View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Calhoun, Institutional Archive of the Naval Postgraduate School

Calhoun
Institutional Archive of the Naval Postgraduate School

Calhoun: The NPS Institutional Archive

Theses and Dissertations | Thesis Collection

2013-09

Combining multiple types of intelligence to generate probability maps of moving targets

Zlatsin, Philip

Monterey, California: Naval Postgraduate School

http://hdl.handle.net/10945/37751

DUDLEY KNOX LIBRARY

Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

COMBINING MULTIPLE TYPES OF INTELLIGENCE TO GENERATE PROBABILITY MAPS OF MOVING TARGETS

by

Philip Zlatsin

September 2013

Thesis Advisors: Moshe Kress
Michael Atkinson

Second Reader: Nedialko B. Dimitrov

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** September 2013	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE COMBINING MULTIPLE TYPES OF INTELLIGENCE TO GENERATE PROBABILITY MAPS OF MOVING TARGETS			**5. FUNDING NUMBERS**
6. AUTHOR(S) Philip Zlatsin			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____N/A____.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

Drug addiction in the United States generates significant health, economic, and social costs. One of the prominent ways in which traffickers smuggle drugs into the United States is by maritime shipments from South America. In 1989 Joint Interagency Task Force South (JIATF-S) was established to fight these traffickers. JIATF-S collects information from multiple sources, which can be broadly classified into two categories. The first category is sensor-based sources that produce observations about possible targets (e.g., radar, sonar). These observations provide precise location and time but are susceptible to false positive and false negative errors regarding their content. The second category is human-based sources, including tips, messages and intercepted communications among humans. In addition to possible misinformation regarding the content of an event, such inputs are also susceptible to errors regarding the location and time of the event.

In this thesis we develop a data fusion model that can assist JIATF-S in estimating the likelihood that a certain target (i.e., drug-smuggling vessel) is present at a certain location at a certain time and evaluate the reliability of the information source. The novelty of this thesis is manifested in a new probabilistic approach for utilizing human-generated intelligence, and in the way it is combined with sensor-generated intelligence.

14. SUBJECT TERMS Intelligence Cycle, Intelligence Processing, Data Fusion, Dempster-Shafer, HUMINT, SIGINT, Bayesian Update			**15. NUMBER OF PAGES** 105
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

COMBINING MULTIPLE TYPES OF INTELLIGENCE TO GENERATE PROBABILITY MAPS OF MOVING TARGETS

Philip Zlatsin
Captain, Israel Defense Forces
B.Sc., The Hebrew University of Jerusalem, 2006

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2013

Author: Philip Zlatsin

Approved by: Moshe Kress
Thesis Advisor

Michael Atkinson
Thesis Advisor

Nedialko B. Dimitrov
Second Reader

Robert F. Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Drug addiction in the United States generates significant health, economic, and social costs. One of the prominent ways in which traffickers smuggle drugs into the United States is by maritime shipments from South America. In 1989 Joint Interagency Task Force South (JIATF-S) was established to fight these traffickers. JIATF-S collects information from multiple sources, which can be broadly classified into two categories. The first category is sensor-based sources that produce observations about possible targets (e.g., radar, sonar). These observations provide precise location and time but are susceptible to false positive and false negative errors regarding their content. The second category is human-based sources, including tips, messages and intercepted communications among humans. In addition to possible misinformation regarding the content of an event, such inputs are also susceptible to errors regarding the location and time of the event.

In this thesis we develop a data fusion model that can assist JIATF-S in estimating the likelihood that a certain target (i.e., drug-smuggling vessel) is present at a certain location at a certain time and evaluate the reliability of the information source.

The novelty of this thesis is manifested in a new probabilistic approach for utilizing human-generated intelligence, and in the way it is combined with sensor-generated intelligence.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

COMINT	Communication Intelligence
DoD	Department of Defense
DST	Dempster-Shafer Theory
ELINT	Electronic Intelligence
HUMINT	Human Intelligence
JDL	Joint Directors of Laboratories
JIATF-S	Joint Interagency Task Force South
MOE	Measure Of Effectiveness
NDIC	The National Drug Intelligence Center
SIGINT	Signal Intelligence
SPSS	Self-Propelled Semi-Submersible
VISINT	Visual Intelligence

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Data fusion from various sources is a common problem for intelligence organizations around the world. In this thesis we explore the efforts of the Joint Interagency Task Force South, an organization established in 1989 to fight drug traffickers originating from South America, to combine different sources of intelligence into a coherent picture to seize the smuggled drugs.

In this thesis we examine the combination of two categories of intelligence sources regarding drug smugglers: (1) sensor-based sources such SIGINT (signal intelligence) and VISINT (visual intelligence), and (2) human-based sources such as HUMINT (human intelligent) and COMINT (communication intelligence). Sensor-based sources typically have high precision regarding location and time of an observation but are susceptible to false positive and false negative errors. Human-based sources, including tips, messages and communications generated by humans are susceptible to these same errors. In addition to possible misinformation regarding the description of a reported event, these sources also tend to have low precision regarding the location and time of the event.

We explore several methods for combining information from sensor-based and human-based sources. In addition to the traditional Bayesian update mechanism, which is commonly used for sensor fusion, we also examine applying Dempster-Shafer theory. The Bayes' method is mathematically rigorous but requires a number of assumptions not needed for the Dempster-Shafer methods, namely assuming that the distribution of the messages received from the informant is known, and uniform. The Dempster-Shafer theory does not make those assumptions explicitly. Moreover, there are several ways to implement the Dempster-Shafer theory, and it is not clear in advance which implementation would be most appropriate for a given scenario. We compare the methods both qualitatively and quantitatively using a simulation.

Our analysis shows that even when the assumptions of the Bayes' update process are violated, it still manages to yield the best results. The Dempster-Shafer methods did

not perform better than Bayes even though they do not explicitly make as many assumptions as the Bayes update. As expected, when the reliability of the informant is low or is mistaken to be low, and there is non-uniformity in the way he produces messages, all the methods performed poorly.

In addition we develop a Bayesian model to assess the quality of the informant and update a vessel's location simultaneously. We formulate update procedures both when the informants' messages can be verified and when they cannot be verified and we must rely only on the current perception about the location of the vessel and the informant's reliability for the update. We suggest a combined scheme that allows simultaneous estimation of both the location of the vessel and the reliability of the informant as new information becomes available.

ACKNOWLEDGMENTS

I wish to thank my advisors, Moshe Kress and Michael Atkinson, for the hours and effort spent in guiding me and improving my work and to my second reader, Nedialko Dimitrov, for his inspiring ideas.

I wish also to thank my wife, Rebecca, for her love and support.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. MOTIVATION

Drug addiction in the United States generates significant health, economic, and social costs. According to the National Drug Intelligence Center (NDIC) report from 2011, “In 2007 alone, the estimated cost of illicit drug use to society was $193 billion, including direct and indirect public costs related to crime, health, and productivity… an increasing number of individuals, particularly young adults, are abusing illicit drugs. In 2009, an estimated 8.7 percent of Americans aged 12 or older (21.8 million individuals) were current illicit drug users” (NDIC, 2011).

One of the primary ways drugs arrive to the United States is via smuggled maritime shipments from South America. Most of the marijuana seized and more than 50% of the methamphetamine, cocaine and heroin seized are detected on the Southwest border (NDIC, 2011). Figure 1 illustrates that more than 99% of the cocaine flow from South America to the United States in 2007 was smuggled through the Caribbean Sea or Pacific Ocean via Mexico (Palter, 2009).

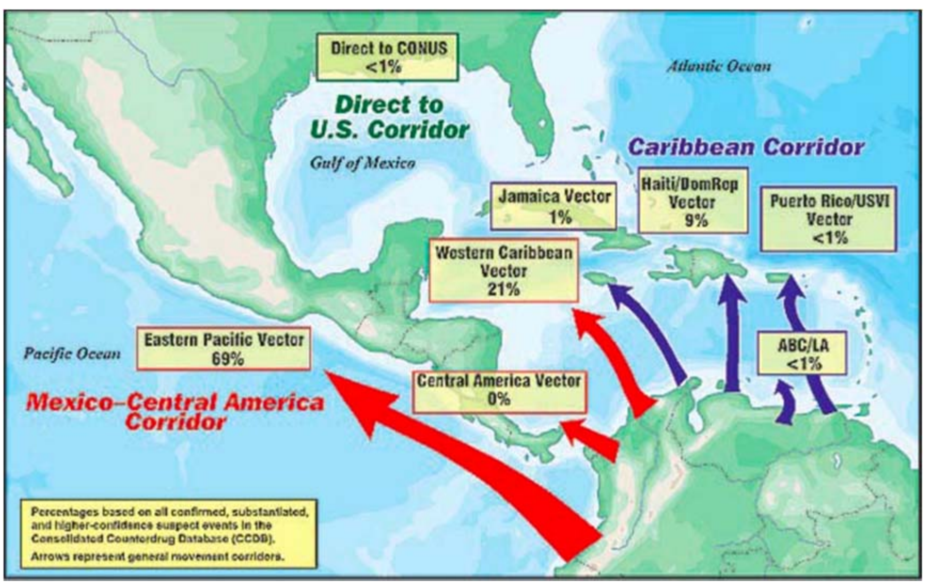

Figure 1. Northward-bound Cocaine Flows (From Palter, 2009).

In October 1989 the Joint Interagency Task Force South (JIATF-S) was established in order to fight these traffickers. JIATF-S is a multiservice, multiagency national task force that conducts counter-illicit trafficking operations, intelligence fusion, and multi-sensor correlation to detect, monitor, and hand off suspected illicit trafficking targets (JIATF-S, 2013). JIATF-S collects information from multiple sources with different characteristics and of different quality. The types of information collected can be broadly classified into two categories:

1) Sensor-based intelligence: Sensor observations are typically characterized by high precision regarding location and time of the observation but are also susceptible to false positive and false negative errors regarding its outcome. Typical examples for sensor-based observation sources are electronic intelligence (ELINT), electronic intelligence obtained from sources such as RADAR and non-content signals from communication devices, and visual intelligence (VISINT), visual intelligence such as video, images and the naked eye.

