked for a similar analysis regarding four other measures: RTL-RTL Transmission (Tx) Success Rate, Channel Access Success Rate, Channel Reception Success Rate and Average ReTx/Tx Ratio. It appeared that the randomness in the simulation did not affect as much these four measures; the variations in the measurements between consecutive runs of the same scenario were not significant. The analysis suggested that approximately 15 runs with the LBTB would be enough for these four measures.



Some examples of the use of the MOPs tool were documented in this memorandum. However, the objective of the OR support was not to analyze scenarios or experimental data. The goal was rather to provide the members of HCTCN with a tool to analyze the results of scenarios ran with the LBTB. Analysis of experimental results will be documented in the final report for the information management segment of the HCTCN TD project.

Caron, J-D. and Stemate, L. 2007. Measures of performance for analysis of tactical communication simulations conducted using a low bandwidth test bed. DRDC Valcartier TM 2006-791 Defence R&D Canada – Valcartier.

## Sommaire

---

Sur le champ de bataille, au niveau tactique, la communication entre véhicules ou entre les soldats à pied se fait généralement à l'aide de radios à portée directe fonctionnant dans la bande VHF ou UHF. Le partage de l'information sous forme numérique selon le principe de "l'information disponible à tous" est à privilégier afin d'éviter la défaillance d'un système centralisé et d'assurer la poursuite des opérations. Dans ce modèle d'échange, les nœuds tentent de maintenir des copies exactes de chacune de leurs bases de données par la réplication asynchrone de données (les contenus des bases de données sont entièrement synchronisés). Lorsque les canaux de communication présentent un débit faible et variable ainsi qu'une connectivité peu fiable, il peut s'avérer impossible de conserver une synchronisation parfaite. Jusqu'à maintenant, les canaux sans fil se sont avérés inadéquats lors de la transmission de données tactiques sous forme numérique par les radios de l'armée.

En avril 2000, un projet de démonstration de technologies, intitulé "Réseaux de communications tactiques à grande capacité", connu sous le nom de HCTCN, a vu le jour à Recherche et développement pour la défense Canada (RDDC) - Valcartier. L'objectif du projet était de démontrer le potentiel de certaines technologies relatives aux communications sans fil et à la gestion de l'information visant à accroître la capacité limitée des systèmes de communications tactiques. Le projet comprenait trois principaux domaines de recherche: réseautage tactique, radios tactiques mobiles à grande capacité et gestion de l'information. Pour la partie gestion de l'information du projet (dont RDDC Valcartier avait la responsabilité), un banc d'essai, appelé Low Bandwidth Test Bed (LBTB) conçu par RDDC Valcartier et implanté par un entrepreneur, afin d'étudier l'incidence de techniques de gestion de l'information sur la qualité et la rapidité de diffusion des données sur un réseau radio tactique à bande passante étroite. Dans la spécification technique initiale du contrat, les trois mesures de la performance (MP) de base suivantes devaient être incluses:

- Temps entre les mises à jour – Temps depuis la dernière mise à jour d'un type d'information donné dans la base de données tactique d'un nœud;
- Délai de transmission – Temps pris pour transférer une réplication (message et information de tout genre) entre deux nœuds; et
- Cohérence de l'information – Comparaison entre des données de rangées ou colonnes sélectionnées, dans deux ou plusieurs bases de données, pour déterminer si les valeurs sont identiques.

En cours du projet, la mesure de la cohérence de l'information a été enlevée étant donné sa complexité d'implantation. Par contre, une quatrième mesure, fidélité de la position, a été ajoutée dans la spécification après que le contrat a commencé:

- Fidélité de la position – Comparaison entre la vraie position d'un nœud et sa position telle que perçue par les autres nœuds tactiques.

A l'automne 2004, les membres de HCTCN ont demandé du soutien de recherche opérationnelle (RO) de l'équipe de RO de RDDC Valcartier. Le soutien de RO consistait en:

- Une analyse ayant pour but de déterminer les nombres de fois qu'un scénario devrait être répété avec le LBTB, afin de produire des résultats statistiquement valides;
- Analyser les MP implantées dans LBTB, et proposer une amélioration de celles-ci pour qu'elles soient mieux adaptées aux types d'analyses envisagés dans le projet HCTCN. Ces améliorations incluaient différentes interprétations de la mesure de fidélité de la position ainsi que la possibilité de tenir compte d'autres aspects dans l'analyse, tels que la priorité de transmission et le type de réplication; et
- Développement d'un outil permettant une analyse détaillée basées sur les MP améliorées.

Dans LBTB, lorsqu'un scénario est exécuté, une grande quantité d'information sur les événements, des changements dans les bases de données, des mesures de la performance et autres informations présélectionnées est sauvegardée dans les bases de données pour fins d'analyse. Par contre, ces données enregistrées ne sont pas toutes pertinentes pour le calcul des MP. Les auteurs ont créé une requête Structured Query Language (SQL) permettant de retirer des bases de données toute l'information nécessaire pour le calcul des trois MP. Cette information est exportée dans une série de fichiers Excel, lesquels sont utilisés par un outil développé en Visual Basic for Applications Microsoft Excel, pour le calcul des MP proposées dans ce document.

Suite à une analyse statistique, il a été démontré qu'il existait une grande variabilité dans les résultats causée par les aspects stochastiques du LBTB. Une méthode, appelée "independent replications" a été utilisée afin de déterminer, pour chacune des mesures, l'erreur relative attendue en fonction du nombre de fois qu'un scénario donné est exécuté avec le LBTB. On a démontré que, pour les mesures de fidélité de la position et délai de transmission, étant donné la grande variabilité dans les résultats, les usagers devraient effectuer un très grand nombre d'exécutions (au moins 30) pour obtenir des résultats statistiquement valides. Pour ce qui est de la mesure du temps entre les mises à jour, ce nombre se situe autour de 20.

Les membres HCTCN ont aussi demandé une analyse similaire pour quatre autres mesures: RTL-RTL Transmission (Tx) Success Rate, Channel Access Success Rate, Channel Reception Success Rate et Average ReTx/Tx Ratio. Il semble que ces quatre mesures ne soient pas autant affectées par les aspects stochastiques du banc d'essai. En effet, les variations observées entre les résultats d'exécutions consécutives ne sont pas très significatives. Dans le cas de ces mesures, l'analyse statistique a suggéré que 15 exécutions avec le LBTB d'un scénario donné seraient suffisantes pour produire des estimations statistiquement valides.

Quelques exemples de l'utilisation de l'outil développé ont été inclus dans ce mémorandum. Cependant, l'objectif de la contribution de la RO au projet n'était pas d'analyser les scénarios ou les données expérimentales. Le but premier était plutôt de fournir aux membres de HCTCN, un outil permettant l'analyse de scénarios exécutés à l'aide du LBTB. Les résultats d'analyse de la partie gestion de l'information du projet HCTCN, obtenus avec l'outil développé, seront documentés dans un rapport technique final.

Caron, J-D. et Stemate, L. 2007. Measures of performance for analysis of tactical communication simulations conducted using a low bandwidth test bed. DRDC Valcartier TR 2006-791 R et D pour la défense Canada – Valcartier.

# Table of contents

---

Abstract / Résumé .....	i
Executive summary .....	iii
Sommaire .....	vi
Table of contents .....	ix
List of figures .....	xi
List of tables .....	xiii
Acknowledgements .....	xiv
1. Introduction .....	1
1.1 Tactical Communication Environment.....	1
1.2 High Capacity Technical Communications Network.....	1
1.3 LBTB Overview .....	2
1.4 Operational Research Support to HCTCN .....	3
2. Proposed Improvements to Measures of Performance .....	4
2.1 General Thoughts on Measures of Performance .....	4
2.2 Location Fidelity .....	5
2.3 Currency .....	11
2.4 Latency .....	13
3. Number of Runs Required to Ensure Statistical Validity of the Results .....	17
3.1 Stochastic Nature of LBTB .....	17
3.2 Discussion on the Variability .....	18
3.3 Method for Estimating the Sample Size .....	20
3.4 Sample Size – Location Fidelity.....	21
3.5 Sample Size – Currency .....	22
3.6 Sample Size – Latency .....	22
3.7 Sample Size – Network Measures .....	23

3.8	Comments.....	23
4.	Examples of the Use of the Measures .....	25
4.1	Transmission Delays .....	25
4.2	Location Fidelity .....	26
5.	Conclusion.....	29
	References .....	30
	Annex A: User Guide .....	31
	Annex B: SQL Script.....	46
	Annex C: Graphical User Interface .....	50
	List of Symbols/Abbreviations/Acronyms .....	53
	Distribution List.....	54

## List of figures

---

Figure 1: Location Fidelity – Smallest Circle that Includes All Points .....	6
Figure 2: Location Fidelity – Maximum Location Error .....	7
Figure 3: Location Fidelity – The "Acceptable" Circle .....	8
Figure 4: Location Fidelity Example .....	11
Figure 5: Latency Implementation .....	16
Figure 6: Latency Sample Set Histogram with a Bin Size of 3 Seconds .....	19
Figure 7: Location Fidelity – Expected Relative Error .....	21
Figure 8: Currency – Expected Relative Error .....	22
Figure 9: Latency – Expected Relative Error .....	23
Figure 10: Network Measures – Expected Relative Error .....	24
Figure 11: Average Latency per Node (Three Scenarios) .....	26
Figure 12: Location Fidelity (Scenarios #3 and #4) .....	27
Figure A.1: Folders and Files Structure of a Run .....	31
Figure A.2: Input Parameters in the "GenerateExcelFiles.sql " Script .....	32
Figure A.3: Example of "LF.xls" file .....	34
Figure A.4: Example of "CURRENCY_LOG.xls" File .....	35
Figure A.5: Example of "LATENCY_LOG.xls" file .....	35
Figure A.6: Average Latency (in Seconds) per Node .....	45
Figure B.1: "GenerateExcelFiles.sql" Script (Part I) .....	46
Figure B.2: "GenerateExcelFiles.sql" Script (Part II) .....	47
Figure B.3: "GenerateExcelFiles.sql" Script (Part III) .....	48
Figure B.4: "GenerateExcelFiles.sql" Script (Part IV) .....	49

Figure C.1: Graphical User Interface – Part 1 .....	50
Figure C.2: Graphical User Interface – Part 2 .....	51
Figure C.3: Graphical User Interface – Part 3 .....	52



## List of tables

---

Table 1: Example of Currency Measurement.....	12
Table 2: Average Latency per Scenario.....	26
Table A.1: Example of Location Fidelity Output.....	41
Table A.2: Example of Currency Output.....	44

## Acknowledgements

---

The authors wish to acknowledge the contributions and express sincere gratitude to the members of the HCTCN TD project team, particularly Dr. Allan Gibb and Mr. Jean-Claude St-Jacques, for providing guidance and sharing their knowledge and expertise throughout the entire study, as well as for their careful revision of this report. The authors would also like to extend their thanks to Mr. Phillippe Plante and Mr. Frédéric Nadeau for their much appreciated help during the development phase of the VBA tool, and for the identification, implementation and validation of the distance formula used for the location fidelity measurements.

# 1. Introduction

---

## 1.1 Tactical Communication Environment

On the tactical battlefield, the means of communication between vehicles or dismounted soldiers is generally a line-of-sight radio operating in the VHF or UHF band. The advent of digital command and control systems in the tactical domain offers the promise of increased battlefield awareness through systematic and automated distribution of relevant data. To deliver on this promise, it must be possible to distribute digital data over the radio system with high fidelity and with a timeliness appropriate to the operation. Initial experience with transmitting tactical data in digital form over Army radios has highlighted the inadequate data capacity of these wireless channels [1].

Sharing information in digital form on an all-informed basis is highly desirable to avoid a single point of failure and to ensure continuity of operations. Under this exchange model, nodes aim to maintain exact copies of key information in each other's database via asynchronous data replication. When communication channels have low and variable throughput and unreliable connectivity, maintaining full synchronization is next to impossible. It is likely that, over time, the databases will drift out of synchronization. Users may believe that they are sharing the same situation picture, when, in fact, they are not. If detected, such a characteristic can undermine confidence in the system; undetected, it may have deadly consequences.

## 1.2 High Capacity Technical Communications Network

In April 2000, a Technology Demonstration Project<sup>1</sup> (TDP) entitled “High Capacity Technical Communications Network” (HCTCN) [1] was initiated. The objective of the project was to demonstrate the potential for selected technologies in wireless communication networks and information management to increase the limited capacity of tactical communication systems to support command and control requirements. The project featured three areas of investigation: tactical networking, high capacity radio bearers and information management. For the information management segment of the project (for which DRDC Valcartier had responsibility), a research test bed concept called the Low Bandwidth Test Bed (LBTB) was conceived by project staff at DRDC Valcartier and implemented through contract, to evaluate battlefield information management strategies applied in a low bandwidth tactical wireless communications environment. As part of the technical specifications for this contract [2], the capacity to support three basic measures of performance was defined for inclusion in the LBTB. These were:

---

<sup>1</sup> The overall objective of a TDP is to demonstrate technologies fostered by DRDC and Canadian industry in the context of real and potential future Canadian Forces capabilities, concepts, doctrine, operations, and equipment. The TDP is aimed at concept development and evaluation for force design purposes and is therefore typically not focused on hardware development.

- Currency – Time elapsed since the last update of a given piece of information in a node’s tactical database;
- Latency – Time taken to transfer a data replication (messages and information of any type) between two nodes;
- Data Consistency – Comparison of data values in designated table row(s) or in column(s) in a table row, in two or more different databases to determine whether data values are equal.

During the course of the contract, the data consistency measurement was dropped due to the technical complexity of its implementation. A fourth measurement, location fidelity, was added to the specifications after the contract had begun:

- Location Fidelity – Comparison between the real location of a node and its perceived location by other tactical nodes.

### 1.3 LBTB Overview

LBTB is a synthetic environment that simulates the use of a tactical command and control system across a single radio tactical communications network. The LBTB application enables the user to define scenarios and execute them in a controlled manner. The user has the capability to control a variety of parameters for each execution, including the following:

- Scenario script used;
- Type of radio being simulated;
- Quality of the radio network;
- Amount of traffic being broadcast;
- Amount of data replication that will occur between the various nodes in the network;
- Parameters of the Replication Transport Layer; and
- Measurements taken.

Given the number of possible outcomes that may be generated for a scenario script, it is important for the user to be able to capture the state of the simulation for subsequent analysis. The user is able to generate data captures at any point during the course of an execution, or at the very end of the process. In addition, there are a variety of statistics

relating to the data that were broadcast across the network that are being collected and stored in various databases.

As the scenario is executed, the nodes update their local command and control databases with position updates, enemy contact reports and the like. Intelligent rules residing on each node determine which of these local updates, are replicated over a simulated combat radio network (the radio communication system simulator) to other nodes.

## **1.4 Operational Research Support to HCTCN**

In autumn 2004, members of the HCTCN project requested support from the DRDC Valcartier Operational Research (OR) team to address the question of the number of times each scenario would have to be executed with the LBTB to ensure statistical validity of the results.

The involvement of the OR team has gone beyond the initial scope. The OR team was also asked to provide an analysis of the existing Measures of Performance (MOPs) implemented in the LBTB, to propose refinements to the existing MOPs to better suit the detailed analysis requirements of the HCTCN TDP, and to take the lead in building a software tool to implement the refined MOPs. These refinements included several possible interpretations of the Location Fidelity measure as well as the consideration of more aspects such as transmission priority and replication message type. This ultimately led to a significant improvement of the ability of the existing MOPs to evaluate the impact of the battlefield information management strategies. More precisely, the improvement consisted of providing a way to analyze the performance of the system from a global perspective, which required finding a means to summarize in a meaningful way the large quantity of information (pre-defined measurements) that were collected during the simulation and stored in various log files within LBTB.

To summarize the contributions of the OR team, the report is organized in five sections. Following this introduction section is Section 2, which contains a general discussion on MOPs and describes the refined MOPs that were proposed to evaluate various battlefield information management strategies. The question regarding the number of runs required to ensure the statistical validity of the results is tackled in Section 3 through a method that correlates the expected relative error to the number of runs executed. The following section presents two examples of scenarios that were executed with the LBTB and analyzed using the MOPs presented in this paper. Section 5 presents the main conclusions arising from the OR support. Finally, this document includes three annexes: a user's guide (Annex A), the Structured Query Language (SQL) scripts developed to retrieve the information from the database (Annex B) and the Graphical User Interface (GUI) of the tool (Annex C).