2) Human-based intelligence: Tips, messages and communications generated by humans, which in addition to errors in content are also susceptible to low precision regarding the location and time of the event. The error rate depends upon the reliability of the source, which is less well defined than the error rates of sensors. Typical examples of human-based intelligence (HUMINT) are intelligence (such as tips and messages) gathered from human sources, and communications intelligence (COMINT) are content-based intelligence gathered from intercepted communications.

A main challenge for JIATF-S is the integration of information about different spatial locations and time ranges from multiple sources in a consistent and coherent manner in order to locate illicit drugs trafficking vessels.

In this thesis we develop data fusion techniques to assist JIATF-S in estimating the likelihood that a certain target (i.e., a drug-smuggling vessel) is present at a certain location at a certain time. This information can provide JIATF-S with better situational awareness and inform decision makers about where they choose to send search and interdiction assets. More specifically, we provide a probability distribution for the location and departure time of possible targets. We also evaluate the quality of the information sources in order to give their inputs the proper weight (e.g., the reliability of a human informant).

B. CONTRIBUTIONS OF THIS WORK

Exploitation of human intelligence for targeting has been known since ancient times and discussed by many strategists and military theorists, as described for example in the famous "The Art of War" written by Sun Tzu in the 6th century BC. Processing HUMINT intelligence into spatial information, for instance, to create crime maps for policing applications is also known (Ratcliffe, 2000)

There is also vast literature on the fusion of sensor-based information (Hall & Llinas, 2001); however, there is still a need to combine human intelligence with intelligence from other sources - a process that it is traditionally done manually by trained professionals due to the difficulties of implementing automated algorithms.

In this work we suggest automated algorithms to combine human-based intelligence and sensor-based intelligence in a single framework. We also provide a way to estimate and update the perceived reliability of our sources.

C. THE INTELLIGENCE PROCESS

Since this thesis relates to the effective utilization of intelligence, it is useful to frame this thesis within the intelligence processing paradigm. The intelligence process comprises the following six categories of intelligence operations (United States Dept. of the Army, 2007):

- Planning and direction - Planning operations to acquire new or better data or develop intelligence sources.
- Collection - Acquisition of the required data.
- Processing and exploitation – Converting the collected raw data into information that can be used by commanders.
- Analysis and production – Analyzing the information and producing higher-level intelligence from the information gathered via interpretation and integration with other relevant information.

- Dissemination and integration – Disseminating the intelligence to appropriate users.

- Evaluation and feedback – Evaluation of the intelligence performance.

The following figure represents graphically the intelligence process:

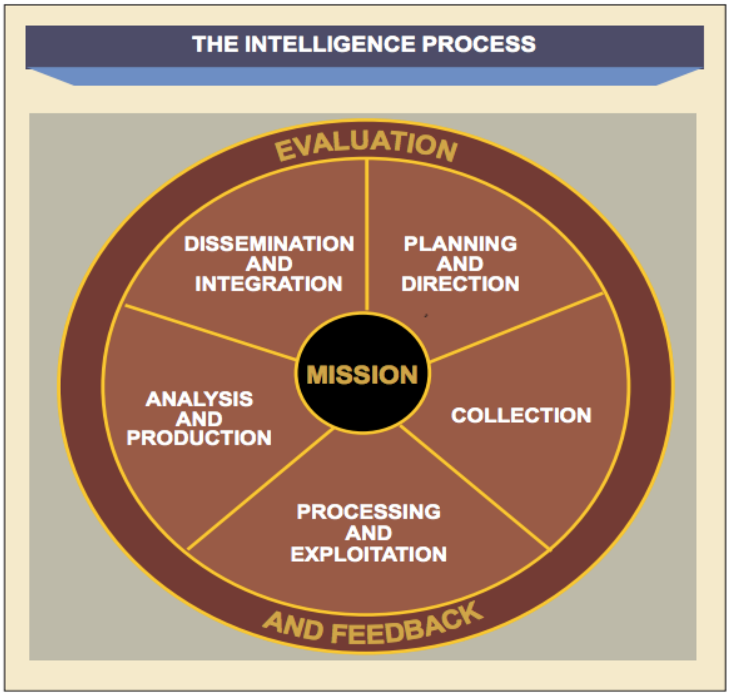

Figure 2. The Intelligence Process (From United States Dept. of the Army, 2007).

Since the evaluation and feedback component is used to initialize the other intelligence activities, the intelligence process is also referred to as the "intelligence cycle." In this thesis, we concentrate on the processing stage and examine how to effectively integrate new pieces of information into the intelligence profile using data fusion methods.

Combining information from different data sources is commonly called data fusion, which is defined as a "process dealing with the association, correlation, and combination of data and information from single and multiple sources to achieve refined position and identity estimates, and complete and timely assessments of situations and threats as well as their significance" (White, 1991).

In this thesis, we concentrate on combining the two types of inputs - sensor data and HUMINT input into a single fused intelligence picture. The feedback between these two sources can be used, in turn, to reevaluate the quality of the intelligence obtained from each source, thus improving the fusion of future intelligence. The following figure illustrates, in simple terms, the processing problem considered in this thesis.

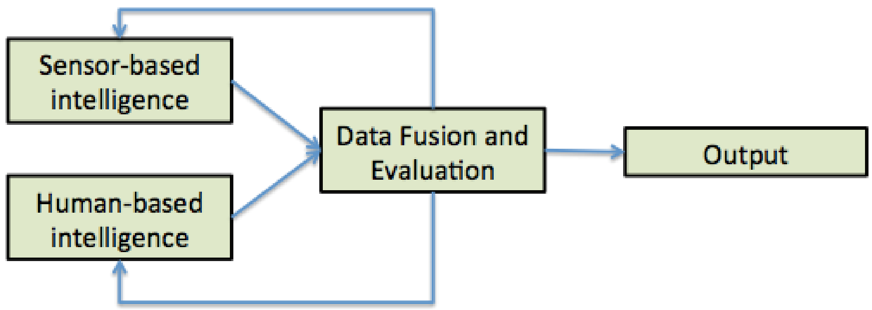

Figure 3. Simplified intelligence process model.

D. DATA FUSION

Data fusion may relate to fusion of information of different levels and for different purposes. In order to encompass and categorize the different levels, the Data Fusion Subpanel (which later became known as the Data Fusion Group) of the Joint Directors of Laboratories (JDL) developed the Data Fusion model in 1985 (Hall & Llinas, 2001).

The JDL model as revised in (Steinberg, Bowman & White, 1999) categorizes the fusion according to the relation of the information to the entity of interest (in our case the drug smuggling vessel) and the purpose of the outcome of the fusion. The following levels are included in the model:

- Level 0 — Sub-Object Data Assessment: prediction of entities that are not recognized as an object yet, such as pixels and radio signals.
- Level 1 — Object Assessment: estimation and prediction of entity states on the basis of inferences from observations.
- Level 2 — Situation Assessment: estimation and prediction of entity states on the basis of inferred relations among entities.

- Level 3 — Impact Assessment: estimation and prediction of effects on situations of planned or estimated/predicted actions by the participants.
- Level 4 — Process Refinement (an element of Resource Management): adaptive data acquisition and processing to support mission objectives.

The flow of information, from raw measurements to assessment of the entire picture with regard to the JDL model levels is described in the following figure:

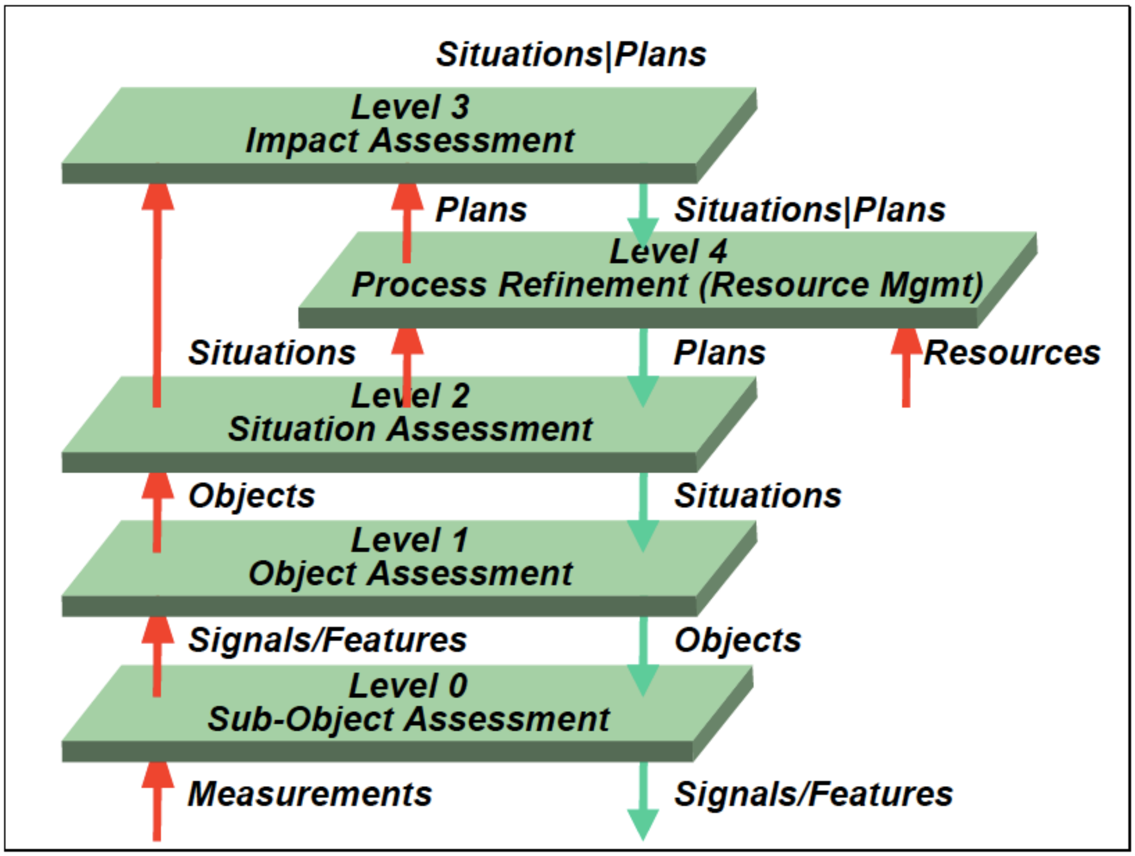

Figure 4. JDL model information flow (From Steinberg, Bowman & White 1999).