## 2. Proposed Improvements to Measures of Performance

---

### 2.1 General Thoughts on Measures of Performance

The information exchanged between nodes can have various levels of importance. This is handled in LBTB through the concept of priority level that is associated to a given piece of information and which allows for a distinction to be made between information that is critical, important, or just desirable. It should also be noted that the level of importance of a given piece of information may also change with time. For example, information about the position of friendly nodes is very important when it is recent but it becomes obsolete beyond a certain age. Good measures of performance should take into account the fact that all information is not equal. This will be especially important when evaluating the various information management rules. Therefore, measures should not be global, but specific to a given importance level (or priority level) of the information being exchanged.

Most real life military missions can be described as having multiple phases. During the first phase, there is generally relatively little going on, then there is a build-up towards the peak of the conflict, followed by the peak phase itself characterized by a significant amount of activity, and finally a defusing phase of the conflict until the mission is considered completed. The communication needs will vary according to the phase of the conflict. This should be taken into account when computing measures of performance. As such, meaningful measures of performance should be calculated for each phase separately, rather than globally for the whole duration of the mission. This will avoid the phenomenon of global averaging of data, a phenomenon that can mask important effects within a phase and can produce misleading results.

Measures should be simple to understand and meaningful for the military. There are several types of measures of performance that could be envisaged, depending on the type of effects that one would like to capture or on the phenomena that one would like to better understand. If we were to look at the problem in terms of cause and effect<sup>2</sup>, a fair assumption to be made is that less than perfect communication between vehicles will cause some undesirable operational effects. However, although easily acceptable, this assumption has little practical value, unless a good set of measures is implemented. A good set of measures would be one that would allow for a better

---

<sup>2</sup> The terms of “cause” and “effect” should be interpreted loosely. They are just labels that are used in this paper to convey an intuitive view of the system: things that have to do with technology and implementation of the communication network will be considered under the label “cause”, while things that have to do with what is happening in theatre and have a direct bearing on the humans involved will be referred to as “effects”. One should avoid getting caught into semantics issues such as the fact that a certain effect can be considered the cause for another effect, since the point is not to describe the ultimate effect, nor the primary cause. The point is rather to suggest an intuitive but methodical way to look at the system as a whole and understand that a good comprehension of the system can only be obtained through adequate probing, or, in other words, through a good set of measures.

understanding of the operational impacts (which includes first of all recognizing these impacts and second providing a way to evaluate them), as well as a good understanding of the potential causes. One should aim for a balanced set of measures, some focusing on the causes and some focusing on the effects, which will allow for a better understanding of the system as a whole and ultimately improving the system.

In this paper the authors analyze a set of three basic measures implemented in LBTB and propose refinements to these measures. The first one, *location fidelity*, described in subsection 2.2, deals with one of the possible effects of less than perfect communication, namely the fact that a vehicle is not always aware of the correct location of the other friendly vehicles. The second measure, described in subsection 2.3, is *currency*, tackling another potential effect of imperfect communications, namely the fact that the information available to a vehicle is not necessarily the most current available. Finally, the third measure, *latency*, explained in subsection 2.4 is addressing one of the potential causes for problems, namely the delays that can appear when transferring information between nodes.

## 2.2 Location Fidelity

The measure of "Location Fidelity" is intended to capture the fact that, due to less than perfect communication between vehicles, there could be a gap between the real position of a given vehicle A and the perceived position (i.e., the last known position) of that vehicle by another vehicle B. Measuring this gap is important because it allows for a quantitative assessment of the situational awareness experienced by the vehicles in the field, which is key in estimating the risk of friendly fire, for example.

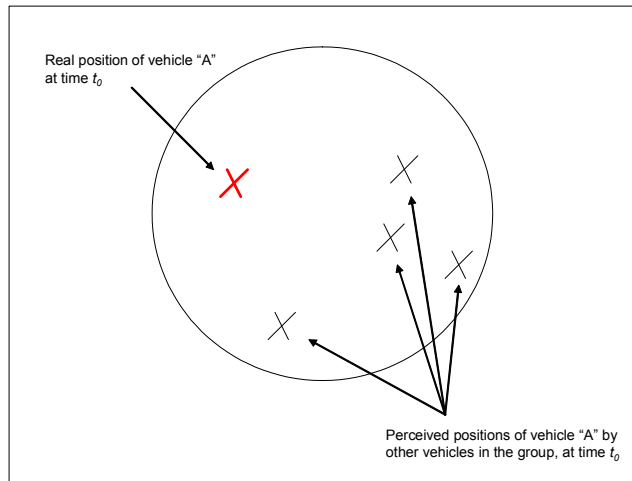
There are multiple ways of capturing this phenomenon. The more basic ones will attempt to describe the phenomenon through some quantitative measure. Others can provide some "value added" by attempting to describe the phenomenon in such a way to allow for actions to be taken to improve the system. So the first category is mainly for information purposes, while the second one provides information that one can act upon to make the system better.

The authors proposed a total of four interpretations for location fidelity to the project sponsors, for their final selection. These measures, some belonging to the first category, some to the other, will be described in the following paragraphs in general terms. Upon discussion with the sponsors and consideration of the implementation challenges raised by the measures proposed, a "compromise" was reached. The formulae actually implemented will be discussed later in the document. Still, all proposals were included here since they may prove useful in case LBTB will be used in other projects, for example.

### 2.2.1 Smallest Circle that Includes All Points

The idea of this measure is to provide a way to quantify how reliable the location information is within the network. Perfect location information would mean that the real position of a given vehicle at a given point in time is known by all other vehicles.

Imperfect information would mean that the real position of a given vehicle may not be known by some of the other vehicles which may cause them to believe the vehicle is in a different place. Figure 1 illustrates this case of imperfect communication.



**Figure 1:** Location Fidelity – Smallest Circle that Includes All Points

Consider the situation of a certain node of the network (say, vehicle A) being situated in the position marked in red in Figure 1. This position (the real position) is certainly known by the node itself and maybe by some other nodes. However, in a case of imperfect communication, there may be nodes that perceive node A in different locations (the perceived or last reported positions). Therefore, for each given node, at a given point in time, one can imagine a circle (or a sphere) that includes all points representing the real and perceived positions of that particular node and which is the smallest one that has this property. The smaller the radius of this circle, the more reliable the location information is (i.e., the general knowledge about the position of that vehicle is quite precise, the image is "sharp"). The larger the radius, the less reliable the location information is (i.e., the general knowledge about the position of that vehicle is less precise, the image is fuzzy).

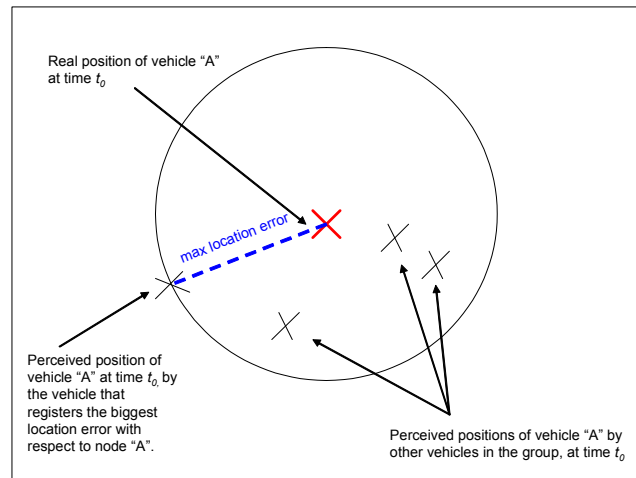
Therefore, the radius of the smallest circle that includes all points representing the real and perceived positions of a given node could be one way to describe the consequences of less than perfect communications, in terms of location fidelity, for a given node, at a given point in time. However, to describe the whole network and not only one node, another measure would be needed. The authors proposed calculating either the average radius (as defined earlier) for all nodes of the network, or, alternatively, monitoring the largest radius (which basically amounts to monitoring the worst case). Such measures would provide a good way to describe the consequences of less than perfect communications, in terms of location fidelity, for the network, at a given point in time. Finally, the same basic measure of the radius of the smallest circle could be used to describe the performance, in terms of location fidelity, for the network, over a given period of time (which could be one phase of the mission, for example). To do that, it would be sufficient to do several measurements at various points in time during the desired period of time and consider the average radius based on all of these measurements (or, alternatively, the largest radius).



### 2.2.2 Maximum Location Error

This measure is very similar to the first one proposed, the only difference being that in this case the smallest circle that includes all points (real and perceived positions of a given node) has the real location of the node as its centre (while for the first measure the real location could have been anywhere inside the circle or on its circumference), as shown in Figure 2. As such, this measure shows the maximum error, how far the perceived location of a node can be from its real location. This measure has the advantage (over the first measure) of being more "natural", or easier to grasp, but the two measures are in fact very similar.

The Maximum Location Error, as discussed, is defined for a given node, at a given point in time. Variations of this measure for the whole network at a given point in time and for the whole network over a given period of time can be obtained in a similar way to those suggested for the previous measure proposed for location fidelity.



**Figure 2: Location Fidelity – Maximum Location Error**

### 2.2.3 The "Acceptable" Circle

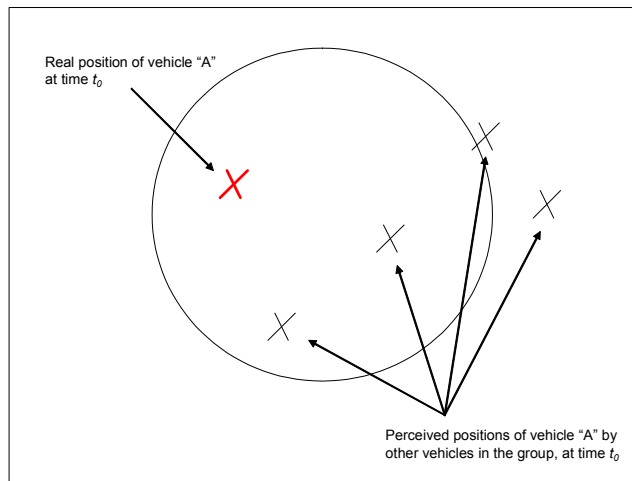
The idea of this measure is to describe the reliability of the location information within the network with respect to an acceptable threshold. Basically, if it could be decided what is "good enough" in terms of location fidelity, then some measures could be imagined to describe the performance of the system as either "good enough" or to quantify how far we are from the acceptable threshold.

An example of such a measure, for a given node at a given point in time, could be to simply count how many points representing the real and perceived positions of that particular node are inside the "acceptable" circle and how many are outside this circle and calculate the ratio (see Figure 3). The "acceptable" circle is defined as having its center in the real location of a given node and a radius equal to a pre-defined value suggested by military experts, according to their vision of what is an acceptable error in terms of the location of a vehicle. It should be noted that, depending on the activities

carried out at various moments during the mission, the tolerance to errors can vary. For example, in a firing situation the acceptable error will likely be very small (i.e., frequent position updates will be required), to reduce the risk of friendly fire. In situations with little or no risk involved, the tolerance to location errors will likely be larger (i.e., less frequent updates will be necessary).

A central idea explored in the information management segment of the HCTCN TDP is the use of application-layer information management rules that can modify the way that information is packaged, and can automatically suppress replication of certain information, in response to network conditions. The rule implemented for the HCTCN TDP experiments is one which suppresses replication of position information for any vehicles moving more slowly than a certain speed. Developing information management rules that would automatically change the frequency of position updates during the various phases of the scenario based on some criteria could also improve the system by limiting the amount of unnecessary traffic (i.e., more position updates than are really necessary during a certain period of time).

For example, if during some phases of the scenario it is not so important to have good location information and the Information Management (IM) rules recognize this and order less frequent location updates, this measure is flexible enough not to penalize the fact that more points would fall outside the circle. It will simply adjust the radius of the circle and display the proper result – that the situation is satisfactory, under the circumstances, rather than in some absolute way. As such, this measure provides some value added when compared to the previous one, because it can be used to evaluate some IM rules, and not just give a portrait of how good or bad the location fidelity is.



**Figure 3: Location Fidelity – The "Acceptable" Circle**

The measure of the acceptable circle, as discussed, is defined for a given node, at a given point in time. Variations of this measure for the whole network at a given point in time and for the whole network over a given period of time can be obtained in a similar way to those presented for the first of the measures proposed for location fidelity.

## 2.2.4 The Location Knowledge Ratio

This measure will be first described for a given node, at a given point in time. The idea of this measure is to capture how much a given node knows about the location of all of the other nodes, at a given point in time. In other words, it is meant to capture the degree of confidence that the operator of a vehicle can have in the (location) information that is available to him for making decisions. A measure that could potentially serve this purpose is the ratio of the number of nodes that a given node knows the exact position of, over the total number of nodes in the network. Note that knowing the "exact" position of a node at any point in time is probably impossible, given the inherent communication delays that exist in any communication network. Therefore, for the purpose of calculating this measure, "exact" knowledge should be replaced by "reasonably good" knowledge, which in turn should be clearly defined mathematically through a precision factor. For example, if operational requirements dictate that location knowledge within five meters is sufficient, then the Location Knowledge Ratio will be defined as the ratio of the number of nodes that a given node can locate within five meters, over the total number of nodes.

As such, if this ratio is 1, then a node knows (reasonably well) where all of the other nodes are located. A link could be made between the value of this ratio and the risk of friendly fire. The closer the ratio is to 1, the smaller the risk of friendly fire. A whole range of other decisions (besides the decision to fire) could be affected by the level of knowledge that an operator has regarding the location of the other friendly vehicles, so this measure as the one proposed here could be very valuable.

It can be of interest to take several such measurements for a given node, over a period of time. The resulting graph would show how this ratio changes over time, depending on the phase of the conflict and on the communication requirements at the different stages of the mission.

A variation of this measure can also be used to describe the network, over a period of time (as opposed to a single node). An example of such a measure could be to calculate the percentage of nodes that have a ratio of more than a specified value for more than 90% of the time. This will provide a good description of the level of knowledge that vehicles in the network have about each other with respect to their geographical positioning.

## 2.2.5 Location Fidelity Implementation

Upon discussion with the sponsors, and considering the type of analysis they were interested in conducting, along with the implementation challenges and computation time associated to the various location fidelity measures considered, an agreement has been reached to implement the *average location fidelity*, as described further.

As a scenario is being executed in the LBTB, information about location fidelity measures is recorded in two different tables: "loc\_fidelity\_log" and "loc\_fidelity\_log\_postions" [3]. At the end of the execution, a SQL script cross-

referring these two tables is automatically called to create a view that summarizes the location fidelity data into a single table named "Loc\_Fidelity\_Measurement".

The view contains the following columns: the logging time (the time when the measurement was taken), the observing node, the observed node, latitude1 and longitude1 corresponding to the position where the observing node believed the observed node was at the time the measure was taken and finally, latitude2 and longitude2, which represent the true position of the observed node at the time the measure was recorded. The only information missing in order to implement the average location fidelity is the actual error between the perceived and the true locations. This error corresponds to the distance coordinates (longitude1, latitude1,) and (longitude2, latitude2). Note that the distance formula that was implemented is named Haversine, which is discussed in depth in [4]. This formula is particularly well-conditioned for numerical computation even at small distances, unlike calculations based on the spherical law of cosines.

Before the execution of a scenario in LBTB, the user specifies the frequency at which the location fidelity measurements will be taken. For instance, consider a 60-minute long scenario, including 5 nodes and where location fidelity measurements are recorded every 5 minutes. It means that for this example, there would be a total of the 240 entries (12 times 20) in the "Loc\_Fidelity\_Measurement" view. This file can become very large, especially considering that the members of the TDP envisioned 2-hour scenarios with 20 nodes with measurements taken every minute.

The location fidelity measurements implemented are given as a confidence interval, i.e. as an estimated range of values which is likely to include the error parameter (in this case, distance between believed and true location). Given a time interval (scenario start time and end time) input by the user, these confidence intervals are calculated every time the location fidelity measures are recorded in the log file, for instance, 12 times (60 divided by 5) in the example presented in the previous paragraphs. Confidence intervals are usually calculated so that this percentage is 95%, but we can produce 90%, 99%, 99.9%, confidence intervals for the unknown parameter. More details about confidence interval and confidence level are given in [5].