In this thesis, we explore levels 0 and 1 of the JDL model combining raw data information from multiple sources in order to estimate a target's location and departure time. Higher data fusion levels, such as assessing the characteristics and intentions of multiple targets are not discussed in this work.

We will examine two information-fusion approaches in this thesis: Bayesian update and Dempster-Shafer Theory.

a. Bayesian Update

Bayes' formula was first introduced in the 18th century (Bayes, 1763), and its applications are found in many fields that range from diagnosing the medical situation of patients (Lincoln & Parker, 1967) to artificial intelligence (Korb & Nicholson, 2011).

Many tracking and location algorithms are based on Bayesian methods, such as the well-known Kalman filter. The manuscript "Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond" (Chen, 2003) includes an exhaustive review of filters and (Morelande, Kreucher & Kastella ,2007) reviews other Bayesian tracking algorithms.

Bayesian methods are also applied extensively to sensor fusion. The book *Handbook of Multisensor Data Fusion* (Hall & Llinas, 2001) presents an overview of data fusion methods that includes several chapters regarding Bayesian updates and treats it as the basic method for data fusion.

In this work, the Bayesian method is used in a similar manner to update the probability that a target is at a particular location at a certain time.

However, Bayesian methods also have limitations. In particular, they require intimate knowledge of sensor capabilities, such as estimates of the error rates, a notion of the distribution of the possible errors of the sensor and assumptions regarding the state of the world, such as a prior distribution. For those reasons, we also consider other information updating methods.

b. Dempster-Shafer Belief Method

Dempster-Shafer theory (DST) was developed by Arthur Dempster and Glenn Shafer (Dempster, 1967; Shafer, 1976). This theory, which is in some sense a generalization of probability theory, allows for assigning "belief" values (and not probabilities) to events and sets of events, thus requiring fewer assumptions and axioms.

Due to the theory's ability to deal with complicated types of variability in belief, Dempster-Shafer theory has been used widely for decision making algorithms and data-fusion. *An Introduction to Bayesian and Dempster-Shafer Data Fusion* (Koks and

Challa, 2003) is a good introductory summary to Dempster-Shafer theory in comparison with Bayesian methods.

Hall and Llinas (2001 also includes several chapters about Dempster-Shafer theory while Sentz (2002) is an extensive report about different Dempster-Shafer methods and their applications.

E. OTHER POSSIBLE APPLICATIONS

In this thesis the framework suggested for locating drug traffickers may be applied to a range of related applications.

1) Locating friendly forces: A similar problem to the one described above on the sea can occur on land as well, when data from sensors such as radars are combined with information from human sources in order to locate a friendly force in need of assistance on the battlefield.

2) Other types of sensors: Traditional updating mechanisms require an intimate knowledge of the technical parameters that determine the performance of a sensor and the environment in which it is used. When this knowledge is lacking, more robust methods, such as the ones explored in this thesis, can be of use. Those methods can be applied not only to SIGINT and HUMINT, but also to other types of intelligence.

F. THESIS OUTLINE

This chapter includes the background and problem description. Chapter II describes the problem, the basic integration methods used, the assumptions, and the details of the Bayesian update and Dempster-Shafer belief theory and their application to the problem. In Chapter III we compare those two models and develop a simulation to gain additional insights. Chapters IV and V include a detailed mathematical framework of extensions to the base model described in Chapter II. While Chapter IV includes a model that handles multiple routes, Chapter V includes a framework that allows estimating and updating the reliability of the sources.

II. THE MODEL

In this chapter we describe the scenario, the theater of operations and the goals of the JIATF-S operator. We define the different types of intelligence received and describe a model that updates the situational awareness regarding this scenario by combining intelligence from different sources together.

A. SCENARIO

We consider a drug-smuggling situation from the northern part of South America to Central America. Drug smugglers may leave from multiple points of embarkation in the northern part of South America towards one of multiple final destinations in Central America and Mexico. The smugglers use three kinds of vessels for their operations:

- GO-FAST small boats - designed to reach high velocities but with relatively low capacities,
- Merchant Vessels - high capacity but slow and easy to detect, and
- SPSS (self-propelled semi-submersible) are partly submersible. These vessels are difficult to detect by radar, but their velocity is slow.

The types of vessels have very different characteristics, and therefore, it is usually easy to distinguish among them. There are also several categories of typical routes the smugglers use:

- Close to the shore in the Pacific ocean,
- Close to the shore in the Caribbean sea,
- In the Pacific ocean, via Galapagos Islands,
- Straight routes between the embarkation point and the destination, and
- Piece-wise linear routes between the embarkation point and the destination. (This category essentially covers all possibilities).

Figure 5 shows examples of those routes.

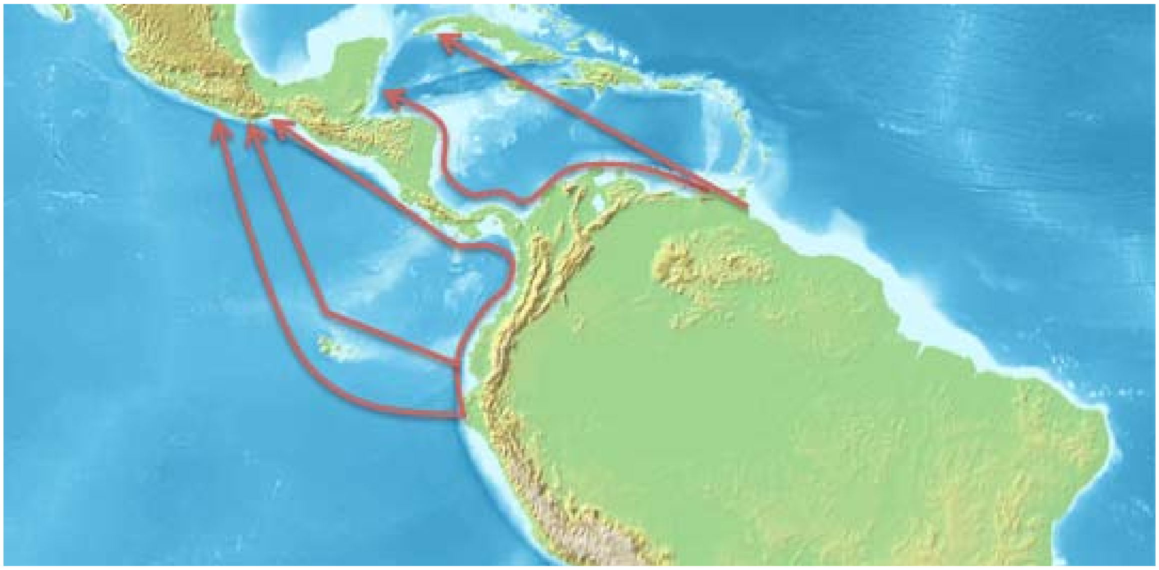

Figure 5. Typical smuggling routes.

JIATF-S operators desire enhanced situational awareness about the location of targets in order to more effectively direct interdicting assets that can seize the smuggled drugs. JIATF-S operators increase situational awareness by obtaining information from sensors such as radars and cameras, as well as from human sources. Often, new intelligence arrives in real time. In those cases, the operators should update their perceived probability that a vessel is located in a specific place according to the new intelligence.

JIATF-S use their situational awareness to send a surveillance aircraft or surface vessels to look for smugglers in the suspected location. If a smuggler is positively identified, the surveillance vehicle holds contact with the smuggler's vessel until a maritime force boards it and confiscates the drugs.

The operators must also evaluate the quality of the different intelligence sources in order to weight their information contributions appropriately.

B. ASSUMPTIONS

For tractability, we initially make the following assumptions, some of which we relax later. These assumptions reduce the problem to estimating a single parameter: departure time.

1. Single Target Vessel

There is at most one target vessel in the theater at any given time. Although in reality multiple vessels might be present in the theater at the same time, it is assumed that JIATF-S's operators are able to associate incoming information with the correct vessel. In other words, in this thesis we do not consider data-association problems that might be a subject for future research.

2. Constant and Known Speed

The speed of the vessel is constant and known. This assumption is reasonable because the speed of the vessel depends mainly on the vessel's type (which is usually known), and may not change much during the course of its movement. Since the velocity of each vessel type is known and its variance is rather small, this assumption is reasonable. Future work may consider variable velocity due to weather conditions, refueling stops, strategic considerations or other factors.

3. Discrete Departure Time Distribution

The set of departure times is discrete and finite; the vessel can leave the harbor at any one of several possible time slots. As the discretization can be as fine as necessary, this assumption does not affect the results of the model. A reasonable size of a time slot would be one to three hours.

4. A Single Known Route

To start, we assume there is only one possible route for simplicity. This assumption is relaxed later on to include multiple routes. Since the route and the speed of the vessel are known and fixed, the location of the vessel is uniquely defined by the time of departure.

C. DEFINITIONS

A random variable, T_d, denotes the departure time of the vessel.

For simplicity we assume that T_d is discrete, and its possible values are in the set $T = \{t_1, t_2, ..., t_n\}$, so there are n possible departure times.

The probability mass function of the departure times is $f(t) \equiv P(T_d = t)$. We assume a prior for this distribution, and our objective is to update this prior as new information arrives. As more intelligence arrives, the posterior should narrow around a handful of most likely departure times to aid in the routing of surveillance aircrafts.

1. Sensor-based Intelligence – Observations

An observation is a random event associated with a certain time, $t' \in T$. An example of an observation might be the event that the operator received a radar reading regarding a certain time and location. Since we assume a single route and fixed velocity, an observation that is made at any location along the route can be trivially translated to an observation made about a perceived departure time. For instance, if the speed is 30 knots, and we have an observation at distance 60 NM on the route at 4 p.m. it is equivalent to an observed departure at 2 p.m. This allows us to locate the ship at any desired time after disembarkation.