Due to the large quantity of information provided by the location fidelity measure, there are many graphs that can be drawn from the outputs. An example of a graph showing the evolution of the location fidelity measurement as a function of the time in the scenario is presented in Figure 4. In this example, the average location fidelity was calculated every minute during the 45 minutes scenario (from 7:15 to 8:00). Similar graphs can also be drawn for individual nodes versus the others nodes (how good or how bad a node perceived to other ones) or even for one node versus another node. All the details about the location fidelity measurements, inputs, outputs and examples are included in Annex A and Section 4.

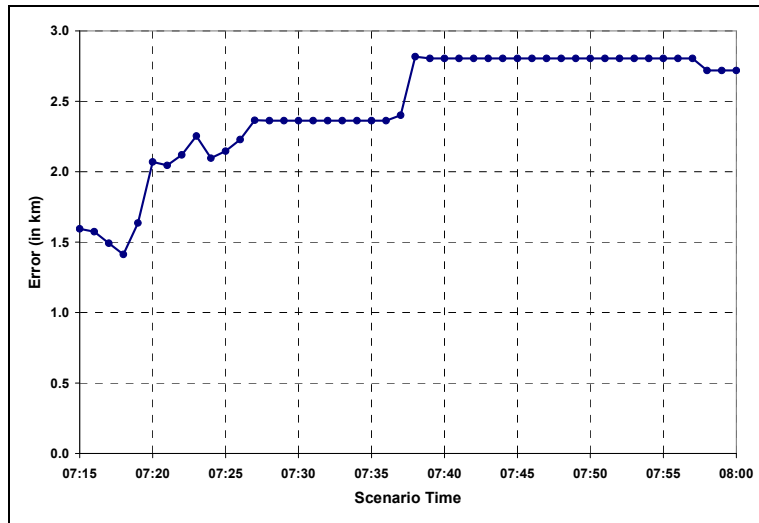


Figure 4: Location Fidelity Example

## 2.3 Currency

### 2.3.1 Discussion

Another side effect of imperfect communications within the network is that the data available to a given node may not be the most current. Since decisions can only be made based on the available data, lack of currency of the available data will clearly have an impact on the quality of these decisions. Decisions that are good in the context of old data (which is no longer accurate) may prove to be bad decisions in the actual (changed) context. Therefore, measuring and monitoring this parameter becomes very important.

The currency measure initially implemented within LBTB records the time elapsed since the last successful update of a given piece of information and compares the elapsed time with a threshold set by the user. However, since the updates are not necessarily sent at regular intervals, the authors found the comparison with a pre-set threshold not relevant, and dropped it from the final implementation of this measure, while keeping the basic idea of measuring the time between two successful updates.

The following example will help clarify the measurement process that is being carried out to compute the currency measure. Let us assume a network of four nodes where node #1 sends four updates of its position to the other three nodes. Node #1 will be called the source node. The other nodes are the destination nodes, also referred to as local nodes. The updates can be successfully applied to the tactical databases of the destination nodes, or not. Suppose the situation synthesized in Table 1 is occurring, where "X" indicates that the update was not successfully received by the destination node.

**Table 1:** Example of Currency Measurement

Update #	Time update is sent by the source node #1	Time update is received by local node #2	Time update is received by local node #3	Time update is received by local node #4
1	$t_1$	X	$t_{13}$	$t_{14}$
2	$t_2$	$t_{22}$	$t_{23}$	X
3	$t_3$	X	$t_{33}$	X
4	$t_4$	$t_{42}$	$t_{43}$	$t_{44}$

Then the following measurements, representing the time elapsed between two consecutive successful updates, will be recorded:

For node #2:  $(t_{42} - t_{22})$  (one measurement for node #2)

For node #3:  $(t_{23} - t_{13}), (t_{33} - t_{23}), (t_{43} - t_{33})$  (three measurements for node #3)

For node #4:  $(t_{44} - t_{14})$  (one measurement for node #4)

These are the basic measurements that are used to compute the various currency measures that were selected for implementation by the authors, upon consultation with the project sponsor. The important things to note right now are that those measurements make sense if they are considered by replication type (such as "Position Update", for example), during a given time interval, and that they describe currency issues related to a given pair of nodes (the source node and the destination, or local node).

While it could be of interest to evaluate the currency of the information exchanged between a given pair of nodes, it could also be of interest to have a more global appreciation of the currency, describing the whole network. As such, in anticipation of several types of analysis that will be required in the future, several currency measures were implemented, covering the case of a given pair of nodes, the case of a unique source node and multiple destination nodes, up to the case of multiple source nodes and multiple destination nodes.

### 2.3.2 Currency Implementation

The proposed implementation of currency uses the information recorded in a table called "currency\_log". This table is generated as the scenario is executed in the LBTB. Entries are added in the table when a successful update of a given piece of information occurs during the simulation. Columns in this table include: record identifier, time since last update (i.e. currency values), the source node, the local node, and more [3].

It was pointed out in the previous subsection that the currency measurements make sense if they are considered by replication type, during a given time interval, and that they describe currency issues related to a given pair of nodes (the source node and

local node). The information in the "currency\_log" allows for a consideration of all these factors with the exception of the replication type (or event type). To implement the currency measure in the way the authors envisioned it, an additional step, consisting of executing a SQL script (steps to follow and script are presented in Annexes A and B), is necessary to associate a replication type to the records in "currency\_log".

Upon discussions with the sponsors and considering the type of analysis to be conducted, it was decided that six currency measures will be implemented. Given a local node *LN* and a source node *SN* input by the user, the currency measures calculated are as follows:

Measure A: The average of currency values for all the rows with local node (Column C) equal *LN* and source node (Column D) equals to *SN*.

Measure B: The average of the currency values for all the rows with source node equals to *SN* and local node different from *SN*.

Measure C: The average of the currency values for all the rows with local node equal to *LN* and source node different from *LN*.

Measure D: The average of the currency values for all the rows where local node is not equal to the source node.

Measure E: The average of the currency values for all the rows where local and source nodes equal to *LN*.

Measure F: The average of the currency values for all the rows where local node equals source node.

Note that the aforementioned measures are calculated based on a given time interval and a replication type, as specified by the user. For each of the currency measures A, B, C, D, E and F, the total number of rows and the sum of the currency values of the rows that meet the criteria specified through the input parameters (replication type, start time, end time, etc.) are also calculated. The currency values are expressed in terms of seconds. Implementation details and currency example are given in Annex A.

## 2.4 Latency

### 2.4.1 Discussion

In the current LBTB implementation, the *data latency* measurement is defined as the time taken to transfer a data replication message between two nodes. The process of transferring data can be broken down into packaging the data at the source node, sending the packets over the network and finally unpackaging the data at the destination node. As such, the sum of the durations of each of these processes defines the end-to-end latency, or simply *latency*.

It is expected that, under conditions of significant traffic (high congestion), the latency will be high. This is mainly due to the limitations imposed by low bandwidth. Messages will have to wait in a queue before they can be sent over the network, leading to high values for latency. Such high values translate into nodes not having the most current data, which ultimately translates into poor situational awareness.

The *raison d'être* of the information management segment of the HCTCN TDP is to demonstrate that by using adequate information management strategies, it is possible to alleviate some of the communication problems associated with a low bandwidth environment. Some of the information management rules to be implemented aim to improve network communication by favouring the transmission of important information (replications) over less important information. In terms of priorities, under conditions of high congestion, the critical items will be sent in priority, to ensure that, when not all information can get through, at least the most important pieces do (i.e. the right information, at the right time).

Latency values depend on several factors including offered load, channel access protocol, or bit error rate on the channel. Latency appears to be a suitable measure to evaluate especially those IM rules that reduce the offered load by suppressing the transmission of unnecessary information or that provide preferred channel access to information of higher priority. Such rules will likely decrease the latency of high priority items, ensuring that the important information gets through in the shortest amount of time possible.

Depending on the purpose intended for the analysis, there are several “variations” of this measure that could be envisaged, starting with the basic definition of latency as being the delay between when a replication is sent by the source node and when it is received by the destination node(s). For example, latency values could be calculated by individual node (with a distinction possible between sending nodes and receiving nodes) or for all nodes in the network. In each case, calculations can be done by replication type, by priority, for all replication types and a given priority, for one replication type and all priorities, and finally for all replication types and all priorities.

It is important to note that the *latency* measurement only captures information that is related to successful replication transfers. Information regarding the unsuccessful attempts is not reflected in any way in this measure. In other words, latency is a good measure to describe one aspect of the network performance, namely how well things worked when they actually worked. For this reason, it should be used in conjunction with other measures to obtain a more comprehensive description of the performance of the communication network.

## 2.4.2 Latency Implementation

The data used for the proposed implementation of the latency measure are contained in a table named "latency\_log", populated during the execution of a simulation with the LBTB. In this table, there is a log entry for each destination node identifier. Since in LBTB each transmission is broadcast to every node, a separate log entry is created when it is received by the destination node. It is therefore possible to determine which

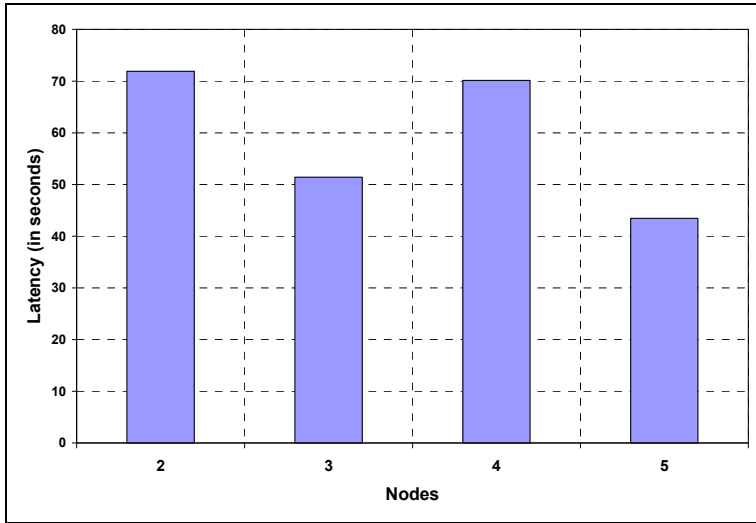


node received the replication message successfully. Columns in the table include: source node, destination node, time when the transmission was sent on the network, transaction time, and more [3].

Two important aspects, message priority and replication type, discussed in the previous subsection, are not available in this table. For this reason, an additional step, consisting of executing a SQL script (steps to follow and script are presented in Annexes A and B), is necessary to associate message priority and replication type.

In the latency implementation, latency values are calculated by individual node, either from a sending node's point of view or from a receiving node's point of view. A global latency value is also calculated; this corresponds to average latency values for all the transmissions that were successful between any source and destination node during the simulation. The calculations are based on a given time interval, a set of replication types (one or more) and a set of message priorities (one or more), as specified by the user.

An example of latency output is presented in Figure 5. The chart should be interpreted as follows: for each node, the latency value corresponds to the average latency of the rows that meet the input parameters specified by the user (replication type, priority type, start time, end time, etc.), either from a sending node's point of view or from a receiving node's point of view. In the example, if the analysis was performed from a sending node's perspective, for the latency corresponding to Node 2, only the records that have Node 2 as source node would have been considered. Otherwise, if from a receiving node's point of view, only the rows where destination node is 2 would have been included in the latency calculations of Node 2. Note that a global value for the latency is also calculated. It corresponds to the average latency (a weighted average) for all the nodes in the scenario (in the example). All the details regarding the latency implementation, inputs, outputs and examples are presented in Annex A and Section 4.



**Figure 5:** Latency Implementation

### **3. Number of Runs Required to Ensure Statistical Validity of the Results**

---

Simulation results analysis involves computing estimates of the parameters of interest using the data generated by the simulation model. The greater the number of times the simulation model is executed, the greater the sample size from which to compute those estimates, the greater the confidence in the results. Ideally, a model should be run enough times to ensure convergence of the results.

In some cases, and the LBTB case is one of them, the simulation time corresponding to one execution of the model is quite large (the number of scenarios to be evaluated is also quite large), which limits the number of runs that can reasonably be done with the resources available. Hence, the question of interest becomes "How many times should the model be executed so that the level of confidence in the results is acceptable?" In addition, sometimes an acceptable convergence cannot be obtained due to the large number of required repetitions. In this case, the level of confidence needs to be specified when reporting results so as to provide the means to evaluate the practicality of the results.

In the context of HCTCN, the parameters of interest are the location fidelity, the latency and the currency measurements. The analysis presented in this section gives an idea of the relative error to expect, for each measure, as a function of the number of runs executed with the test bed. In general terms, the relative error corresponds to the ratio of the absolute error of a measurement to the best estimate. The relative error is especially useful in judging the reliability of estimates.

The members of the TDP also asked the OR team to consider in the analysis of the number of runs, four other network measures: RTL-RTL Transmission (Tx) Success Rate, Channel Access Success Rate, Channel Reception Success Rate and Average ReTx/Tx Ratio. These were not developed by the OR team and therefore they will not be explained in this report.

#### **3.1 Stochastic Nature of LBTB**

In the LBTB, the output data is typically stochastic, because the simulation model derives some values based on random number generators. The aspects of LBTB that introduce randomness are limited though. For instance, in a given scenario, there is no stochastic aspect to the sequence and the timings of the events (position updates, contact reports, etc.) since they are all predefined by the user. This choice was made to create a controlled simulation environment.

The probabilistic component comes from (1) the radio communication system model to account for the fact that radio transmissions sometimes fail, and (2) the channel access protocol. In LBTB, the user can specify levels in the quality of radio transmissions by selecting the bit error distribution. If the percentage of bit error is set

to 0 %, then, there would be no transmission error. In such a case, the model becomes deterministic, i.e. there would be no variability in the transmission results between two consecutive runs of the same scenario.

Consider the case where the percentage of erroneous bits is a value greater than zero. The implementation would be a uniform distribution initialized with a random seed and an array of the percentage of erroneous bits for each state.

This module would increment by 1 bit at a time through the data transmitted and determine if each bit should be erroneous by calling the module which performs a calculation on the uniform distribution indexed by the current position. If the uniform distribution returns a value higher than the percentage of erroneous bits, then the bit is not erroneous, otherwise this bit is erroneous and a bit error must be applied by flipping the bit's value.

The radio transmission errors can also be simulated at a higher level, i.e. at the packet level (array of bits) rather than bit-by-bit. In such case, the stochastic nature of the LBTB results comes from the fact that random numbers that are generated to determine if individual packets of information are loss or not.

The user can set the random number seed that will be used for all the internal calculations. Each random number seed will generate a different set of voice and data transmissions. Running the same scenario consecutively with the same random seed would provide, in principle, the exact same results, i.e. no variability. In practice, some slight variability may be observed because the simulation of the communication system is linked to a model of the user application that employs real databases. Slight variations in the real-time interaction times with the databases on successive simulations may introduce minor variability in the results.

## 3.2 Discussion on the Variability

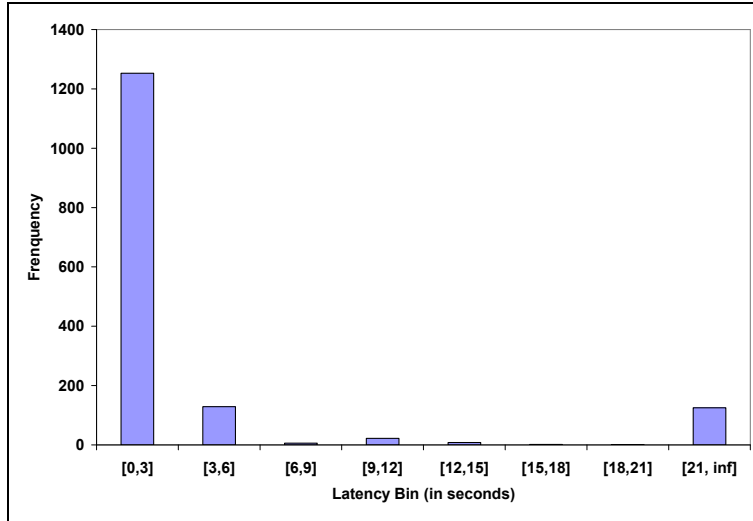
Suppose that the user is interested in one particular measure  $L$ , say latency. In the LBTB, when a scenario is executed once, a series of  $m$  latency measures  $l_{1,1}, l_{1,2}, \dots, l_{1,m}$  are recorded. In this case, the measures correspond to the latency values, i.e. the times taken to transfer the replications between two nodes during Run 1. The latency measure for Run 1 is determined using a sample mean, denoted  $\bar{L}_1$ , given by the following equation:

$$\bar{L}_1 = \frac{\sum_{i=1}^m l_{1,i}}{m}$$

In this case,  $\bar{L}_1$  can be seen as a random variable following a probability distribution dictated by the scenario (list of events, stochastic/randomness, settings, etc.). Before drawing any conclusions about the value  $\bar{L}_1$ , it is important to consider the distribution

and variation in the data used to calculate  $\bar{L}_1$ , i.e. the variation in the latency measures  $\{l_{1,1}, l_{1,2}, \dots, l_{1,m}\}$ .

Histograms of the latency measures for several sample sets were prepared to identify the nature of the probability distribution governing the latency. One of these histograms is shown in Figure 6. It was obtained from a simple scenario including 12 nodes, reporting their position every minute. From the figure, it is clear that the data does not follow a normal distribution. In the data presented in Figure 6, the mean is 19.11 and the standard deviation is 58.30, suggesting that the latency values vary widely.



**Figure 6:** Latency Sample Set Histogram with a Bin Size of 3 Seconds

Consequently, each scenario would have to be executed many times in order to increase the precision of the estimation of the population mean. When the same scenario is repeated  $k$  times,  $k$  sample means are obtained  $\bar{L}_1, \bar{L}_2, \dots, \bar{L}_k$  (one for each run).

To estimate the latency mean value corresponding to the scenario, the average of  $\bar{L}_1, \bar{L}_2, \dots, \bar{L}_k$ , noted  $\bar{L}$ , should be calculated. The fundamental question here is to determine the value of  $k$  such that the confidence interval of  $\bar{L}$  would be acceptable.

Since the random variables are not normally distributed, the study estimated the expected error as a function of the number of runs, a method based on the central limit theorem. The key idea encompassed in the central limit theorem is that when a population is repeatedly sampled, the average value of the attribute obtained by those samples is equal to the true population value [6]. Furthermore, the values obtained by these samples are distributed normally about the true value, with some samples having a higher value and some obtaining a lower score than the true population value. In other words, when the samples become very large, then the sample means will follow a normal distribution even if the respective variable is not normally distributed.

The following subsections present the method used to estimate the expected relative error as a function of the number of run, and the results of the method applied to all the measures.

### 3.3 Method for Estimating the Sample Size

Consider the following definition: a terminating simulation is one that runs for some duration  $T_E$ , where  $E$  is a specified event (or set of events) that stops the simulation. Such a simulated system starts at time 0 under well-specified initial conditions and stops at the time  $T_E$ . The LBTB is an example of a terminating simulation.

The method used in this report to estimate the number of runs of a terminating simulation, such as LBTB, is the method of independent replications explained in [7]. As indicated before, the method is based on the central limit theorem. The question is, if one wants to estimate a parameter  $\theta$ , from simulation output data  $\{Y_1, Y_2, \dots\}$ , with an accuracy criterion  $\varepsilon$  and a high probability  $1 - \alpha$ , how many runs  $R$  of the scenario should be executed? The method consists of the following seven steps:

1. Set  $\alpha$ .
2. Execute the simulation a total of  $R_0$  times, each run using a different random number seed and independently chosen initial conditions, which includes the case that all runs have identical initial conditions. In this analysis, 30 runs ( $R_0=30$ ) were executed.
3. Estimate the sample mean  $\hat{\theta}$  using the following equation:

$$\hat{\theta} = \frac{1}{R_0} \sum_{r=1}^{R_0} Y_r$$

4. Calculate an initial estimate  $S_0^2$ , of the population variance:

$$S_0^2 = \sum_{i=1}^{R_0} \frac{(Y_i - \hat{\theta})^2}{R_0 - 1}$$

5. Assume that if  $\hat{\theta}$  and  $S_0^2$  are valid for  $R=30$ , therefore  $\hat{\theta}$  and  $S_0^2$  are also good estimates of the average and the standard deviation for other values of  $R$ . Calculate the expected relative error  $\varepsilon_R$  as a function of  $R$ , using the following equation:

$$\varepsilon_R = \left( \frac{t_{\alpha/2, R-1} S_0}{\sqrt{R}} \right) / \hat{\theta}$$

The value  $t_{\alpha/2, R-1}$  is taken from the student distribution table. Note that when  $R$  is greater than 30, the value of  $t_{\alpha/2, R-1}$  is replaced by  $z_{\alpha/2}$ , taken from the normal distribution table.

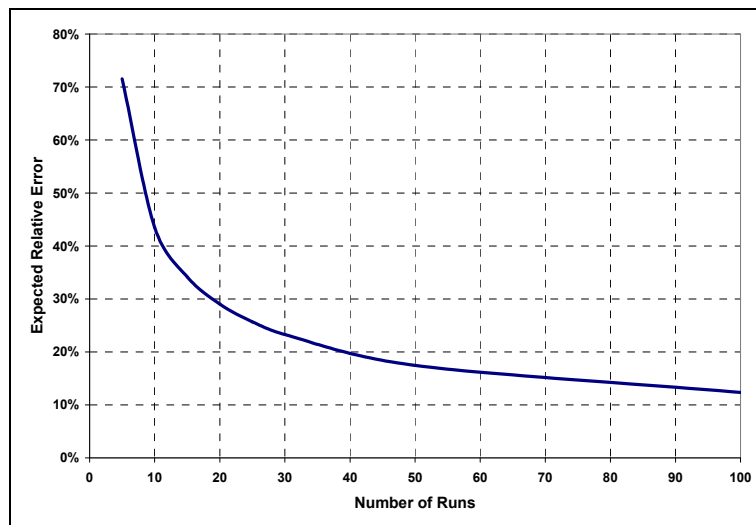
6. Plot the expected relative error ( $\epsilon_R$ ) as a function of the number of runs ( $R$ ).
7. Take decision on how many runs to execute based on the relative error that can be tolerated.

The number of runs may vary depending on what is the MOP of interest. The variability in the measures (latency, currency, etc.) may be affected more or less by the probabilistic aspects of the LBTB. Consequently, the methodology presented above was applied to each MOP separately.

### 3.4 Sample Size – Location Fidelity

The graph representing the expected relative error for the location fidelity measure, with  $\alpha$  equal to 0.10, is shown in Figure 7. The line corresponds to the maximum relative error expected obtained at every time the measures were recorded in the simulation. For instance, in the scenario used for the analysis, measurements were recorded every minute for an hour (at times  $t_1, t_2, \dots, t_{60}$ ), therefore, the values on the graph represent the maximum error of the location fidelity values at time  $t_1, t_2, \dots, t_{60}$ .

It is shown in the figure that for location fidelity, the user should execute as many runs as possible. The results indicate that even with the execution of 100 runs, the relative error to expect is still 12.3 percent of the estimated average. This means that there exists a very high variability in the location fidelity measurement. When the results are more heterogeneous as is the case for location fidelity, a larger sample size is required to obtain an acceptable level of precision.



**Figure 7:** Location Fidelity – Expected Relative Error

### 3.5 Sample Size – Currency

The expected relative error for currency as a function of the number of runs is presented in Figure 8. Note that the value  $\alpha$  used was 0.10. Obviously, the error diminishes as the number of runs increases. The relative error decreases rapidly until  $R$  equals approximately 15, at which point the relative error begins to hit a plateau. The figure should be read as follows. If for instance, 30 runs are executed, the user can expect, 90 % (corresponding to  $(1-\alpha)\times 100$ ) of the time, an error of approximately  $\pm 2.6$  % of the estimated average. This is true only if the variability in the MOPs is less than or equal to the variability obtained with the scenario used for this statistical analysis. The authors considered this a valid assumption as the parameters for the baseline scenario were chosen by the project sponsors such that the variability would be maximal.

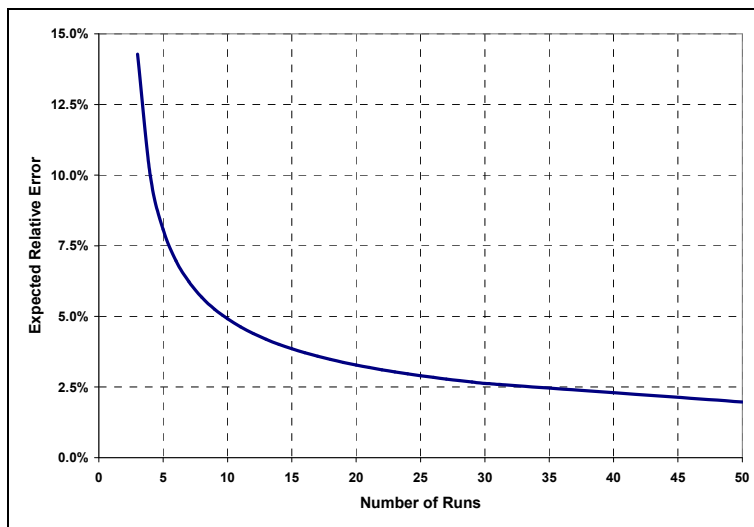
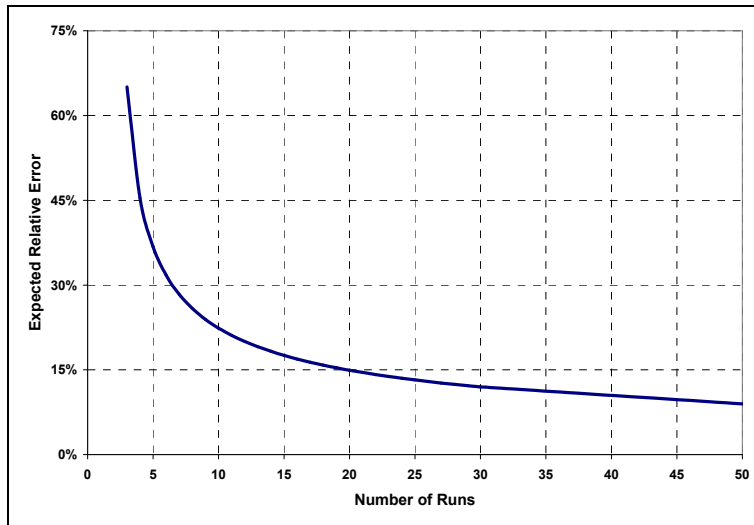


Figure 8: Currency – Expected Relative Error

### 3.6 Sample Size – Latency

Figure 9 presents the expected relative error as a function of the number of runs for the latency measure. The value  $\alpha$  was again set to 0.10. Once again, as shown with the currency, a slight increase in the number of runs makes a big difference when  $R$  is small but makes much less of a difference when  $R$  gets larger. In this case, the number of runs to be executed should be as large as possible since the error is quite large. Even when the curve eases into a plateau at  $R=20$ , the expected error is approximately  $\pm 15$  percent of the estimated average.





**Figure 9:** Latency – Expected Relative Error

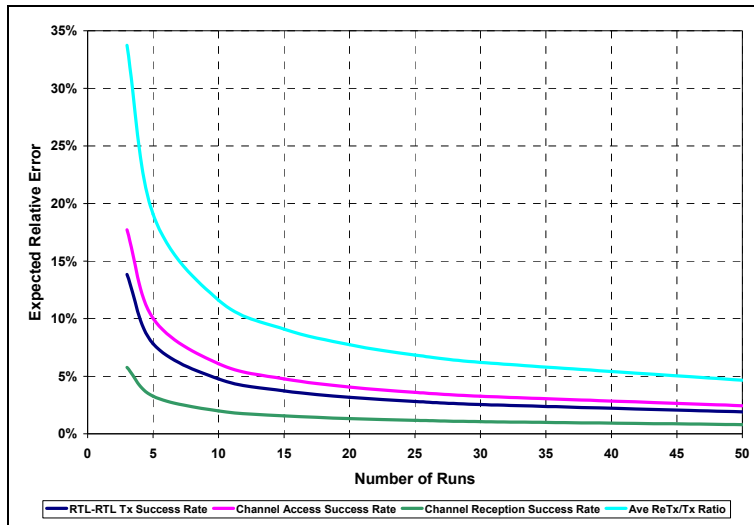
### 3.7 Sample Size – Network Measures

The members of the TDP also requested that this study look at the number of runs to be executed for four other measures not developed by the OR team; namely RTL-RTL Tx Success Rate, Channel Access Success Rate, Channel Reception Success Rate and Average ReTx/Tx Ratio. Figure 10 depicts the expected relative error for the four measures obtained with the application of the aforementioned methodology with  $\alpha$  equal to 0.10.

The figure was included in the document only as a reference for the members of the TDP. For the four network measures, it appears that a number of runs greater or equal to 15 would generate precise enough results, with expected errors varying between 1.5% for Channel Reception Success Rate and 9.1% for Average ReTx/Tx Ratio. It can be noted from this graph that there is no point in executing more than 25 runs since the value added (the precision) is negligible after that mark.

### 3.8 Comments

The method explained in this section allows the user to get an idea of the kind of error that can be expected as a function of the number of runs executed with the LBTB. It is based on the assumption that the scenario used for the analysis produced the largest variability in results that can be envisioned with the test bed (i.e. the worst case scenario with respect to the number of runs required to ensure statistical validity of the results). Assuming that this is a true statement, when executing the same number of runs for a different scenario, one can then expect an error less than or equal to those shown on the various graphs.



**Figure 10: Network Measures – Expected Relative Error**

In a perfect world, the proposed method would be automated in the LBTB for the analysis of the results. For automation of simulation data analysis, suitable stopping rules are necessary so the method presented needs to be done for every single scenario. The number of runs to be executed cannot be predicted in advance, it is rather an iterative process. In reality, for a given scenario, in order to provide statistically valid results, a certain number of runs  $R_0$  should be executed with the LBTB, then estimate  $R$  by performing Steps 3, 4 and 5, and finally execute  $R - R_0$  additional runs of the scenario.

However, such automation was not possible in this project. This capability was not part of the simulation tool (LBTB) when it was first installed at DRDC Valcartier, so additional funding would be required to incorporate it. There was also a time-related constraint in that the contractor had to address fundamental aspects of the simulation before the project ended and that resulted in not time being left for this topic.

This study's recommendation is that the proposed approach be applied manually. This means that the user should run a scenario a certain number of times, then measure the error with confidence intervals. If the error is not small enough, then the simulation should be restarted and run until the desired level of error has been reached.

## 4. Examples of the Use of the Measures

---

The members of the HCTCN TDP have used the tool to compare and analyze various scenarios that were executed with the LBTB. Although the focus of this report is not on the results, the authors decided to present some scenario results to show the reader potential uses of the measures introduced in this paper.

The aim of this section is to present two examples of analysis that can be done with the tool, the first one with the latency measurement and the second one, with the location fidelity measurement. The goal is simply to illustrate how the tool can be used.

### 4.1 Transmission Delays

The first example uses the latency measurements to compare the different delays of transmission between three different scenarios.

The three scenarios are very similar. They all correspond to the same one-hour patrol scenario and include 20 nodes (command and control, Coyote, etc.). In the scenarios, all the nodes are asked to report their position to all other nodes every minute. This results in a very congested communication system due to the high quantity of information broadcast over the network. The differences between the three scenarios are the following:

- Scenario #1: Native format and IM rules – The data format used for the transmission is called native format, which was developed by the contractor who built LBTB [3]. Also included in the scenario is an IM rule consisting of *not sending through the network all position reports if the position of the vehicle has not changed by more than 600 meters since the last position update broadcast*. The application of this rule is expected to reduce the total amount of traffic on the communication system.
- Scenario #2: ODB format and no IM rule – The format employed for data on transmitting over the network is Operation Database Protocol Data Unit (ODB PDU). This is the standard format used by the CF and is bigger in size than the native format. Additionally, no IM rules were used in this scenario.
- Scenario #3: Native format and no IM rule – This scenario is the same as the previous one except that format used is the native format.

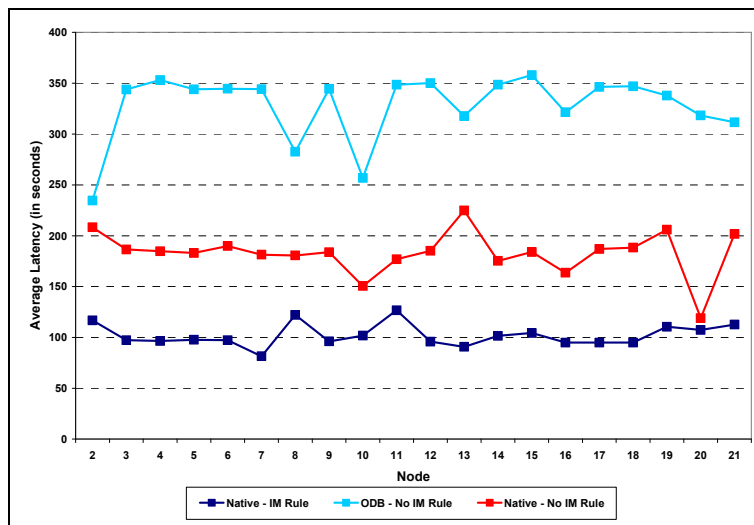
The "OpsMeasures.xls" tool was used to analyze the latency measurements. Table 2 presents a comparison of the average latency values, expressed in seconds, for the three scenarios. It shows that the latency for the Scenario #1 was the best amongst the three scenarios (i.e. the lowest values), almost twice as good as in Scenario #3 and more than three times better than Scenario #2.