The formal definition of an observation is:

$O_{t',+}$ - a positive observation (a vessel departed at time t').

$O_{t',-}$ - a negative observation (no departure at time t').

The observations may be subject to the following errors:

$P(O_{t',+} | T_d \neq t') = P_{f+}$: False positive error, the sensor reports a departure while there is none.

$P(O_{t',-} | T_d = t') = P_{f-}$: False negative error, the sensor fails to detect a departure.

The error probabilities depend on the detection and classification capabilities of the sensor and on the characteristics of the environment. In particular, the probability for false positive error depends on the number of non-target vessels and debris in the area of the point of embarkation.

2. Human-based Intelligence – Messages

While an observation is associated with a specific time of departure, messages are less specific and may include a range of possible departure times. An example of a typical message may be an informant relaying, "I've heard that a ship might embark between 8 a.m. and noon" or perhaps getting a hint via telephone communication that "The drug dealers will leave on one of the following mornings . . ." As mentioned above in Section B, we first consider only the time ambiguity and assume a single known embarkation point. This assumption is relaxed later.

Let M denote the random event that a certain message is received. The sample space of those events (the possible messages) is all the subsets of the departure times $\{t_1, t_2, ..., t_n\} = T$ except for the empty set. Let k denote the cardinality of the message, that is $|M| = k$. Thus, if the event M occurred, the informant claims that the departure time is one of the k values in M.

D. BAYESIAN UPDATE

1. The Update Process

As before, $f(t)$ is the probability mass function of the true departure time, and $P(T_d = t) \equiv f(t)$ is the probability that true departure time is *t*. According to the Bayes' formula, the update probability with new information is defined as:

$$P_{new}(T_d = t \mid \text{New information}) = \frac{P(\text{New information} \mid T_d = t) \cdot P_{prior}(T_d = t)}{P(\text{New information})} \tag{2.1}$$

From Equation (2.1) it follows that in order to calculate the updated probability distribution using Bayes' method, one requires a prior probability $P_{prior}(T_d = t) \equiv f_{prior}(t)$. This probability reflects the prior information we have about the vessel's departure time.

Absent any information we use as the default the uniform prior $f_{prior} \sim U\left[T\right]$. However, if we have some general information about the likelihood of different departure times (for instance, if we know that the likelihood of departing at low tide is much higher than at high tide), we can integrate this knowledge by altering the prior.

Equation (2.1) can be rewritten as a product of an update function and the prior information regarding the distribution: $f_{new}\left(t\right) = f_{update}\left(t\right) \cdot f_{prior}\left(t\right)$, with the update function being:

$$f_{update}\left(t\right) = \frac{P\left(\text{New information} \mid T_d = t\right)}{P\left(\text{New information}\right)} \tag{2.2}$$

In the following chapters we show how to define the update function for different intelligence types.

2. Bayesian Update Following an Observation

As described before, there are two types of observations, positive and negative. We shall calculate the update function for each of those cases.

Applying Equation (2.2) to the case where the new information is a positive observation at time t', and using the law of total probability for the denominator, the updated probability distribution is:

$$f_{update}\left(t\right) = \frac{P\left(\mathrm{O}_{t',+} \mid T_d = t\right)}{\sum_{s \in T} P\left(\mathrm{O}_{t',+} \mid T_d = s\right) \cdot f\left(s\right)} \tag{2.3}$$

By definition of false positive and false negative errors, the probability of receiving a positive observation regarding time t' is $1 - P_{f-}$ if the true departure time is indeed t' and P_{f+} otherwise:

$$P\left(\mathrm{O}_{t',+} \mid T_d = t\right) = \begin{cases} 1 - P_{f-} & t = t' \\ P_{f+} & t \neq t' \end{cases} \tag{2.4}$$

And so we can calculate the total update function, given that we know the values of P_{f+} and P_{f-}:

$$f_{update+}(t) = \begin{cases} \dfrac{1-P_{f-}}{\left(1-P_{f-}\right)\cdot f(t') + P_{f+}\cdot\left(1-f(t')\right)} & t = t' \\ \dfrac{P_{f+}}{\left(1-P_{f-}\right)\cdot f(t') + P_{f+}\cdot\left(1-f(t')\right)} & t \neq t' \end{cases} \tag{2.5}$$

Similarly, we can follow the entire calculation for the case of receiving a negative observation, and we get the following update function:

$$f_{update-}(t) = \begin{cases} \dfrac{P_{f-}}{P_{f-}\cdot f(t') + \left(1-P_{f+}\right)\cdot\left(1-f(t')\right)} & t = t' \\ \dfrac{1-P_{f+}}{P_{f-}\cdot f(t') + \left(1-P_{f+}\right)\cdot\left(1-f(t')\right)} & t \neq t' \end{cases} \tag{2.6}$$

3. Bayesian Update Following a Message

For a given informant and a message of size *k* we define q_k to be the probability that the message is correct: $q_k = P\left(T_d \in M \,||M| = k\right)$. That is, we assume that the quality of the informants depends only on the size of the message. The exact mathematical definition of this parameter will be discussed in the following chapters.

We assume that q_k is monotone non-decreasing in *k* and $q_n = 1$ where the informant gives the entire possible set of departure times as the message. An additional assumption here is that q_k does not depend on the content of the message but only on its size *k*.

The update of the probability distribution of the random variable T_d following a new message *m* is done in a similar way to the observation case:

$$f_{update}\left(t\right)=\frac{P\left(M\mid T_d=t\right)}{\sum_{s\in T}P\left(M\mid T_d=s\right)\cdot f\left(s\right)} \tag{2.7}$$

Differentiating between the case when "*t* is in the message" and "*t* is not in the message" and using the law of total probability, we obtain that the update function is:

$$f_{update}\left(t\right)=\begin{cases}\dfrac{P\left(M\mid T_d=t,t\in M\right)\cdot q_k}{\sum_{s\in M}P\left(M\mid T_d=s,s\in M\right)\cdot q_k\cdot f\left(s\right)+\sum_{s\notin M}P\left(M\mid T_d=s,s\notin M\right)\cdot\left(1-q_k\right)\cdot f\left(s\right)} & t\in M\\ \dfrac{P\left(M\mid T_d=t,t\notin M\right)\cdot\left(1-q_k\right)}{\sum_{s\in M}P\left(M\mid T_d=s,s\in M\right)\cdot q_k\cdot f\left(s\right)+\sum_{s\notin M}P\left(M\mid T_d=s,s\notin M\right)\cdot\left(1-q_k\right)\cdot f\left(s\right)} & t\notin M\end{cases} \tag{2.8}$$

However, in order to calculate the value of $f_{update}\left(t\right)$ we must know the values of the expressions $P\left(M\mid T_d=t,t\in M\right)$ and $P\left(M\mid T_d=t,t\notin M\right)$. In simple words, we need to know the probability of receiving every possible message given every possible departure time. Since this information is practically impossible to acquire, some additional assumptions must be made.

We assume that the probability of receiving a message of length *k* is l_k, and that all messages of a certain size that include the true departure time are equally likely. In other words, if for example the possible departure times are $T=\left\{t_1,t_2,t_3,t_4\right\}$ and the real departure time is t_1, then receiving $\left\{t_1,t_2\right\}$ is as likely as receiving $\left\{t_1,t_3\right\}$ or $\left\{t_1,t_4\right\}$.In general, since the number of messages of size *k* that include a certain departure time *s* is $\binom{n-1}{k-1}$, the conditional probability of receiving a certain message that includes the true departure time *t* is:

$$P\left(M\mid T_d=t,t\in M,\left|M\right|=k\right)=l_k\cdot\frac{1}{\binom{n-1}{k-1}} \tag{2.9}$$

Similarly, there are $\binom{n-1}{k}$ possible messages of size k that do not include a certain departure time t, and assuming they are all equally likely brings us to the following expression for the probability of receiving a certain message given that it does not include the true departure time t:

$$P\left(M \mid T_d = t, t \notin M, \left|M\right| = k\right) = l_k \cdot \frac{1}{\binom{n-1}{k}} \tag{2.10}$$

Combining Equations (2.9) and (2.10) we have the probability of receiving the message m of size k:

$$P\left(M \mid T_d = t, \left|M\right| = k\right) = \begin{cases} l_k \cdot q_k \cdot \dfrac{1}{\binom{n-1}{k-1}} & t \in M \\ l_k \cdot \left(1 - q_k\right) \cdot \dfrac{1}{\binom{n-1}{k}} & t \notin M \end{cases} \tag{2.11}$$

Thus, the denominator of the update function (2.7) becomes

$$\begin{aligned} P\left(M\right) &= \sum_{s \in T} f\left(s\right) \cdot l_k \cdot P\left(M \mid T_d = s, \left|M\right| = k\right) \\ &= \sum_{s \in M} f\left(s\right) \cdot l_k \cdot q_k \cdot \frac{1}{\binom{n-1}{k-1}} + \sum_{s \notin M} f\left(s\right) \cdot l_k \cdot \left(1 - q_k\right) \cdot \frac{1}{\binom{n-1}{k}} \end{aligned} \tag{2.12}$$

Combining the numerator (2.11) and the denominator (2.12) of the update function, we have the following function:

$$P\left(M \mid T_d = t, \left|M\right| = k\right) = \begin{cases} \dfrac{l_k \cdot q_k \cdot \dfrac{1}{\binom{n-1}{k-1}}}{\sum_{s \in M} f(s) \cdot l_k \cdot q_k \cdot \dfrac{1}{\binom{n-1}{k-1}} + \sum_{s \notin M} f(s) \cdot l_k \cdot (1-q_k) \cdot \dfrac{1}{\binom{n-1}{k}}} & t \in M \\ \dfrac{l_k \cdot (1-q_k) \cdot \dfrac{1}{\binom{n-1}{k}}}{\sum_{s \in M} f(s) \cdot l_k \cdot q_k \cdot \dfrac{1}{\binom{n-1}{k-1}} + \sum_{s \notin M} f(s) \cdot l_k \cdot (1-q_k) \cdot \dfrac{1}{\binom{n-1}{k}}} & t \notin M \end{cases} \tag{2.13}$$