**Table 2:** Average Latency per Scenario

Scenario	Description	Average Latency (in seconds)
#1	Native format and IM rules	101.0
#2	ODB format and no IM rules	331.3
#3	Native format and no IM rules	183.2

The noticeable difference between Scenario #1 and Scenario #3 can likely be explained by the fact that in Scenario #1, some undesired transmissions were removed with the application of the IM rule, which resulted in a diminution of the network traffic and therefore, reduced the latency values. As for the difference between Scenario #2 and Scenario #3, which were almost identical except for the data format (ODB PDU versus native), the primary reason why the latency is lesser for Scenario #3 is likely because the size of replication message is smaller in Scenario #3 (native format, 127 bytes per position update) than in Scenario #2 (ODB PDU format, 1127 bytes per position update). The smaller message size takes less time to transmit, resulting in improved channel availability and reduced latency.

More detailed analysis, for instance looking at individual node behavior, could have been done by examining Figure 11, which represents the average latency (in seconds), per node, from a receiving node's perspective.



**Figure 11:** Average Latency per Node (Three Scenarios)

## 4.2 Location Fidelity

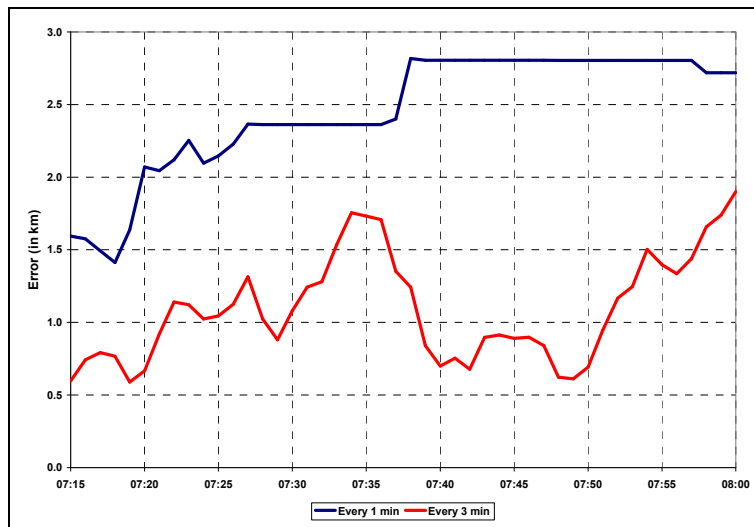
The basic scenario presented in the previous subsection looked at a one-hour patrol scenario with 20 nodes, where each node reports its position every minute. This creates a significant level of traffic on the network. But, what if, instead of every minute, the vehicles would report their position every three minutes? Obviously, the

traffic on the communication system would decrease, but what about the overall situational awareness? The location fidelity measurements presented in this report could partly answer this question.

To demonstrate how the tool can be used to answer such a question, a fourth scenario was defined as follows:

- Scenario #4: The same configuration as Scenario #3, i.e. native data format and IM rules off, except that the vehicles report their position every three minutes.

Figure 12 presents the operational measures tool location fidelity measurements, taken between 07:15 AM until 08:00 AM, obtained for Scenario #3 (every one minute) and Scenario #4 (every three minutes).



**Figure 12: Location Fidelity (Scenarios #3 and #4)**

The results are almost counter-intuitive. One might expect that, if the vehicles report their position at a higher frequency, then the location fidelity measurements (the errors between perceived and real positions) would decrease. However, this example suggests the opposite, i.e. increasing the time between the position updates from one to three minutes appears to improve the location fidelity. In the time period of interest, the average errors for the entire period were respectively 2.5 km for Scenario #3 and 1.1 km for Scenario #4 and, the errors were always better for Scenario #4 (red line under blue line in Figure 12).

These results can likely be explained by the fact that in Scenario #3, the traffic level is so high on the network that probability of channel access is reduced and the transmissions between vehicles take longer. On the other hand, in the case of Scenario #4, the network is not as congested, so the information circulates more efficiently between the vehicles. This can be confirmed by looking at the latency measurements for the "position updates" type events. The average latency measurements were 183

seconds for Scenario #3 versus 72 seconds for Scenario #4. This means that, in the case of Scenario #4, on average, the transmission time of a position update was less than the time in between two consecutive updates. This opposite was true for Scenario #3 where new position updates were transmitted before the latest vehicle positions were received (i.e. average delay of transmission exceeded the time between two consecutive position updates).

This section introduced simple potential uses of the measures and the tool presented in this report. A complete set of experimental results from the information management segment of the HCTCN TD project will be presented in the final project report.

## 5. Conclusion

---

This document summarizes the OR support provided by the DRDC Valcartier OR team to the HCTCN TDP, including the following:

- An analysis regarding the number of times a scenario should be executed with the LBTB to ensure statistically valid results.
- An analysis of the existing Measures of Performance (MOPs) implemented in the LBTB, and the proposal for refinements to the existing MOPs to better suit the detailed analysis requirements of the HCTCN TDP. These refinements included several possible interpretations of the Location Fidelity measures as well as the consideration of more aspects such as transmission priority and replication message type.
- Development of a user-friendly software tool to implement data analysis protocols based on the refined MOPs.

As with all software tools, some improvements are possible and may be recommended in case LBTB will continue to be used for future projects. The study suggests two main areas for improvement:

- Combine the two steps presently needed for the calculation of MOPs into one. This will save some time through automation of some repetitive actions and will also eliminate a potential source of errors by eliminating the human intervention that is presently required.
- Improve the way the analysis determines the number of times to execute the scenario to obtain statistically valid results. The number of runs depends on several factors and it varies for different MOPs, for different scenarios, etc. As such, it cannot be predicted once and for all. It has to be calculated through an iterative process for each case. Implementing this procedure would ensure that statistically valid results are done within LBTB (and not outside it, as it is done presently).

## References

---

1. High Capacity Tactical Communications Networks Technology Demonstration Project, Synopsis Sheet Effective Technology Demonstration Project Approval, February 2001.
2. Low Bandwidth Test Bed Development, Annex C1, Functional Specification, Contract W7701-11632, February 2002.
3. Beauchemin, P., Johnson, T., Taylor, C. and Zheng, L. (2005). HCTCN Low Bandwidth Test Bed: User Reference Manual (IPU Document Number: 021/2003), Version 3.0, Revision 1.1, IP Unwired Inc.
4. Movable Type Ltd., Calculate distance and bearing between two Latitude/Longitude points (Online), <http://www.movable-type.co.uk/scripts/LatLong.html> (4 April 2006)
5. Law, M.L. and Kelton, W.D., Simulation Modeling and Analysis, McGraw-Hill Companies Inc., 2000
6. Israel, G.D., Determining Sample Size, <http://edis.ifas.ufl.edu/PD006> (25 July 2006)
7. Banks, J. and Carson II, J.S., Discrete-Event System Simulation, Prentice-Hall, Englewood Cliffs, New Jersey, 1984



# Annex A: User Guide

## LBTB Folders and Files Structure

As a scenario script is processed in LBTB, the application records a large quantity of information about the events, the changes to the database, the network performance and a set of pre-defined measurements. This information is kept in the log database on the Central Server (CS) for future analysis.

A tool such as the analysis workstation is one way the user can access this information to produce reports and/or consolidate the information for viewing purposes.

During the execution of a scenario with the LBTB, various log information and data to be measured (type and duration of the measurements are *a priori* selected by the user) are written in the database tables. The structure of the folders containing the configuration files and the files generated during a run is presented Figure A.1.

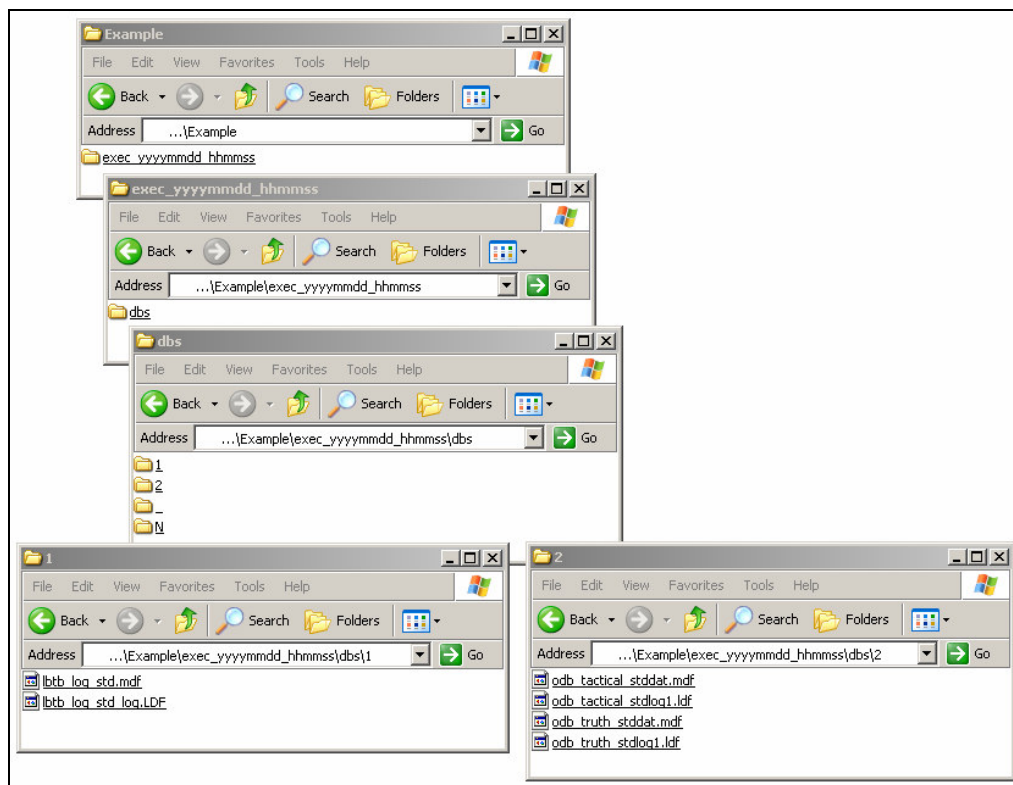


Figure A.1: Folders and Files Structure of a Run

In this example, the folder "exec\_yyyymmdd\_hhmmss" is the main folder containing all the information of the particular run. The folder "dbs" contains a folder for each of the  $N$  nodes in the scenario. The folder named "1" always corresponds to the CS.

Folders "2" to "N" contain the information recorded for each of the tactical nodes in the scenario.

The tables in the ".mdf" files contain a large quantity of information related to the scenario previously executed. However, most of the data captured during the execution of a run is not required for the measures of performance presented in Section 2. For that reason, the authors created a SQL script to retrieve only the necessary information from the database and export it in a series of Excel files.

## SQL Script

The SQL script that was developed to extract the information from the database is named "GenerateExcelFiles.sql". A prerequisite to run the script is to have Microsoft SQL Server installed on the computer.

The first step before running the "GenerateExcelFiles.sql" script is to execute another SQL script, called "CreateTimeDiff.sql", which creates a function used in the "GenerateExcelFiles.sql". In order to run the script that creates this function, the user simply has to open the file using the SQL Query Analyzer tool, which is available within Microsoft SQL Server and then hit the run button (or hit F5). Note that this step only has to be done once.

When the function has been created, the "GenerateExcelFiles.sql" script can be run using the same approach, i.e. by opening the file using the SQL Query Analyzer tool. Once the file is opened, the user must specify the inputs of the script. Figure A.2 corresponds to the code of the SQL script containing the input parameters (the red font characters in quotes).

```
-- INPUTS FROM THE USER
-- LOCATION OF THE SOURCE RUNS AND DESTINATION FOLDERS
SET @uname = 'username'
SET @pwd = 'password'
SET @folderSource = 'E:\Projets\HCTNC\Scenario1\'
SET @folderDest = 'E:\Scenario1\'
INSERT INTO myFileNameTable (nom) VALUES ('exec1')
INSERT INTO myFileNameTable (nom) VALUES ('exec2')
-- INSERT MORE RUNNAME IN myFileNameTable if desired ...
```

**Figure A.2:** Input Parameters in the "GenerateExcelFiles.sql" Script

The following five parameters must be specified:

- Username: The username of the SQL server;
- Password: The password of the SQL server;

- Source Folder: The name of the parent folder where the runs to be analyzed are located;
- Destination Folder: The folder where the extracted information (in Excel files format) will be located; and
- Run Names: The names of the runs from which the information will be retrieved. These runs should be located in the source folder.

Note that for the specification of the source and destination folders, it is mandatory to include the "\" sign (backslash) for the script to run properly. In this example, the information would be extracted from the runs "exec1" and "exec2", both located in the folder "E:\Projets\HCTNC\Scenario1\". To add further runs, more "INSERT INTO" statements shall be added with the appropriate values corresponding to the names.

A point that is worth mentioning is that each individual run should follow the folder and file structure presented in Figure A.1. However, the only tables that are attached, queried, and detached, are all part of the "lbtb\_log\_std.mdf" file located in "...\dbs\1\" folder. In order to limit the disk space due to storage of run results, the user can keep the aforementioned file ("lbtb\_log\_std.mdf") and delete all other files and subfolders ("...\dbs\2", "...\dbs\3", etc.).

Finally, to run the script, the user shall hit the run button. The script automatically initializes the variables, reads the input parameters, attaches the databases, performs the queries to extract the information (for all of the measures), and detaches the databases.

In the example illustrated in Figure A.2, if the script was executed, two folders would be created, one for each run, "Ops\_exec1" and "Ops\_exec2". Both of them would be located in the destination folder specified by the user, which in this case is "E:\Scenario1\". After the execution of the script, a number of Excel files would be generated and added to the new folders, i.e. "E:\Scenario1\Ops\_exec1" and "E:\Scenario1\Ops\_exec2". All the required information for the measures is summarized in the following four Excel files:

- LF.xls;
- CURRENCY\_LOG.xls;
- LATENCY\_LOG.xls; and
- NODE.xls.

Note that, in Annex B, the reader can find the code and the details of the "GenerateExcelFiles.sql" SQL script that was developed.

## Description of Excel Files Generated from the SQL Script

Below is a description of each of the four files. Note that the columns in the Excel file are relatively self-explanatory. If more details are required, consult the reference manual [3]

### Excel File: "LF.xls"

An example of a "LF.xls" file is presented in Figure A.3. The columns in the file, in the same order they appear in the figure, are the following: the logging time (the time when the measurement was taken), the observing node, the observed node, latitude1 and longitude1 corresponding to the position where the observing node believed the observed node was at the time the measure was taken and finally, latitude2 and longitude2, which represent the true position of the observed node at the time the measure was recorded.

	A	B	C	D	E	F	G
1	29/03/2004 12:03:00 AM	FTN2	FTN5	46.86368	-71.164552	46.86368	-71.164552
2	29/03/2004 12:03:00 AM	FTN2	FTN3	46.756547	-71.400408	46.756547	-71.400408
3	29/03/2004 12:03:00 AM	FTN3	FTN5	46.86368	-71.164552	46.86368	-71.164552
4	29/03/2004 12:03:00 AM	FTN3	FTN2	46.855968	-71.417143	46.855968	-71.417143
5	29/03/2004 12:03:00 AM	FTN4	FTN5	46.86368	-71.164552	46.86368	-71.164552
6	...	...	...	...	...	...	...

*Figure A.3: Example of "LF.xls" file*

In the example of Figure A.3, at time 12:03:00 AM, Node 2 (FTN2) knows the exact location of Node 5 (FTN5), since longitude1 equals longitude2 and latitude1 equals latitude2. When measuring the location fidelity, the file is parsed and only the rows that satisfy the inputs parameters (start time, end time, node of interest, etc.) specified by the user are consider. This will become clearer once the inputs will be described, later in this section.