Simplifying this expression yields the following update function:

$$f_{update}(t) = \begin{cases} \dfrac{\dfrac{q_k}{k}}{\sum_{s \in M} f(s) \cdot \dfrac{q_k}{k} + \sum_{s \notin M} f(s) \cdot \dfrac{1-q_k}{n-k}} & t \in M \\ \dfrac{\dfrac{1-q_k}{n-k}}{\sum_{s \in M} f(s) \cdot \dfrac{q_k}{k} + \sum_{s \notin M} f(s) \cdot \dfrac{1-q_k}{n-k}} & t \notin M \end{cases} \tag{2.14}$$

As an example, if the prior is a uniform prior function, $f_{prior}(t) = \dfrac{1}{n}$, the updated distribution can be simplified to be:

$$f_{new}(t) = \begin{cases} \dfrac{\dfrac{1}{n} \cdot \dfrac{q_k}{k}}{\dfrac{k}{n} \cdot \dfrac{q_k}{k} + \dfrac{n-k}{n} \cdot \dfrac{1-q_k}{n-k}} = \dfrac{q_k}{k} & t \in M \\ \dfrac{\dfrac{1}{n} \cdot \dfrac{1-q_k}{n-k}}{\sum_{s \in M} f(s) \cdot \dfrac{q_k}{k} + \sum_{s \notin M} f(s) \cdot \dfrac{1-q_k}{n-k}} = \dfrac{1-q_k}{n-k} & t \notin M \end{cases} \tag{2.15}$$

Thus, we have showed how to update the probability distribution after receiving an observation or a message.

4. Example

The following example demonstrates the process described above. Suppose there are only four possible departure times: $T = \{t_1, t_2, t_3, t_4\}$ of which the true departure time is t_1.

We receive a message of size $k = 2$, $M = \{t_1, t_2\}$, and q_2=0.9. Also, let us assume that the prior probability distribution of the departure times is uniform:

$$f_{prior}(t_1) = f_{prior}(t_2) = f_{prior}(t_3) = f_{prior}(t_4) = 0.25$$

There are $\begin{pmatrix} n-1 \\ k-1 \end{pmatrix} = 3$ possible messages of size 2 that include the true departure time, and the probability of receiving one of them is equal to 0.9. Similarly, there are $\begin{pmatrix} n-1 \\ k \end{pmatrix} = 3$, the probability of receiving one of them is equal to 0.1.

In this case, the updated function is:

$$\begin{aligned} f_{new}(t_1) = f_{new}(t_2) &= \frac{\frac{1}{3} \cdot 0.9}{2 \cdot \frac{1}{3} \cdot 0.9 + 2 \cdot \frac{1}{3} \cdot 0.1} = 0.45 \\ f_{new}(t_3) = f_{new}(t_4) &= \frac{\frac{1}{3} \cdot 0.1}{2 \cdot \frac{1}{3} \cdot 0.9 + 2 \cdot \frac{1}{3} \cdot 0.1} = 0.05 \end{aligned} \tag{2.16}$$

E. DEMPSTER-SHAFER BELIEF THEORY

1. Background

As discussed in the previous section, using the Bayesian method requires us to make significant assumptions that may be difficult to justify. In order to avoid those assumptions and allow for a more robust update mechanism, we explore the non-Bayesian Dempster-Shafer Theory (DST) methods.

The Dempster-Shafer theory defines sets of possible outcomes (or realizations) to a random variable similar to standard probability theory. However, unlike the definition of a probability distribution that assigns probabilities to exclusive outcomes, Dempster-Shafer theory is more general and assigns "mass" values not only to events but also to sets of events. This allows combining pieces of information in a more flexible way.

In the following paragraph, the basic Dempster-Shafer theory is defined mathematically, following Chapter 7.2.3 in (Hall & Llinas, 2001). Let *T* be a set of mutually exclusive outcomes of an experiment ("frame of discernment") and $\Omega = 2^T$ is the power set of *T*. The belief method assigns a "mass of evidence" *m* to elements in Ω.

The mass of evidence allocation, denoted by *m*, obeys the following rules:

$$m(\phi) = 0 \tag{2.17}$$

$$m(A) \geq 0, \forall A \in \Omega \tag{2.18}$$

$$\sum_{A \in \Omega} m(A) = 1 \tag{2.19}$$

Similar to probability, the mass of the empty set is 0 (Equation (2.17)), it is larger than 0 (Equation (2.18)) and the total mass sums to 1 (Equation (2.19)). Unlike probability, the mass can be defined to subsets of $\Omega = 2^T$ (Equation (2.18)). Intuitively it can be viewed as similar to probability theory, but when a mass is assigned to a set, the probability can still "shift" between the elements in the set when new information is acquired.

Now we define useful terms of Dempster-Shafer theory:

belief (*Bel):*

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{2.20}$$

The belief of *A* can be interpreted as the mass of evidence assigned to *A* and all its subsets. This is the minimal probability that is already assigned to *A*.

And *plausibility* (*Pl*):

$$Pl(A) = 1 - \sum_{A \cap B = \phi} m(B) = \sum_{A \cap B \neq \phi} m(B) \tag{2.21}$$

The plausibility is the mass of evidence that can possibly be assigned to *A* in the future (not assigned to any subsets that do not intersect with *A*).

The interval [*bel(A)*, *pl(A)*] can serve as a confidence interval for *A*'s probability (Hall & Llinas, 2001).

As an example, let's assume we have one observation from a source, stating: "With probability 90%, the vessel departed at t_1, and with 10% the vessel could have departed at any time.

In this case, the masses are distributed as follows:

$$\begin{aligned} m(\{t_1\}) &= 0.9 \\ m(T) &= 0.1 \end{aligned} \tag{2.22}$$

And so the belief and plausibility can be calculated to be:

	m	Bel	Pl
$\{t_1\}$	0.9	0.9	0.9
T	0.1	1	1
$\{t_2\}$	0	0	0.1

(2.23)

This is a very simple example, but we can already see that as expected, the "confidence interval" for t_1 is between 0.9 and 1 as expected. The confidence interval for t_2, about which no information was given, is also calculated to be $[0, 0.1]$.

On the other hand, if the source stated that “With probability 90%, I’m sure the vessel departed in t_1, and with 10% probability the vessel could have departed at any **other** time.

In this case, the masses are distributed as follows:

$$m\left(\{t_1\}\right) = 0.9$$
$$m\left(T - \{t_1\}\right) = 0.1 \tag{2.24}$$

And so the belief and plausibility can be calculated to be:

	m	Bel	Pl
$\{t_1\}$	0.9	0.9	0.9
$T-\{t_1\}$	0.1	0.1	0.1
$\{t_2\}$	0	0	0.1

(2.25)

which yields different results as now the confidence value for $T-\{t_1\}$ is $[0.1, 0.1]$. The mass given to this set is 0.1. A message can be defined in a similar manner, but with regard to multiple departure times: For an informant that says that the departure time is one of $\{t_1, t_2\}$, and we know that he is correct only 0.9 of the time, the corresponding mass assignment should be:

$$m\left(\{t_1, t_2\}\right) = 0.9$$
$$m\left(T - \{t_1, t_2\}\right) = 0.1 \tag{2.26}$$

And the belief and plausibility in this case are:

	m	Bel	Pl
$\{t_1, t_2\}$	0.9	0.9	0.9
$\{t_1\}$	0	0	0.9
$\{t_2\}$	0	0	0.9
$T-\{t_1, t_2\}$	0.1	0.1	0.1
$\{t_3\}$	0	0	0.1

(2.27)

2. Combination Rules

Now that we have defined the Dempster-Shafer framework, the next step is to discuss how masses assigned by different sources should be combined in order to make sense out of multiple sources of information. The desired outcome of such a combination of mass assignments would be a new mass assignment.

Probabilistically, if we have the probability assigned for each outcome by multiple sources and we assume independence, the combined probability distribution can be obtained by simply multiplying the probabilities assigned by different sources. In the Dempster-Shafer theory, the situation is more complicated and thus multiple combination rules are suggested with slightly different characteristics. The different combination rules are thoroughly discussed in Sandia Lab's report (Sentz, 2002). In the following paragraphs, we describe a few of the prominent rules in more detail with examples.

a. Dempster-Shafer

The first suggested rule is the Dempster-Shafer rule. Given two mass assignments by different sources m_1 and m_2, the combined mass assignment $m_{1,2}$ of a set *A* is calculated by adding up the multiplication of masses for all the sets *B*, *C* such that their intersection is *A*:

$$m_{1,2}(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - K}, A \neq \phi \tag{2.28}$$

where *K* is a normalization factor that accounts for the conflict - all the pairs of sets that have empty intersection and therefore their corresponding mass can not be assigned to any set:

$$K = \sum_{B \cap C = \phi} m_1(B) m_2(C) \tag{2.29}$$

As an example assume the first message specifies departure times $\{t_1, t_2\}$ with probability of 0.9 and the second message corresponds to times $\{t_1, t_3\}$, out of 5 possible departure times $S = \{t_1, t_2, t_3, t_4, t_5\}$:

$$
\begin{aligned}
&m_1\left(\{t_1, t_2\}\right) = 0.9, m_1\left(\{t_1, t_2, t_3\}\right) = 0.1 \\
&m_2\left(\{t_1, t_3\}\right) = 0.9, m_2\left(\{t_1, t_2, t_3\}\right) = 0.1
\end{aligned}
\tag{2.30}
$$

The combined masses can be calculated by determining the masses of the intersections:

	$m_1\left(\{t_1, t_2\}\right) = 0.9$	$m_1\left(\{t_3, t_4, t_5\}\right) = 0.1$
$m_2\left(\{t_1, t_3\}\right) = 0.9$	$m_{1,2}\left(\{t_1\}\right) = 0.81$	$m_{1,2}\left(\{t_3\}\right) = 0.09$
$m_1\left(\{t_2, t_4, t_5\}\right) = 0.1$	$m_{1,2}\left(\{t_2\}\right) = 0.09$	$m_{1,2}\left(\{t_4, t_5\}\right) = 0.01$

(2.31)

In this case, there is no conflict between the sources; each two sets with positive masses have a non-empty intersection (K=0).