### Excel File: "CURRENCY\_LOG.xls"

An example of a "CURRENCY\_LOG.xls" file is shown in Figure A.4. The file is composed of nine columns: the logging node ID (identifier of the node generating the log entry), the update time (scenario time of the log entry), the local node, the source node, the event identifier, the event type (Position Update, Vector Movement, Contact Report, Movement Order, etc.), the priority<sup>3</sup> associated with the event, the replication identifier and the time since last update (the elapsed scenario time since the last update of that data).

---

<sup>3</sup> Note that in the current version, the priority column is not properly populated by the SQL script. At this time, the default value for all rows in the file is 3. The column has been included in the file for future development (if ever required).

	A	B	C	D	E	F	G	H	I
1	2	29/03/2004 12:02:00 AM	2	2	1	Vector Movement	3	2.0	Day 000 00:00:00.000000
2	2	29/03/2004 12:02:43 AM	2	3	2	Vector Movement	3	3.0	Day 000 00:00:00.000000
3	2	29/03/2004 12:02:57 AM	2	5	3	Vector Movement	3	5.0	Day 000 00:00:00.000000
4	2	29/03/2004 12:03:26 AM	2	4	4	Vector Movement	3	4.0	Day 000 00:00:00.000000
5	2	29/03/2004 12:04:00 AM	2	2	5	Vector Movement	3	2.1	Day 000 00:02:00.000000
6	...	...	...	...	...	...	...	...	...

**Figure A.4:** Example of "CURRENCY\_LOG.xls" File

As indicated by its name, the "CURRENCY\_LOG.xls" file is used for the currency measurement. The currency values that are calculated are based on the last column (Column I), which corresponds to the scenario time since the last update of that data. For the currency measurements, only the rows that meet the different parameters listed by the user (start time, end time, event type, etc.) are considered in the computation of measurements.

### **Excel File: "LATENCY LOG.xls"**

The structure of the "LATENCY\_LOG.xls" Excel file is very similar to "CURRENCY\_LOG.xls". Figure A.5 presents an example of "LATENCY\_LOG.xls". The file contains the following nine columns: the logging node ID (identifier of the node generating the log entry), the source node identifier, the destination node identifier, the event type, the priority associated with the event, the event ID, the replication identifier, the latency value (transaction time) and the update time (scenario time of the log entry).

	A	B	C	D	E	F	G	H	I
1	2	3	2	Vector Movement	2	2	3.0	Day 000 00:00:12.9989	29/03/2004 12:02:42 AM
2	2	3	2	Vector Movement	2	7	3.1	Day 000 00:00:45.1346	29/03/2004 12:05:15 AM
3	2	3	2	Vector Movement	2	14	3.2	Day 000 00:00:14.1772	29/03/2004 12:06:44 AM
4	2	3	2	Vector Movement	2	18	3.3	Day 000 00:00:21.8976	29/03/2004 12:08:51 AM
5	2	3	2	Vector Movement	2	22	3.4	Day 000 00:01:32.8638	29/03/2004 12:12:02 AM
6	2	3	2	Vector Movement	2	26	3.5	Day 000 00:01:14.7514	29/03/2004 12:13:44 AM
7	...	...	...	...	...	...	...	...	...

**Figure A.5:** Example of "LATENCY\_LOG.xls" file

Obviously, the "LATENCY\_LOG.xls" file is the one used for the latency measurements. The latency values that are used for the calculations are in Column H, and only the rows satisfying the input parameters are considered, i.e. start time, end time, event type, priority type, etc.

### **Excel File: "NODE.xls"**

This Excel file is used for all three measures to determine the number and the name of the nodes in the scenario to be analyzed. The file is generated from the NODE\_LIST table, found in "...\dbs\1\lbtb\_log\_std.mdf" file. There are six columns in the file, one for each of the following: the node identifier, the record identifier in the table, the role of the node, the node that logged the information (it normally corresponds to the CS), the type of the node, and the name of the node. Only the node identifier and the name of the node fields are used for the calculations and populated results.

## VBA Excel Tool – "OpsMeseasures.xls"

The four Excel files described above are transparent, meaning that the user does not have to view or edit them at any time. The authors developed a Visual Basic for Applications (VBA) Microsoft Excel spreadsheet tool, named "OpsMeasures.xls", that automatically calculates the MOPs, based on the Excel files generated by the SQL script and on a set of input parameters specified by the user. The GUI of the tool is presented in Figures C.1, C.2 and C.3 of Annex C.

Each measure can be calculated separately by clicking on the corresponding measure button ("Location Fidelity", "Currency" or "Latency") located in Column A of the GUI (see Figure C.1). When clicking on one of the buttons, the inputs are read, the files required to calculate the measure are opened<sup>4</sup>, some preprocessing and formatting actions are performed, and the desired measures are returned. The user also has the possibility of running the three measures at once by clicking on the "Execute All" button also positioned in Column A.

For the three measures, there are more than 50 input parameters in "OpsMeasures.xls" that must be specified by the user. Note that there is very little error checking done by the model, so it is important that the user understands the various inputs for the model to run properly and to return valid outputs.

The remainder of this section serves as a user's guide. The different inputs necessary for all three measurements as well as the general inputs (common for all MOPs) are described. The format of the outputs returned for location fidelity, currency and latency are also presented.

### General Inputs

In the GUI, there are some parameters that are common for all three MOPs. Here is a description of each of them:

- Parent Folder (Cell A1) – The folder that contains the different runs to be analyzed. Note that the "\" (backslash) sign is required at the end.
- Name of Runs to Analyze (Cells A2:A19) – The name of the runs to be analyzed. They correspond to folders containing the four Excel files previously depicted. Note that the application stops reading at the first blank cell.
- Results File Name (Cell D3) – The name of the Excel file in which the outputs will be recorded. The default name (template) is "Results.xls". The template has a predefined format, with five worksheets ("LocFid", "Latency",

---

<sup>4</sup> For instance, when clicking on the "Location Fidelity" button, the "LF.xls" file will open as well as "Node.xls". Note that the file containing the nodes information ("Node.xls") opens for all three measurements.

"Currency", "Temp" and "Inputs"), that are used by the model. Note that this file must be located in the same folder as the application ("OpsMeasures.xls").

- Results File Name Location (Cell D4) – The user can choose to save the results (after the measures are computed) in a file which can be found at the location specified in this cell. Note that if the user chooses not to save the results, the results file will remain open.
- Close Support Files (dropdown list near Cell C1) – A "Yes/No" parameter indicating if the supporting Excel files ("LF.xls", "CURRENCY\_LOG.xls", etc.) will be closed after the execution. It is recommended to set it to "No" when working in debugging mode, otherwise, it should be set to "Yes" for normal use of the tool. Note that the files are not saved before they are closed.
- Close Results Files (dropdown list near Cell C2) – A "Yes/No" parameter indicating if the results file will be saved after the execution. If the parameter is set to "No", the file will remain open and it will be to the user's discretion to save the file or to discard the run. When the parameter is set to "Yes":
  - For location fidelity analysis: A file for each run will be saved with the name "OpsResults\_dd\_mm\_yyyy\_hh\_mm\_ss\_AM/PM.xls" in the same folder results File Name Location (Cell D4). Note that the date and time used in the name are the ones when the results analysis was performed.
  - For currency and latency analysis: Only one file will be saved at the end of the last run, even if there were multiple runs. The file will include the measurements of all the runs under the "Currency" or "Latency" worksheet, depending on the type of analysis. The name of the file will be "OpsResults\_dd\_mm\_yyyy\_hh\_mm\_ss\_AM/PM.xls".

Note that after the computation of a measure, the inputs are automatically copied in the worksheet named "inputs" of the results file workbook, and also a copy of "OpsMeasures.xls" is made with the name "OpsParam\_dd\_mm\_yyyy\_hh\_mm\_ss\_AM/PM.xls". This is useful when the user wants to re-calculate a measure or to review the list of inputs that were used (priority, event type, etc.)

## Location Fidelity

When the user clicks on the "Location Fidelity" button, the "LF.xls" file opens and there is a preprocessing step automatically done by the application. This step consists of adding a column in "LF.xls" corresponding to the actual error (distance between true location and the perceived location). This is necessary because the file only has raw data (latitude1, longitude1, latitude2 and longitude2) and not the actual error or distance. Note that the distance is calculated using the Haversine equation described in depth in [4]. This formula is particularly well-conditioned for numerical

computation even at small distances, unlike calculations based on the spherical law of cosines.

## **Inputs**

Consider the following notation:  $D(N_1, N_2, t_k)$  is the distance, calculated using the Haversine formula, between where  $N_1$  believes  $N_2$  is at time  $t_k$  and where  $N_2$  really is (its true location).

The inputs specific to the location fidelity measure are the following:

- Type of Analysis (dropdown list near Cell C20) – The user has to select one of the three possible types:
  - Compute Averages and All Nodes Information: This type of analysis will provide a complete analysis, i.e. the output will include all possible location fidelity measurements. For instance, if a simulation included four nodes, for each time  $t_k$  (time when the measures were taken), LBTB records  $D(1, 2, t_k)$ ,  $D(1, 3, t_k)$ ,  $D(1, 4, t_k)$ ,  $D(2, 1, t_k)$ ,  $D(2, 3, t_k)$ ,  $D(2, 4, t_k)$ , ...,  $D(4, 3, t_k)$ . Also, for each  $t_k$  and for each node, the application will calculate the maximum value of the location fidelity measurements, the number of nodes seen by the current node at time  $t_k$ , the average, the standard deviation and the confidence interval (upper bound and lower bound).
  - Compute Only Averages: If this type is selected, for each node at time  $t_k$ , the application will only return the maximum values, the averages, the standard deviations and the confidence intervals of the location fidelity measurements for the times the measures were taken. The difference between this mode and the previous one is that no individual node information will be populated. In the example presented for the previous type,  $D(1, 2, t_k)$ ,  $D(1, 3, t_k)$ ,  $D(1, 4, t_k)$ ,  $D(2, 1, t_k)$ ,  $D(2, 3, t_k)$ ,  $D(2, 4, t_k)$ , ...,  $D(4, 3, t_k)$ , would not be displayed in the results.
  - Compute Averages and the Information on the Nodes Specified in Chart: In this case, the application will populate all the maximum values, the averages, the standard deviations and the confidence intervals of the location fidelity measurements. In addition, the application will calculate the specific measures defined in Graph for Location Fidelity parameter (see below). This mode can be used if the user only wants to see the location fidelity measures for particular nodes and not for other ones (for instance, information how the command and control node perceived the other nodes in the scenario).
- Confidence Level (Cell D21) – The location fidelity measurements are given in terms of a confidence interval (average, upper bound and lower bound), therefore, a confidence level must be specified by the user. Note that the



confidence level is a value between "0.00" and "1.00", with a default value of "0.95".

- Use Bar Chart (Cell D22) – The user has the possibility to include some bar chart animation to see the evolution of the location fidelity through time, from a single node's point of view. If this functionality is desired, the value must be set to "Y", otherwise to "N".
- Use Time Interval (Cell D23) – If this parameter is set to "Y", the location fidelity measure will be done for the rows logged between start time (Cell D24) and end time (Cell D25), otherwise, all of the records will be considered.
- Start Time (Cell D24) – The start time of the interval. All rows recorded before start time will be automatically ignored. If the above parameter is set to "N", this input is ignored.
- End Time (Cell D25) – The end time of the interval. Note that all rows logged after end time will be discarded. If the "Use Time Interval" parameter is set to "N", this input is ignored.
- Wait Time (Cell D26) – The wait time is an integer between 1 and 59 seconds corresponding to the time between two consecutive bar chart updates (refresh time).
- Y-axis Maximum (Cell D27) – If the value is set to -1, the Y-axis maximum will be automatically determined by the application. Otherwise (different than -1), the value entered by the user will be used as the maximum.
- Step By (Cell D28) – Consider the following example. If in the LBTB scenario, the measurements were taken every 5 minutes starting at midnight, there would be measures logged in "LF.xls" at "12:05:00 AM", "12:10:00 AM", "12:15:00 AM", etc. The "Step By" parameter indicates the bar chart values that will be plotted. In other words, if "Step By" was set to 2 in the aforementioned example, the bar chart would correspond to every other time step, in this case, "12:05:00 AM", "12:15:00 AM", etc. Note that the time one particular chart will be visible on the screen is given in the "Wait Time" parameter (presented earlier).
- Observing Node (Cell D29) – This parameter is also linked to the bar chart. This is the node for which the bar chart will be constructed. Most of the time, the observing node will be the one corresponding to the command and control centre.
- Graph for Location Fidelity (Columns F:H) – The user can enter which graphs (of type XY-line chart) will be automatically created by the application. In the example of Figure C.2, the user specified two charts to be plotted (1 and 2 in Column H). In this case the first chart would present the error in kilometers for how Node 2 sees the other nodes (3, 4 and 5) versus their true location.

Chart 2 would show the evolution of the average error through time, along with the upper and lower bound estimates of the confidence interval.

## **Outputs**

The location fidelity measurements are given as a confidence interval, i.e. as an estimated range of values which is likely to include the error parameter (in this case, distance between believed and true location), the estimated range being calculated from a given set of sample data. Confidence intervals are usually calculated so that this percentage is 95%, but we can produce 90%, 99%, 99.9%, confidence intervals for the unknown parameter. More details about confidence interval and confidence level are given in [5].

Due to the large quantity of information provided by the location fidelity measure, there are many graphs that can be drawn from the outputs. The user has the possibility to automatically have charts drawn by the application. The generated charts show the evolution of the location fidelity measurement as a function of the time in the scenario. This capability can be used with the input parameter "Graph for Location Fidelity" presented in the previous subsection. However, it is recommended that the user execute the location fidelity measure and then manually create the chart that is relevant for the analysis.

The outputs are represented in a table. Consider the case where the type of analysis is "Averages and All Nodes". It is better to present this case as an example because it includes all of the other types. As mentioned earlier, a complete analysis of location fidelity is performed when this type is selected. An example of a full location fidelity measurement output is presented in Table A.1. The results were truncated due to the size of the table. In this case, the scenario included 20 nodes (Nodes 2 to 21), the time interval used was "07:15:00 AM" (start time) to "07:45:00 AM" (end time) and the measurements in the LBTB, were taken every two minutes.

**Table A.1:** Example of Location Fidelity Output

		2000-04-12 07:15:00 AM	2000-04-12 07:17:00 AM	...	2000-04-12 07:45:00 AM
2	N	19	19	...	19
	Maximum	5.000047	5.500017	...	7.833319
	Average	3.807023	4.192988	...	5.789482
	Std. Dev.	0.724918	0.724915	...	1.382039
	LB	3.457623	3.843590	...	5.123361
	UB	4.156422	4.542386	...	6.455603
	2				
3	0.000000	0.000000	...	0.000000	
4	3.333301	3.666722	...	7.333365	
...	...	...	...	...	
21	5.000010	5.500017	...	5.666638	
...	...	...	...	...	
...	...	...	...	...	
21	N	19	19	...	19
	Maximum	5.000047	5.500017	...	10.666729
	Average	3.780710	4.131578	...	5.982467
	Std. Dev.	0.700647	0.688483	...	1.783771
	LB	3.443009	3.799740	...	5.122717
	UB	4.118411	4.463416	...	6.842217
	2	3.666725	4.000008	...	10.666729
	3	0.000000	0.000000	...	0.000000
	...	...	...	...	...
	20	4.666712	4.999941	...	6.166711
21					
All	N	380	380	...	380
	Maximum	5.333300	5.666737	...	10.833371
	Average	3.800841	4.185837	...	6.025426
	Std. Dev.	0.682603	0.694326	...	1.699724
	LB	3.733947	4.117794	...	5.858856
	UB	3.867735	4.253880	...	6.191996

There is one major block for each node in the simulation (see first column containing 2, 3, ..., 21) and one block at the bottom (labeled "All"), which presents the global measures. The major blocks have the following information:

- **N**: At the time the location fidelity measurements are taken, a particular node may not have any information in its tactical database about the position of one or many other nodes. Consider the following definition. At time  $t_k$ , a Node B is defined as "known" by Node A, if Node A has information about the position of Node B, although it may not be the true position. The value N represents the number of nodes "known" by the node of interest. In the example presented in Table A.1, at time "07:15:00 AM", Node 2 has information about all of the other 19 nodes.