However if informants submit conflicting messages, the situation will be different. Let us see what happens if in addition to the information before, the first informant is positive that the true departure time is not t_3, and therefore he does not assign any masses to sets that include t_3, resulting in the following mass assignments:

$$
\begin{aligned}
&m_1\left(\{t_1, t_2\}\right) = 0.9, m_1\left(\{t_4, t_5\}\right) = 0.1 \\
&m_2\left(\{t_1, t_3\}\right) = 0.9, m_2\left(\{t_2, t_4, t_5\}\right) = 0.1
\end{aligned}
\tag{2.32}
$$

	$m_1\left(\{t_1, t_2\}\right) = 0.9$	$m_1\left(\{t_4, t_5\}\right) = 0.1$
$m_2\left(\{t_1, t_3\}\right) = 0.9$	$m_{1,2}\left(\{t_1\}\right) = 0.81$	$K = 0.09$
$m_1\left(\{t_2, t_4, t_5\}\right) = 0.1$	$m_{1,2}\left(\{t_2\}\right) = 0.09$	$m_{1,2}\left(\{t_4, t_5\}\right) = 0.01$

(2.33)

Now there are two sets in the different assignments that have an empty intersection (the upper-right cell in the table). This mass is added to K, that measures the amount of “conflict.” The way Dempster-Shafer handles this situation is by normalizing the conflict out, eventually assigning:

$$
\begin{aligned}
m_{1,2}\left(\{t_1\}\right) &= \frac{0.81}{1-0.09} \approx 0.89 \\
m_{1,2}\left(\{t_2\}\right) &= \frac{0.09}{1-0.09} \approx 0.1 \\
m_{1,2}\left(\{t_{4,}t_5\}\right) &= \frac{0.01}{1-0.09} \approx 0.01
\end{aligned} \tag{2.34}
$$

Although this method may work well when the conflict is small, some paradoxes arise when the conflict between assignments is substantial. For example, if the first informant is almost certain that the departure time is t_1 but it might be t_3, and the second informant is quite sure about t_2, then:

$$
\begin{aligned}
m_1\left(\{t_1\}\right) &= 0.9, m_1\left(\{t_3\}\right) = 0.1 \\
m_2\left(\{t_2\}\right) &= 0.9, m_2\left(\{t_3\}\right) = 0.1
\end{aligned} \tag{2.35}
$$

	$m_1\left(\{t_1\}\right)=0.9$	$m_1\left(\{t_3\}\right)=0.1$
$m_2\left(\{t_2\}\right)=0.9$	$K=0.81$	$K=0.09$
$m_1\left(\{t_3\}\right)=0.1$	$K=0.09$	$m_{1,2}\left(\{t_3\}\right)=0.01$

(2.36)

Paradoxically, the combined assignment will be $m_{1,2}\left(\{t_3\}\right)=\frac{0.01}{1-0.99}=1$, disregarding the possibility of t_1 or t_2. Resolving this paradox is one of the main incentives in examining other combination rules.

This rule can be generalized to combine more than two messages:

$$
m_{1,2,\ldots,n}\left(A\right) = \frac{\sum_{\bigcap_{1\le i\le n} B_i = A} m_1\left(B_1\right) m_2\left(B_2\right) \cdots m_2\left(B_n\right)}{1-K}, A \neq \phi \tag{2.37}
$$

with K being:

$$K = \sum_{\substack{\bigcap B_i = \phi \\ 1 \le i \le n}} m_1\left(B_1\right) m_2\left(B_2\right) \cdots m_2\left(B_n\right) \tag{2.38}$$

b. *Yager's Modified Dempster-Shafer Rule*

This method is similar to the basic Dempster-Shafer rule, with one difference: instead of normalizing the masses m by 1-K, K is added to the mass of the entire set T. The mass of the entire set T can be interpreted as the "ignorance," since this mass does not help in distinguishing between different departure times.

We now revisit the high conflict example defined in (2.35) and (2.36). Following the same calculation, the final assignment would be $m_{1,2}\left(\left\{t_3\right\}\right) = 0.01, m_{1,2}\left(T\right) = 0.99$ with confidence intervals:

	m	Bel	Pl
$\left\{t_3\right\}$	0.01	0.01	1
$\left\{T\right\}$	0.99	0.99	0.99
$\left\{t_1\right\},\left\{t_2\right\}$	0	0	0.99

(2.39)

Suggesting that since the conflict is so large, every outcome is basically possible. This rule can also be extended for more than two evidences, but it is not associative. It is, however, commutative as Equation (2.37) is symmetric.

c. *Zhang's Center Combination Rule*

This rule is yet another extension of the Dempster-Shafer combination rule. While the Dempster-Shafer rule does not account for the intersection between two sets, Zhang's rule does by multiplying the assigned outcome mass by a metric $r\left(B,C\right)$, as logically the mass assigned to the intersection of B and C should increase with its size. A common metric is the cardinality of the intersection:

$$r\left(B,C\right) = \frac{\left|B \cap C\right|}{\left|B\right| \cdot \left|C\right|} \tag{2.40}$$

and the combined mass is:

$$m_{1,2}\left(A\right)=k\sum_{B\cap C=A} r\left(B,C\right)m_1\left(B\right)m_2\left(C\right) \tag{2.41}$$

with k a normalization factor such that $\sum_{A\subseteq S} m_{1,2}\left(A\right)=1$, (not the same as K that accounted for the conflict in Dempster-Shafer and Yager's rules).

To see the difference between Zhang's rule and the regular Dempster-Shafer combination rule, let's look at the case where we have a positive observation regarding time t_1 and message about times t_1 or t_2.

$$\begin{aligned} &m_1\left(\left\{t_1\right\}\right)=0.9, m_1\left(\left\{t_2,t_3,t_4,t_5\right\}\right)=0.1 \\ &m_2\left(\left\{t_1,t_2\right\}\right)=0.9, m_2\left(\left\{t_3,t_4,t_5\right\}\right)=0.1 \end{aligned} \tag{2.42}$$

As before, we calculate the masses of the intersections:

	$m_1\left(\left\{t_1\right\}\right)=0.9$	$m_1\left(\left\{t_2,t_3,t_4,t_5\right\}\right)=0.1$
$m_2\left(\left\{t_1,t_2\right\}\right)=0.9$	$m_{1,2}\left(\left\{t_1\right\}\right)=0.81$	$m_{1,2}\left(\left\{t_2\right\}\right)=0.09$
$m_2\left(\left\{t_3,t_4,t_5\right\}\right)=0.1$	$K=0.09$	$m_{1,2}\left(\left\{t_3,t_4,t_5\right\}\right)=0.01$

(2.43)

So the intervals for the departure times are:

	m	Bel	Pl
$\left\{t_1\right\}$	0.89	0.89	0.89
$\left\{t_2\right\}$	0.1	0.99	0.99
$\left\{t_3,t_4,t_5\right\}$	0.01	0.01	0.01

(2.44)

However, with Zhang's rule we also calculate the value of $r\left(B,C\right)=\frac{\left|B\cap C\right|}{\left|B\right|\cdot\left|C\right|}$ for each intersection:

$$
\begin{array}{c|cc}
 & m_1\left(\left\{t_1\right\}\right)=0.9 & m_1\left(\left\{t_2,t_3,t_4,t_5\right\}\right)=0.1 \\
\hline
m_2\left(\left\{t_1,t_2\right\}\right)=0.9 & m_{1,2}\left(\left\{t_1\right\}\right)=0.81, r=\frac{1}{2} & m_{1,2}\left(\left\{t_2\right\}\right)=0.09, r=\frac{1}{8} \\
m_2\left(\left\{t_3,t_4,t_5\right\}\right)=0.1 & - & m_{1,2}\left(\left\{t_3,t_4,t_5\right\}\right)=0.01, r=\frac{1}{4}
\end{array}
\tag{2.45}
$$

And by using Equation (2.41), the final combined masses according to Zhang's rule are:

	m	Bel	Pl
$\{t_1\}$	0.967	0.967	0.967
$\{t_2\}$	0.027	0.027	0.027
$\{t_3,t_4,t_5\}$	0.006	0.006	0.006

(2.46)

Since the size of the intersection between $m_1\left(\left\{t_1\right\}\right)$ and $m_2\left(\left\{t_1,t_2\right\}\right)$ is much bigger than the others (which can be interpreted as a better agreement), $m_{1,2}\left(\left\{t_1\right\}\right)$ receives a higher mass in Zhang's rule in comparison with Dempster-Shafer's combination rule.

A conflict between assignments is resolved by the normalization coefficient *k* similar to Demspter-Shafer's combination rule.

d. *Mass Mean*

This rule of combination is the most straightforward one. According to the mass average rule, the mass of a set in the combined mass assignment is simply the average of masses the set received:

$$
m_{1,2,..,n}\left(A\right)=k\sum_{1\le i\le n} m_i\left(A\right) \tag{2.47}
$$

Weighting the average according to some measure of confidence or reliability is also common, but this issue is not discussed in this thesis.

3. Observations and Messages

Defining observations and messages according to Dempster-Shafer scheme is not straightforward. We define these terms as follows:

We translate a positive observation with "correctness" q regarding time t_i into a mass assignment of:

$$m\left(\left\{t_i\right\}\right)=q, m\left(T-\left\{t_i\right\}\right)=1-q \tag{2.48}$$

This means that according to this observation, with probability q the true departure time is t_i and with probability $1-q$ it is any other time.

Similarly, a message M with "correctness" q_k ($\left|M\right|=k$) can be translated to the following mass assignment:

$$m\left(M\right)=q_k, m\left(T-M\right)=1-q_k \tag{2.49}$$

In this case the true departure time is included in the message M with probability q_k , and with probability $1-q_k$ it is not in the message.