- Maximum: For a given node, this represents the maximum error (distance between the perceived and the true positions) amongst the N "known" nodes (nodes for which information are available).
- Average: The average error amongst the known nodes. In Table A.1, for Node 21 and at time "07:15:00 AM", the average is 3.78 kilometers, which is the mean of 3.67 (Node 2), 0.00 (Node 3), ..., 4.67 (Node 4).
- Std. Dev.: The standard deviation is calculated from the errors of the known nodes. The standard deviation value is used to populate the lower and upper bound of the confidence interval.
- LB and UB: These values corresponds to the lower and upper bounds of the confidence interval calculated based on the average, the standard deviation, the value N, and the confidence level specified by the user in the GUI.

The block at the bottom has the same output parameters but is obtained from a global perspective (all of the nodes) and not only from one node's point of view.

If the type of analysis selected is "Compute Only Averages", all of the rows containing individual nodes information will not be displayed in the output table. These rows correspond to rows in which the values in Column 2 are numbers. When "Compute Averages and the Information on the Nodes Specified in Chart" is the selected type, only individual node information as specified in "Graph for Location Fidelity" input parameter will be presented in the table.

## Currency

When the user clicks on the "Currency" button, the application opens the "NODE.xls" and "CURRENCY\_LOG.xls" file. The first step that is done by the macro is a reformatting of the values in Column I of "CURRENCY\_LOG.xls". The reformatting consists of converting the time since last update in a "Day 000 hh:mm:ss.sssss" format, in a float format, corresponding to the time since last update expressed in seconds. This is necessary because "Day 000 hh:mm:ss.sssss" is not a valid format in Microsoft Excel.

### Inputs

The following input parameters are specific to the currency MOP:

- Replication Type (checkboxes between Cells A31:A42) – Specification of which replication types will be considered in the analysis. There is a total of 13 types that can be selected, the most common being "Position Update", "Vector Movement" and "IIM Transfer Data". The user can select more than one replication type for a given analysis.

- Start Time (Cell D6) – The start logging time from which the rows will be considered in the analysis.
- End Time (Cell D7) – The end logging time from which the rows will be considered in the analysis<sup>5</sup>.
- Local Node (Cell D9) – The local node to consider in the analysis for currency measures A, C and E described in the "Outputs" subsection.
- Source Node (Cell D10) – The source node to consider in the calculation for currency measures A and B introduced in "Outputs" subsection.
- IIM Data Transfer Correspondence (Columns J:K) – This is a table of correspondence between the events of type "IIM Transfer Data" and the other event types. The user can specify, for instance, that a given event entered as "IIM Transfer Data" in the LBTB, is in fact, corresponding to a "Vector Movement" or a "Position Update". In the example presented in Table B.2, Events 6 and 11 should be considered in the calculations as events of type "Vector Movement". This input is necessary because of some limitations in the LBTB regarding the implementation of the intelligent information management process.

## **Outputs**

The outputs provided by the currency measurement are presented in one table. Consider the example depicted in Table A.2, where the "Local Node" input parameter was set to 2 and "Source Node" to 4. In this case, the currency results were calculated as follows:

- Measure A: The average of currency values (Column I) for all the rows with local node (Column C) equal 2 and source node (Column D) equals to 4.
- Measure B: The average of the currency values for all the rows with source node equals to 4 and local node different from 4.
- Measure C: The average of the currency values for all the rows with local node equal to 2 and source node different from 2.
- Measure D: The average of the currency values for all the rows where local node is not equal to the source node.
- Measure E: The average of the currency values for all the rows where local and source nodes equal to 2.

---

<sup>5</sup> Note that in order to analyze all the rows in the file, the user can set the start time to a very small value and end time to an extremely high value. For instance, if start time="12/04/1980 12:00:00 AM" and end time="12/04/2020 12:00:00 AM", all the rows in "CURRENCY\_LOG.xls" will be considered.

- Measure E: The average of the currency values for all the rows where local node equals source node.

For each of the currency measures (A, B, C, D, E and F), the tool returns the number of rows, the sum of the currency values and the average currency time of the rows that meet the criteria specified through the input parameters (replication type, start time, end time, etc.). Note that the currency values are expressed in terms of seconds.

**Table A.2:** Example of Currency Output

E:\Scenario8\exec1					
	Local	Source	Nbr of Rows	Sum	Average
Measure A	2	4	14	1944.211	138.872
Measure B	All	4	42	5832.633	138.872
Measure C	2	All	31	3813.439	123.014
Measure D	All	All	138	17585.694	127.433
Measure E	2	2	11	1200.000	109.091
Measure F	N	N	38	4080.000	107.368

## Latency

A similar reformatting preprocessing step as the one done for the currency measurement is executed when the user clicks on the "Latency" button. The values in Column H of the "LATENCY\_LOG.xls" file are converted to a float format (in seconds), which is readable by Microsoft Excel.

## Inputs

The following parameters with regards to latency must be specified by the user:

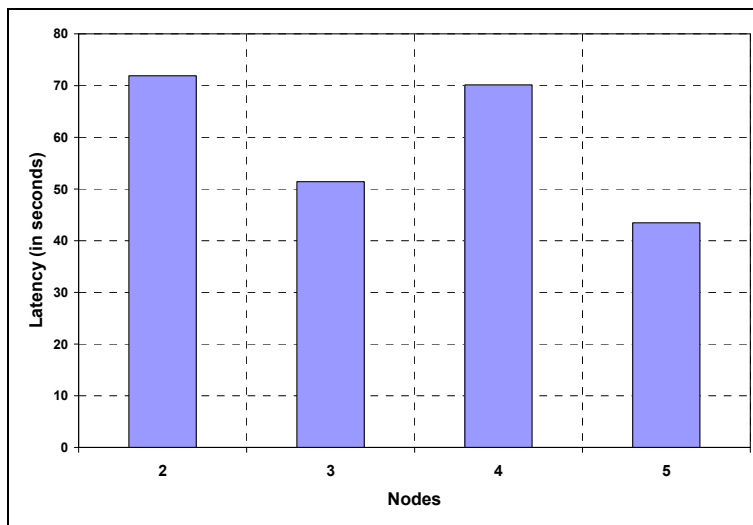
- Replication Type (checkboxes between Cells A31:A42) – Same as replication type described for currency.
- Priority Type (checkboxes between Cells C12:C16) – The replication priorities that will be included in this measurement. There are four possible values: 0, 1, 2 and 3. Note that the user can check more than one priority, for instance, one may check priorities 0 and 1 and ignore all other replications with priorities 2 and 3.
- Start Time (Cell D6) – Same as start time described for currency.
- End Time (Cell D7) – Same as end time described for currency.
- Type of Analysis (dropdown list near Cell D17) – This dropdown list has two possible values. The first one, "Sending Node", simply indicates that the latency measurement will be done from a sending node's point of view. This means that only the replications originating from the node will be considered in the calculation. The second possible value, "Receiving Node", is the

opposite of the first one, meaning that the latency calculation will be made from the perspective of the receiving node (replications received by the node). This will be clarified in the next subsection.

- IIM Data Transfer Correspondence (Columns J:K) – Same as "IIM Data Transfer Correspondence" input parameter described for currency.

## Outputs

The outputs that are populated from the latency measurements can be summarized in one bar chart such as the example presented in Figure A.6. After the execution of a latency analysis, latency values for each node are written in the worksheet labeled "Latency" in the results file, and the bar chart is automatically plotted.



**Figure A.6:** Average Latency (in Seconds) per Node

The chart should be interpreted as follows: for each node, the latency value corresponds to the average latency of the rows that meet the input parameters specified by the user (replication type, priority type, start time, end time, etc.), either from a sending node's point of view or from a receiving node's point of view. In the example presented in Figure A.6, if the analysis was performed from a sending node's perspective, for the latency corresponding to Node 2, only the records that have Node 2 as source node would have been considered. Otherwise, if from a receiving node's point of view, only the rows where destination node equalled 2 would have been included in the latency calculations of Node 2.

Note that the VBA Excel tool also calculates a more global value for the latency which corresponds to the average latency (a weighted average) for all the nodes in the scenario. This value is outputted in the worksheet labeled "Latency" in the results file.

## Annex B: SQL Script

```
-- VARIABLES DECLARATION
DECLARE @folderSource varchar(2048)
DECLARE @folderDest varchar(2048)
DECLARE @nFiles INT
DECLARE @Filename varchar(2048)
IF EXISTS (SELECT * FROM INFORMATION_SCHEMA.tables WHERE table_name = 'myFileNameTable') BEGIN
    DROP TABLE myFileNameTable
END
CREATE TABLE myFileNameTable (nom varchar(2048))
DECLARE @pwd varchar(50)
DECLARE @uname varchar(50)

-- INPUTS FROM THE USER
-- LOCATION OF THE SOURCE RUNS AND DESTINATION FOLDERS
SET @uname = 'uname'
SET @pwd = 'pwd'
SET @folderSource = 'e:\Jean-Denis\Projets\HCTNC\Scenario8\
SET @folderDest = 'e:\Scenario8\
INSERT INTO myFileNameTable (nom) VALUES ('exec1')

-- GET THE NUMBER OF FILES THAT WERE ADDED BY THE USER
SELECT @nFiles = COUNT(*) FROM myFileNameTable
PRINT 'Number of files = ' + CONVERT(char(5), @nFiles)

-- Specify the information for the nodes
-- Declare temporary VARIABLES
PRINT 'Begin SCRIPT ...'
DECLARE @counterFile INT
SET @counterFile = 1
DECLARE @temp1 varchar(300)
DECLARE @temp2 varchar(300)
DECLARE @temp3 varchar(2056)
CLOSE file_cursor
DEALLOCATE file_cursor
DECLARE file_cursor CURSOR FOR SELECT * FROM myFileNameTable
-- Specify THE information for THE nodes
DECLARE @nNodes int
SET @nNodes = 20
DECLARE @counterNode int

-- Main Loop, for each run
OPEN file_cursor
WHILE (@counterFile <= @nFiles) BEGIN

    -- Reset the counter for each run
    SET @counterNode = 2

    -- Retrieve, from the temporary table
    -- the name of the run
    FETCH NEXT FROM file_cursor INTO @Filename

    -- Create the folder for the run
    SET @temp1 = 'mkdir ' + @folderDest + @Filename
    EXEC master..xp_cmdshell @temp1

    -- Attach the *.mdf file that contains the latency table
    PRINT 'Attach DATABASE ...'
    SET @temp1 = @folderSource + @Filename +
        '\dbs1\lbtb_log_std.mdf'
    SET @temp2 = 'lbtb_log_std'
    EXEC sp_attach_db @temp2, @temp1

    -- EXPORT THE Latency TABLE TO A FILE
    PRINT 'Exporting LATENCY_LOG table ...'

    -- Creates the temp table for Latency
    DROP TABLE ##tableTemp
    CREATE TABLE ##tableTemp
    (
        logging_node_id numeric (15,0),
        source_ftn_id numeric (15,0),
        dest_ftn_id numeric (15,0),
        type varchar (100) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
        priority int,
        event_id numeric (15,0),
        rep_id varchar (50) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
        tx_time varchar (40) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
        logging_datetime datetime)

```

Figure B.1: "GenerateExcelFiles.sql" Script (Part I)



```

-- Inserts in the temp table the rows that are in latency_log table and adds
-- the fields for type of replication and priority
INSERT INTO ##TableTemp
SELECT a.logging_node_id, a.source_ftn_id, a.dest_ftn_id, c.type, b.priority,
       c.event_id, a.rep_id, a.tx_time, d.logging_datetime
FROM
  lbtb_log_std.dbo.latency_log a, lbtb_log_std.dbo.rep_msg_tx_log e, lbtb_log_std.dbo.rep_msg_rx_log d,
  lbtb_log_std.dbo.rtl_packet_log f, lbtb_log_std.dbo.rtl_fragment_log g, lbtb_log_std.dbo.rtl_frame_log h,
  lbtb_log_std.dbo.hids_packet_log b, lbtb_log_std.dbo.event_list c
WHERE
  a.tx_time != '' and a.source_ftn_id = e.src_node_id and a.source_ftn_id = d.src_node_id and
  a.source_ftn_id = f.logging_node_id and a.source_ftn_id = g.logging_node_id and a.source_ftn_id = h.logging_node_id and
  a.source_ftn_id = b.logging_node_id and a.logging_node_id = d.logging_node_id and a.rep_id = e.replication_id and
  a.rep_id = d.replication_id and e.event_id = c.event_id and d.event_id = c.event_id and
  e.replication_id = f.packet_id and f.tx_status = 'CREATE' and f.packet_id = g.data_packet_id and
  g.tx_status = 'ATTEMPT' and g.tx_attempt = '1' and g.logical_frame_id = h.logical_frame_id and
  g.logging_time = h.logging_time and h.tx_status = 'ATTEMPT' and h.packet_id = b.data_packet_id and
  h.logging_time = b.logging_time and b.tx_status = 'SUBMIT'

-- Creates the IIM latency log raw table which contains the info about the
-- event that are of type IIM TRANSFER DATA, without the priority
DROP TABLE ##iim_latency_log_raw
CREATE TABLE ##iim_latency_log_raw
(
  logging_node_id numeric (15,0),
  source_ftn_id   numeric (15,0),
  dest_ftn_id    numeric (15,0),
  type           varchar (100) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
  priority       int,
  event_id      numeric (15,0),
  rep_id        varchar (50) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
  tx_time       varchar (40) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
  send_time     varchar (50) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
  logging_datetime datetime)

-- Insert data in IIM latency log raw table
INSERT INTO ##iim_latency_log_raw
SELECT
  a.logging_node_id, a.src_node_id,
  a.logging_node_id, c.type, '0', a.event_id
  a.replication_id, a.arrival_time,
  b.send_time, a.logging_datetime
FROM
  lbtb_log_std.dbo.rep_msg_rx_log a,
  lbtb_log_std.dbo.rep_msg_tx_log b,
  lbtb_log_std.dbo.event_list c
WHERE
  CAST(c.type AS CHAR(17)) like 'IIM Transfer Data' and
  c.event_id = a.event_id and c.event_id = b.event_id and
  c.node_id = a.src_node_id and c.node_id = b.src_node_id and
  a.replication_id = b.replication_id

UPDATE ##iim_latency_log_raw
SET tx_time = dbo.TimeDiff(send_time, tx_time)

-- Creates the IIM Latency Log table which will contain the info
-- of the IIM data (IIM latency log raw) and populate the values
-- in the column "priority"
DROP TABLE ##iim_latency_log
CREATE TABLE ##iim_latency_log
(
  logging_node_id numeric (15,0),
  source_ftn_id   numeric (15,0),
  dest_ftn_id    numeric (15,0),
  type           varchar (100) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
  priority       int,
  event_id      numeric (15,0),
  rep_id        varchar (50) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
  tx_time       varchar (40) COLLATE SQL_Latin1_General_Cp1_CI_AS NOT NULL,
  logging_datetime datetime);

-- Populate the priority field
-- For now, it does not work
SELECT * FROM ##iim_latency_log