4. Transforming Dempster-Shafer Belief and Plausibility Measures to Probability Values

In order to support decision makers we must have some well-defined probability about the departure time of the target. This is straightforward with the Bayesian method, but the Dempster-Shafer theory only allows us to obtain *Belief-Plausibility* intervals. The following subsection describes a few methods of translating the belief function (or mass function) to a probability value.

a. Pignistic Transformation

The most popular transformation is the Pignistic Transformation (Smets, 1990) that distributes the mass assigned to a set uniformly among all the members of the set:

$$P_{pig}(t) = \sum \left\{ \frac{m(A)}{|A|} \mid A \in \Omega\, s.t.\, t \in A \right\} \tag{2.50}$$

where $P_{pig}(t)$ is the estimated probability of departure time *t* by this transformation. This transformation is intuitive, when we are given the mass of a subset and we are required to estimate the probability of each member of the set, a natural assumption is that the probability of all the members is equal.

Let's examine the outcome of this transformation in the case of a single message:

$$\begin{aligned} m(\{t_1, t_2\}) &= 0.9 \\ m(\{t_3, t_4, t_5\}) &= 0.1 \end{aligned} \tag{2.51}$$

And the Belief and Plausibility in this case:

	m	Bel	Pl
$\{t_1, t_2\}$	0.9	0.9	0.9
$\{t_1\}$	0	0	0.9
$\{t_2\}$	0	0	0.9
$\{t_3, t_4, t_5\}$	0.1	0.1	0.1
$\{t_3\}$	0	0	0.1

(2.52)

Then the pignistic probability of each time is:

$$P_{pig}(t_1) = P_{pig}(t_2) = \frac{0.9}{2} = 0.45$$

$$P_{pig}(t_3) = \frac{0.1}{3} = P_{pig}(t_4) = P_{pig}(t_5) \approx 0.033$$

Smets (2002) claims that the pignistic is the transformation adequate for making decisions; however, it does not necessarily represent your belief:

> At the creedal level, beliefs are represented by a belief function; at the pignistic level, this belief function induces a probability function that is used to make decisions. This probability function should not be understood as representing your beliefs, it is nothing but the additive measure needed to make decision, i.e., to compute the expected utilities.

However, one of the principles of Dempster-Shafer theory is that with further information the mass can shift within the set. This transformation, however intuitive, disregards this flexibility by dividing the mass equally among the members of the set (Cobb & Shenoy, 2003).

b. Plausibility Transformation

Another transformation is the Plausibility transformation, which is basically a normalization of the plausibility function of the singletons.

$$P_{pl}(t) = K^{-1} Pl(\{t\}) \tag{2.53}$$

with K being the normalization factor:

$$K = \sum Pl(\{t\} \mid t \in \Omega) \tag{2.54}$$

This transformation tries to keep the essence of Dempster-Shafer theory by considering the plausibility value of the departure times (that can be interpreted as the potential, or the biggest mass that can possibly be assigned to it if the right information is received) and normalizing those plausibilities. Looking at the same example as in the previous subsection, defined by Equation (2.51), we obtain:

$$\begin{aligned} K &= \sum_{1 \le i \le n} Pl(t_i) = 0.9 + 0.9 + 0.1 + 0.1 + 0.1 = 2.1 \\ P_{pl}(t_1) &= P_{pl}(t_2) = \frac{0.9}{2.1} \approx 0.42 \\ P_{pl}(t_3) &= P_{pl}(t_4) = P_{pl}(t_5) = \frac{0.1}{2.1} \approx 0.047 \end{aligned} \tag{2.55}$$

The probability assigned to t_1 and t_2 is lower now, to account for the fact that there are three other departure times that are plausible.

c. Belief Transformation

A less useful transformation is the Belief transformation that normalizes the beliefs of the singletons but disregards information about non-singletons. This transformation will not be considered in this work (and in the example discussed it does not make any sense).

F. DISCUSSION

We have suggested two different probability update models: The Bayesian update and Dempster-Shafer Theory. Both models allow updating observation and messages. The Bayesian method is well known, popular and mathematically rigorous. However, it requires multiple assumptions that are often difficult to justify.

The Demspter-Shafer Theory includes multiple combination rules. There is no clear method that is considered appropriate for all situations According to our mass assignments as defined in (2.24) and (2.26) we do not expect to have any conflict since all the possible departure times are members of sets that have a positive mass, and therefore, Dempster-Shafer and Zhang's rules are reliable and justifiable (Sentz, 2002). Yager's rule is useful whenever there is large conflict; however in our case, since there is no conflict, it is no different than Dempster-Shafer's rule. The mean combination rule may not be appropriate when averaging extremes, but it is easy to compute and might provide satisfactory results in certain cases, and therefore, is also of interest.

The transformation step from Dempster-Shafer theory measures to probabilities is also possible via several distinct methods. Although the pignistic method is more common in the literature the Plausibility method has some appealing characteristics, and it is more consistent with Dempster-Shafer theory (Cobb & Shenoy, 2003).

In the following chapter we will compare Dempster-Shafer, Zhang's and the mean methods, each of those transformed into probabilities by the two transformations models and the Bayesian approach.

III. BAYES AND DEMPSTER SHEFER THEORY COMPARISON

In this chapter, we compare the Bayesian update process and Dempster-Shafer methods. We first discuss the qualitative pros and cons of each method and show the equivalence between the two in certain situations. We conclude with a section describing the results from a simulation experiment.

A. QUALITATIVE COMPARISON

As discussed in the previous chapter, the Bayes' method requires a prior probability distribution $f_{prior}(t)$. Dempster-Shafer theory does not require a prior for computing the update distribution although it can incorporate such a prior as an additional piece of information. The biggest advantage of Dempster-Shafer theory in our context is that it does not require one to specify the probabilities of receiving a message given the actual departure time: $P(M \mid T_d = t, t \in M)$ and $P(M \mid T_d = t, t \notin M)$. Data to estimate these conditional probabilities may not be available and so assumptions that are difficult to justify have to be made to obtain them. However, once we impose these assumptions, calculating the updated probabilities is straightforward using Bayes' theorem, and the result of the update is unique. Dempster-Shafer theory may utilize any one of multiple combination rules, and those rules may produce very different results depending on the agreement between the different pieces of information received.

The output of the Bayesian update gives us the estimated probability that a given departure time is in fact the true one. The Dempster-Shafer theory output is a distribution of masses that allows us to calculate Belief-Plausibility confidence intervals regarding the time departure. Those intervals give some insight about the probability of the vessel departing at a certain time but in order to make decisions, and in particular when those intervals are large, an additional transformation is required to obtain probabilities. This transformation results in a degradation of the flexibility that makes Dempster-Shafer theory so appealing.

Computation-wise, the Dempster-Shafer theory machinery is much more intensive since each member in the power set of *T* may be assigned a mass. A Bayesian update assigns probabilities only to the members of the set itself, not its power set. However, the Bayes' method may incur additional computational costs if calculating $P\left(M \mid T_d = t, t \in M\right)$ and $P\left(M \mid T_d = t, t \notin M\right)$ from existing data is difficult. The following table summarizes the main differences between the Bayes' updating process and the Dempster-Shafer theory:

	Bayes' Method	**Dempster-Shafer Theory**
Prior	A prior distribution of the outcomes is required	Prior distribution not required
Event distribution	The probability of receiving each message in every state of the world is required	Only mass distribution is required
Combination Rules	Bayes' Formula	Several different combination rules
Output	Probability distribution of outcomes	Results in Belief and Plausibility of outcomes and sets of outcomes. Requires transformation to obtain probabilities
Computation	Computationally easy, assigns probability values to the members of *T*.	Computationally intensive, requires assigning values to members of the power-set of *T*.

Table 1. Bayesian update - DST comparison.

B. BAYESIAN UPDATE – DEMPSTER SHAFER ZHANG EQUIVALENCE

As we saw in Chapter II, the update process can be performed using multiple methods. In this section we show that the Bayesian update and the Dempster-Shafer Zhang method with a pignistic transformation produce the same probabilities under the model assumptions described in Chapter II. As stated, for the Bayesian model we assume that (a) q_k is known, (b) the probabilities of receiving true messages of a certain size are equal and the probabilities of receiving false messages of the same size are also equal, and (c) the Bayes' prior is a uniform distribution. The Zhang combination method is performed as described in Chapter II, Equation (2.41). Assuming that (a) q_k is known and (b) receiving message of size *k*, M_k yields the following mass assignment:

$$\begin{aligned} m\left(M_k\right) &= q_k \\ m\left(T - M_k\right) &= 1 - q_k \end{aligned} \tag{3.1}$$

In order to show the Bayes - Zhang pignistic equivalence we calculate the probabilities assigned to a certain departure time t by both methods after N messages are received. We assume that the informant included the specific departure time t in exactly N_{in} messages out of the N received, and $N_{out} = N - N_{in}$.