```

**Figure B.2: "GenerateExcelFiles.sql" Script (Part II)**

```

INSERT INTO ##iim_latency_log
SELECT a.logging_node_id, a.source_ftn_id, a.dest_ftn_id, a.type, g.priority,
      a.event_id, a.rep_id, a.tx_time, a.logging_datetime
FROM
##iim_latency_log_raw a, lbtb_log_std.dbo.event_list c,
lbtb_log_std.dbo.rep_msg_tx_log b, lbtb_log_std.dbo.rtl_packet_log d,
lbtb_log_std.dbo.rtl_fragment_log e, lbtb_log_std.dbo.rtl_frame_log f,
lbtb_log_std.dbo.hids_packet_log g

WHERE
a.tx_time != " and c.event_id = a.event_id and
c.event_id = b.event_id and a.source_ftn_id = b.src_node_id and
a.source_ftn_id = d.logging_node_id and a.source_ftn_id = e.logging_node_id and
a.source_ftn_id = f.logging_node_id and a.source_ftn_id = g.logging_node_id and
a.rep_id = b.replication_id and b.replication_id = d.packet_id and
d.tx_status = 'CREATE' and d.packet_id = e.data_packet_id and
e.tx_status = 'ATTEMPT' and e.tx_attempt = '1' and
e.logical_frame_id = f.logical_frame_id and
f.tx_status = 'ATTEMPT' and f.packet_id = g.data_packet_id and
f.logging_time = g.logging_time and g.tx_status = 'SUBMIT'

INSERT INTO ##tableTemp
SELECT * FROM ##iim_latency_log

SET @temp3 = 'bcp ' + "SELECT DISTINCT * FROM ##tableTemp" +
' queryout "' + @folderDest + @Filename +
'\LATENCY_LOG.xls' -c -U' + @uname + '-P' + @pwd
EXEC master..xp_cmdshell @temp3
DROP TABLE ##tableTemp
DROP TABLE ##iim_latency_log_raw
DROP TABLE ##iim_latency_log

-- EXPORT THE CURRENCY TABLE TO A FILE
PRINT 'Exporting CURRENCY table ...'

IF EXISTS (SELECT * FROM INFORMATION_SCHEMA.tables WHERE table_name = 'tableTemp2') BEGIN
    DROP TABLE tableTemp2
END

CREATE TABLE tableTemp2
(
    logging_node_id numeric (15,0),
    update_time varchar(40),
    local_ftn_id numeric (15,0),
    source_ftn_id numeric (15,0),
    event_id numeric (15,0),
    type varchar (100),
    priority int,
    rep_id varchar(50),
    time_since_update varchar(40))

INSERT INTO tableTemp2
SELECT
    a.logging_node_id, a.update_time, a.local_ftn_id, a.source_ftn_id,
    a.event_id, c.type, '3', b.replication_id, a.time_since_update
FROM
    lbtb_log_std.dbo.currency_log a,
    lbtb_log_std.dbo.rep_msg_tx_log b,
    lbtb_log_std.dbo.event_list c
WHERE
    a.update_time != " and
    a.source_ftn_id = c.node_id and
    a.source_ftn_id = b.src_node_id and
    a.event_id = c.event_id and
    a.event_id = b.event_id

IF EXISTS (SELECT * FROM INFORMATION_SCHEMA.tables WHERE table_name = 'iim_currency_log_raw') BEGIN
    DROP TABLE iim_currency_log_raw
END

CREATE TABLE iim_currency_log_raw
(
    logging_node_id numeric (15,0), update_time varchar(40),
    local_ftn_id numeric (15,0), source_ftn_id numeric (15,0),
    event_id numeric (15,0), type varchar (100),
    priority int, rep_id varchar(50),
    time_since_update varchar(40), sequence bigint );

```

**Figure B.3:** "GenerateExcelFiles.sql" Script (Part III)

```

INSERT INTO iim_currency_log_raw(logging_node_id, update_time, local_ftn_id, source_ftn_id, event_id,
                                type, priority, rep_id, time_since_update, sequence)

SELECT
    a.logging_node_id, a.recv_time, a.logging_node_id, a.src_node_id,
    a.event_id, CAST(c.type AS VARCHAR(100)), '3', b.replication_id, a.recv_time, a.sequence
FROM
    lbtb_log_std.dbo.iim_rx_log a,
    lbtb_log_std.dbo.rep_msg_rx_log b,
    lbtb_log_std.dbo.event_list c
WHERE
    CAST(c.type AS CHAR(17)) like 'IIM Transfer Data' AND a.event_type like 'Recursive Data Transfer with no Ack'
    AND c.event_id = a.event_id AND c.event_id = b.event_id
    AND c.node_id = a.src_node_id AND c.node_id = b.src_node_id
    AND a.logging_node_id = b.logging_node_id AND a.logging_datetime = b.logging_datetime

IF EXISTS (SELECT * FROM INFORMATION_SCHEMA.tables WHERE table_name = 'iim_currency_log ') BEGIN
    DROP TABLE iim_currency_log
END

CREATE TABLE iim_currency_log
(
    row_id          INTEGER IDENTITY, logging_node_id NUMERIC (15,0),
    update_time     VARCHAR (40), local_ftn_id      NUMERIC (15,0),
    source_ftn_id   NUMERIC (15,0), event_id        NUMERIC (15,0),
    type            VARCHAR (100), priority         INT,
    rep_id          VARCHAR(50), time_since_update VARCHAR (40),
    sequence        BIGINT );

INSERT INTO iim_currency_log (logging_node_id, update_time, local_ftn_id,
                              source_ftn_id, event_id, type, priority, rep_id, time_since_update, sequence)
SELECT * FROM iim_currency_log_raw
ORDER BY logging_node_id, event_id, sequence

UPDATE a
SET a.time_since_update = b.time_since_update
FROM iim_currency_log a, iim_currency_log b
WHERE a.logging_node_id = b.logging_node_id
    AND a.event_id = b.event_id
    AND a.row_id = b.row_id + 1

UPDATE iim_currency_log
SET time_since_update = dbo.TimeDiff (time_since_update, update_time)

INSERT INTO tableTemp2
SELECT logging_node_id, update_time, local_ftn_id, source_ftn_id, event_id, type,
       priority, rep_id, time_since_update
FROM iim_currency_log

SET @temp3 = 'bcp ' + "'SELECT DISTINCT * FROM tableTemp2'" +
' queryout "' + @folderDest + @Filename +
'\CURRENCY_LOG.xls" -c -U' + @uname + '-P' + @pwd
EXEC master..xp_cmdshell @temp3
DROP TABLE tableTemp2
DROP TABLE iim_currency_log_raw
DROP TABLE iim_currency_log

-- EXPORT THE LOCATION FIDELITY MEASUREMENT VIEW
PRINT 'Exporting LOCATION_FIDELITY_MEASUREMENT view ...'
SET @temp3 = 'bcp ' + "'Use ' + @temp2 + ' SELECT * FROM DBO.LOC_FIDELITY_MEASUREMENT '" +
' queryout "' + @folderDest + @Filename +
'\LF.xls" -c -U' + @uname + '-P' + @pwd
EXEC master..xp_cmdshell @temp3

-- EXPORT THE NODE_LIST TABLE
PRINT 'Exporting NODE_LIST table ...'
SET @temp3 = 'bcp ' + "'Use ' + @temp2 + ' SELECT * FROM DBO.NODE_LIST '" +
' queryout "' + @folderDest + @Filename +
'\NODE.xls" -c -U' + @uname + '-P' + @pwd
EXEC master..xp_cmdshell @temp3

-- Detach the *.mdf file that contains the latency table
PRINT 'Detach DATABASE ...'
EXEC sp_detach_db @temp2, 'true'

SET @counterFile = @counterFile + 1
END

-- GET rid off THE cursor
CLOSE file_cursor
DEALLOCATE file_cursor

DROP TABLE myFileNameTable
PRINT 'End SCRIPT ...'

```

**Figure B.4:** "GenerateExcelFiles.sql" Script (Part IV)

## Annex C: Graphical User Interface

This annex includes three figures (B.1, B.2 and B.3) corresponding to screenshots of the Graphical User Interface (GUI) of the operational measure tool called "OpsMeasures.xls" that was developed to calculate the MOPs for the HCTCN project.

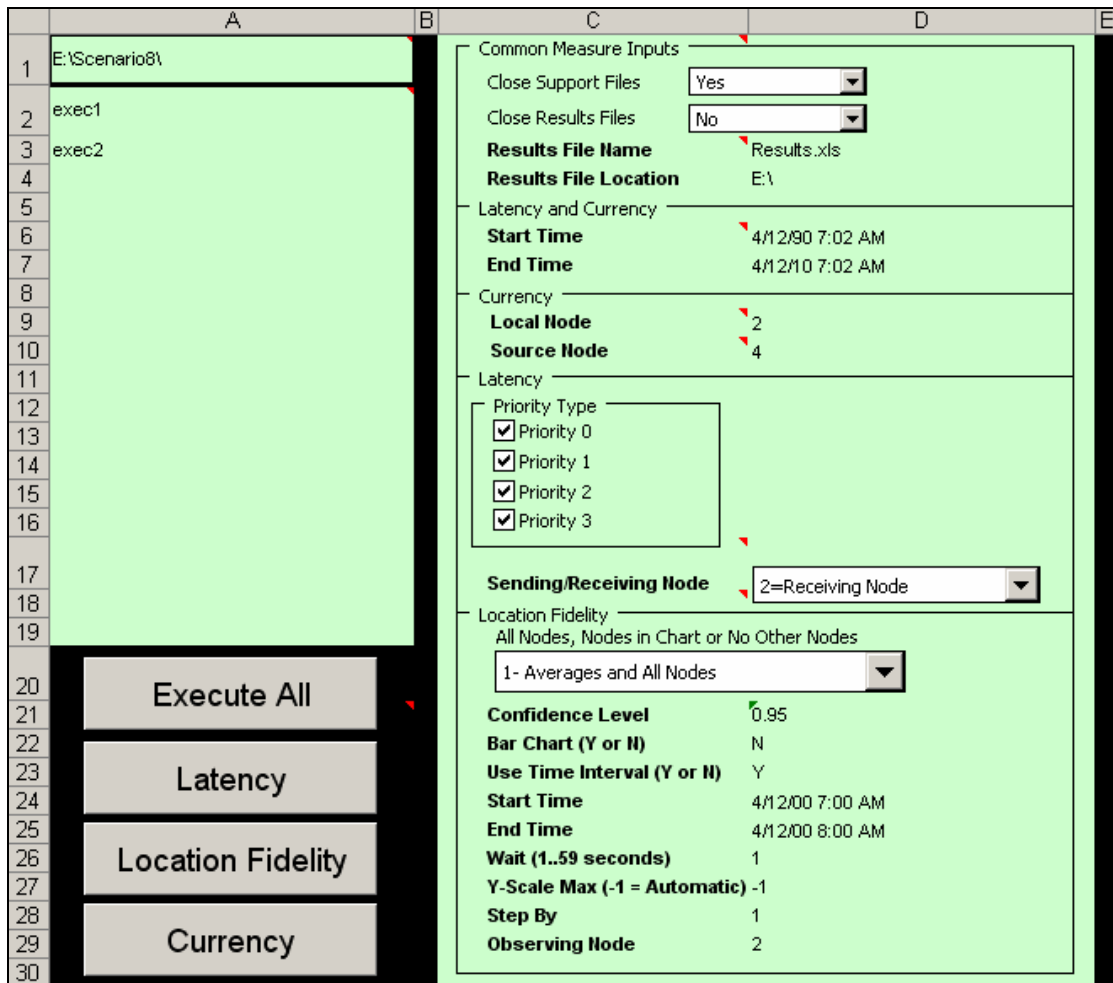


Figure C.1: Graphical User Interface – Part 1

	F	G	H	I	J	K
1	<b>Graphs for Location Fidelity</b>				<b>IIM Data</b>	
2	<b>Observing</b>	<b>Perceived</b>	<b>Chart</b>		<b>Event ID</b>	<b>IER Type</b>
3	2	3	1		6	Vector Movement
4	2	4	1		11	Vector Movement
5	2	5	1			
6	All	Average	2			
7	All	LB	2			
8	All	UB	2			
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						

**Figure C.2:** Graphical User Interface – Part 2

	A	B
31	<div style="border: 1px solid black; padding: 5px;"> <p>Replication Type</p> <input checked="" type="checkbox"/> Position Update  <input checked="" type="checkbox"/> Vector Movement  <input type="checkbox"/> Situation Report  <input type="checkbox"/> Movement Order  <input type="checkbox"/> Fire Mission  <input type="checkbox"/> CPR  <input type="checkbox"/> Contact Report  <input type="checkbox"/> Discrete Voice Message  <input type="checkbox"/> Fragmented Order  <input checked="" type="checkbox"/> IIM Transfer Data  <input type="checkbox"/> Intelligence Report  <input type="checkbox"/> Patrol Report  <input type="checkbox"/> IIM Data Request </div>	
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		

**Figure C.3:** Graphical User Interface – Part 3

## List of Symbols/Abbreviations/Acronyms

---

CS	Central Server
DRDC	Defence Research and Development Canada
FTN	Foreground Tactical Node
GUI	Graphical User Interface
HCTCN	High Capacity Technical Communications Network
IM	Information Management
LBTB	Low Bandwidth Test Bed
MOP	Measure of Performance
MP	Mesure de performance
ODB PDU	Operation Database Protocol Data Unit
OR	Operational Research
PPCT	Probability Plot Correlation Test
RDDC	Recherche et Développement pour la Défense Canada
RO	Recherche Opérationnelle
SQL	Structured Query Language
TDP	Technology Demonstration Project
Tx	Transmission
VBA	Visual Basic for Applications

## Distribution List

---

### Internal

1-Director General  
3-Document Library  
1-J.-D. Caron (Author)  
1-L. Stemate (Author)  
1-A. Gibb  
1-J.-C. St-Jacques  
1- N. Chafai  
1-Operational Research Team (Attn. L. Stemate)

### External

1-DRDKIM (PDF file)

1-M. Rey  
Director General DRDC-CORA  
National Defence Headquarters  
Ottawa ON K1A 0K2

1-J. Evans  
Director Operational Research (Maritime/Land/Air)  
National Defence Headquarters  
Ottawa ON K1A 0K2

1-B. Dickinson  
Director Operational Research (Joint)  
National Defence Headquarters  
Ottawa ON K1A 0K2

1-Dean Haslip  
Section Head (Land OR)  
National Defence Headquarters  
Ottawa ON K1A 0K2

1-B. Nguyen  
DRDC-Atlantic  
9 Grove Street  
Dartmouth, NS B2Y 3Z7



1-J. Schlesak  
Section Head  
Communications Research Center  
3701 Carling Avenue  
P.O. Box 11490, Station H  
Ottawa, ON K2H 8S2

1-DRDC CORA Library  
Attn. Renita Repsys  
National Defence Headquarters, 6 CBS  
Ottawa, ON K1A 0K2

This page intentionally left blank.

UNCLASSIFIED  
 SECURITY CLASSIFICATION OF FORM  
 (Highest Classification of Title, Abstract, Keywords)

<b>DOCUMENT CONTROL DATA</b>		
1. ORIGINATOR (name and address) DRDC Valcartier, Operational Research Group 2459, Pie XI Nord, Québec Québec (QC) G3J 1X5	2. SECURITY CLASSIFICATION (Including special warning terms if applicable) UNCLASSIFIED	
3. TITLE (Its classification should be indicated by the appropriate abbreviation (S, C, R or U) Measures of performance for analysis of tactical communication simulations conducted using a low bandwidth test bed (U)		
4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.) Caron, Jean-Denis and Stemate, Luminita		
5. DATE OF PUBLICATION (month and year) February 2007	6a. NO. OF PAGES 70	6b. NO. OF REFERENCES 7
7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. Give the inclusive dates when a specific reporting period is covered.) Technical Memorandum		
8. SPONSORING ACTIVITY (name and address) High Capacity Tactical Communications Network (HCTCN) – Technology Demonstration		
9a. PROJECT OR GRANT NO. (Please specify whether project or grant) Technology Demonstration Project (12oc)	9b. CONTRACT NO. N/A	
10a. ORIGINATOR'S DOCUMENT NUMBER DRDC Technical Memorandum TM 2006-791	10b. OTHER DOCUMENT NOS N/A	
11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)		
<input checked="" type="checkbox"/> Unlimited distribution <input type="checkbox"/> Restricted to contractors in approved countries (specify) <input type="checkbox"/> Restricted to Canadian contractors (with need-to-know) <input type="checkbox"/> Restricted to Government (with need-to-know) <input type="checkbox"/> Restricted to Defense departments <input type="checkbox"/> Others		
12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)		

UNCLASSIFIED  
SECURITY CLASSIFICATION OF FORM  
(Highest Classification of Title, Abstract, Keywords)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

This memorandum summarizes the three-fold Operational Research (OR) support provided to the Technology Demonstration Project entitled High Capacity Technical Communications Network (HCTCN), by the Defence Research and Development Canada Valcartier OR team. First, refinements were proposed to three basic measures of performances implemented in the Low Bandwidth Test Bed simulation environment developed for this project: location fidelity, currency and latency. Second chronologically, but the main effort of this study was to develop a tool that implements these refined measures. This tool, fully documented in this document, including a user's guide, represents one of the key requirements for the successful analysis and delivery of the results obtained through experimentation by the HCTCN project team. Third, the memorandum also addresses the question of the number of times each scenario needs to be executed within the test bed in order to ensure statistically valid results.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Communication System  
Radio  
Measure of performance  
MOP  
Location fidelity  
Latency  
Currency  
Sample Size  
Number of Runs  
Independent Replications Method  
Design Experimentation



## **Defence R&D Canada**

Canada's Leader in Defence  
and National Security  
Science and Technology

## **R & D pour la défense Canada**

Chef de file au Canada en matière  
de science et de technologie pour  
la défense et la sécurité nationale



[WWW.drdc-rddc.gc.ca](http://WWW.drdc-rddc.gc.ca)