1. Bayesian Update

Recall from Equation 2.14 the update equation when the message size is k :

$$f_{update}\left(t\right) = \begin{cases} \dfrac{\dfrac{q_k}{k}}{\sum\limits_{s\in M} f\left(s\right)\cdot\dfrac{q_k}{k} + \sum\limits_{s\notin M} f\left(s\right)\cdot\dfrac{1-q_k}{n-k}} & t \in M \\ \dfrac{\dfrac{1-q_k}{n-k}}{\sum\limits_{s\in M} f\left(s\right)\cdot\dfrac{q_k}{k} + \sum\limits_{s\notin M} f\left(s\right)\cdot\dfrac{1-q_k}{n-k}} & t \notin M \end{cases} \tag{3.2}$$

where $f\left(s\right)$ is the probability distribution before the update. Since the denominator of the update function is merely a normalization coefficient, we can state that without normalizing, the updated distribution is:

$$f_{update}\left(t\right) \propto \begin{cases} \dfrac{q_k}{k} & t \in M \\ \dfrac{1-q_k}{n-k} & t \notin M \end{cases} \tag{3.3}$$

Applying Bayes' theorem and using the fact that the message probability depends only on the true departure time and q_k, and that the messages are independent given the departure time, allows us to formulate the update function after N messages:

$$P_{new}\left(T_d = t \mid M_k^1, M_k^2, \ldots, M_k^N\right) = \frac{P\left(M_k^1, M_k^2, \ldots, M_k^N \mid T_d = t\right) \cdot P_{prior}\left(T_d = t\right)}{P\left(M_k^1, M_k^2, \ldots, M_k^N\right)} =$$
$$= \frac{P\left(M_k^1 \mid T_d = t\right) \cdot P\left(M_k^2 \mid T_d = t\right) \cdots P\left(M_k^N \mid T_d = t\right) \cdot P_{prior}\left(T_d = t\right)}{P\left(M_k^1, M_k^2, \ldots, M_k^N\right)} \tag{3.4}$$

Since $P\left(M_k \mid T_d = t\right) \propto \frac{q_k}{k}$ for messages that include *t* and $P\left(M_k \mid T_d = t\right) \propto \frac{1-q_k}{n-k}$ for messages that do not include *t*, the update function after N_{in} messages that include t and N_{out} that do not include *t* is:

$$f_{update}\left(t\right) \propto \left(\frac{q_k}{k}\right)^{N_{in}} \cdot \left(\frac{1-q_k}{n-k}\right)^{N_{out}} \tag{3.5}$$

This update function can later be normalized, although it is not required for proving the equivalence to the Pignistic Zhang method.

2. Pignistic Zhang formulation

Recall the Zhang's combination rule from Equations (2.40) and (2.41):

$$m_{1,2}\left(C\right) = k \sum_{A \cap B = C} m_1\left(A\right) m_2\left(B\right) \frac{\left|A \cap B\right|}{\left|A\right|\left|B\right|} \tag{3.6}$$

The specific departure time *t* is included in M_k in N_{in} of the messages, and included in T-M_k in N_{out} of the messages. In order to eventually calculate the probability of departure at time *t* let us first look at the intersection of the sets that include *t* for all the messages, C:

$$C = M_k^1 \cap M_k^2 \cap \ldots \cap M_k^{N_{in}} \cap \left\{T - M_k^{N_{in}+1}\right\} \cap \ldots \cap \left\{T - M_k^N\right\} \tag{3.7}$$

Since for every message M_k, t is included either in M_k or in $\{T-M_k\}$, the sum in Equation (3.6) is reduced to a single term and C can be interpreted as the set that includes t after the combination of all the messages received. The mass of C is:

$$m_{combined}(C) \propto \prod_{i:t\in M_k} m_i\left(M_k^i\right)\cdot \prod_{j:t\notin M_k^j} m_i\left(\left\{T-M_k^j\right\}\right)\cdot \frac{|C|}{\prod_{i:t\in M_k}\left|M_k^i\right|\cdot \prod_{j:t\notin M_k^j}\left|\left\{T-M_k^j\right\}\right|} \tag{3.8}$$

We know that the mass assignment is $m(M_k)=q_k$, $m(T-M_k)=1-q_k$ and that the size of a sets are $|M_k|=k$ and $|\{T-M_k\}|=n-k$. By substituting those expressions in Equation (3.8) we have:

$$m_{combined}(C) \propto \frac{q^{N_{in}}\cdot(1-q)^{N_{out}}}{k^{N_{in}}\cdot(n-k)^{N_{out}}}\cdot|C| \tag{3.9}$$

Once we have the combined mass assignment, we can calculate the pignistic probability of time t according to Equation (2.50):

$$P_{pig}(t)=\sum\left\{\frac{m(C)}{|C|}\,|\,C\in\Omega\, s.t.\, t\in C\right\} \propto \frac{q^{N_{in}}\cdot(1-q)^{N_{out}}}{k^{N_{in}}\cdot(n-k)^{N_{out}}}\cdot\frac{|C|}{|C|}=\frac{q^{N_{in}}\cdot(1-q)^{N_{out}}}{k^{N_{in}}\cdot(n-k)^{N_{out}}} \tag{3.10}$$

which equates to the expression achieved via the Bayesian update in Equation (3.5).

C. SIMULATION

We construct a simulation experiment, built on the simulation described in (Martin, 2009) and implemented in MATLAB, to further compare the Bayesian and Dempster-Shafer methods. Using the simulation we study the process of combining pieces of evidence. The simulation mimics the production of different messages and its main output is the distribution specifying the probability the target left at any particulate departure time (probability distribution over *T*). We construct the simulation in two parts: 1) generating the stream of messages that represents the state of the world and it does not depend on the update mechanism, and 2) updating the departure time distribution using different methods described thus far. For a fixed number of messages received, we run the simulation multiple times and calculate the fraction of time the correct departure time has the highest probability after performing the updates with a certain method.

1. Generating Messages

First, the simulation generates a stream of messages of size k out of n possible departure times, assuming the informant's reliability is q_k. That is, on average, a fraction q_k of the messages include the true departure time. Next, we examine the update methods on different streams of messages. We examine streams of messages that are generated both according to the assumption that all true messages are equally likely (see Chapter II), and when this assumption is relaxed. We define an input parameter $NU \in [0,1)$ to describe the measure of non-uniformity – $NU = 0$ implies that the messages are created uniformly, and as NU increases to 1, the non-uniformity of messages also increases. The exact effect of this parameter is described in the following paragraph.

We assume, without loss of generality that the true departure time T_d is the last one possible: $T_d = n$.

The method for generating the messages of size *k* proceeds as follows:

- First, we determine whether the message includes the true departure time. We do this by generating a random Bernoulli variable with parameter q_k.
- If the message is true, we include the true departure time in it. If it is false, we make sure that it does not include the true departure time. Next, we populate the rest of the message with departure times:
- If $NU = 0$, the rest of the departure times are picked uniformly, that is, each of the possible times have the same probability of being included in the message.
- If *NU>0*, a random departure time t is drawn from a geometric distribution with a parameter NU. If $t \in [1,..,n]$, t is not the true departure time and t is

not included in the message already, t is added to message. Else, t is drawn again. This process repeats itself until the message is filled up with k departure times. As an example, let us look at the probability of picking different departure times for some values of NU when $n = 9$, as depicted in Figure 6:

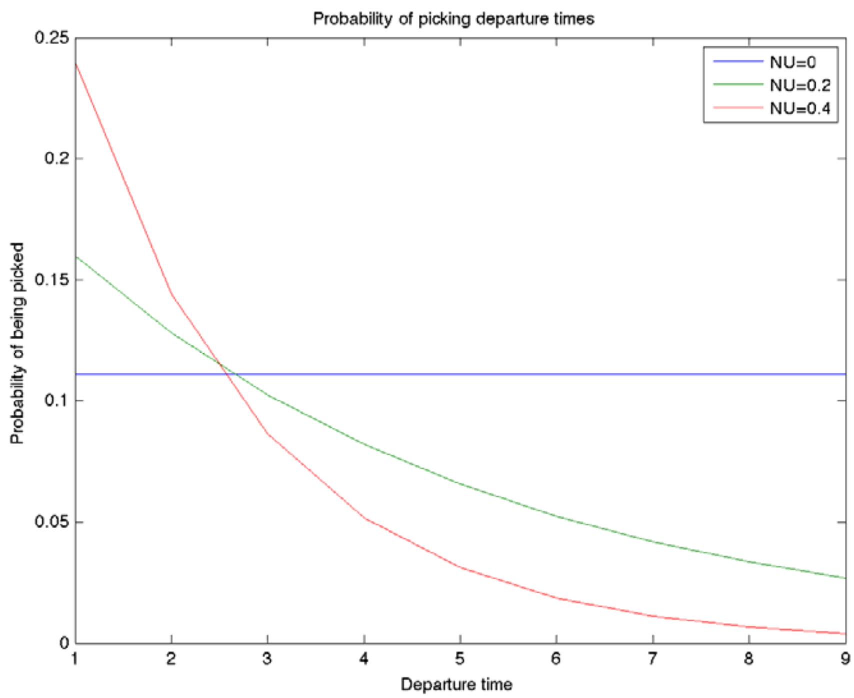

Figure 6. Probability of populating the message with departure times for different NU values.

The bigger *NU*, the more likely smaller values will populate the message. In other words, messages with small values of *t* will be more likely to be generated. Once the messages are created, the different update methods are used in order to estimate the true departure time.

2. Estimating the Probabilities of the Departure Times

We use six different methods to generate the combined probabilities:

1. Bayesian method, with a uniform prior and an updating process that assumes that all true message of a certain size are equally likely, and the same is true for false messages. The update process uses this assumption, but the messages generated might not be drawn according to this assumption if *NU>0*.
2. Dampster-Shafer rule, where the combined probability is derived from the pignistic method.
3. Dampster-Shafer rule, where the resulting probability is derived from the plausibility method.
4. Zhang's rule, where the resulting probability is derived from the plausibility method.
5. Mean rule with a pignistic transformation.
6. Mean rule with a plausibility transformation.

Each of these methods is applied according to the description in Chapter II, assuming that there is an estimate of the informant's reliability q_e. Note that this parameter does not have to equate to the true informant's reliability q_k, since the estimation of the informant's reliability might not be correct. Different values of q_e and q_k will be tested in the simulation

Since the messages are not necessarily created uniformly the update processes might use the wrong distributional assumptions. However, we are interested in examining how well Bayes' update method performs even when it uses the wrong distributions for its updating process in comparison with Dempster-Shafer methods. While the Dempster-Shafer methods have less explicit assumptions than the Bayesian update approach, the combination rules are somewhat arbitrary and may have hidden assumptions that are manifested in the different combination rule.

3. Constructing the Results

We focus on two measures of effectiveness (MOE) in our analysis: (1) the average probability assigned to the true departure time, and (2) the percent of the runs in which the true departure time has the highest probability. After each simulation run we record the probability specified for the true departure time for each of the updating methods. We also tabulate for each method whether the true departure time has the highest final probability associated with it. We do this because if the decision maker needs to take action, he would select the time with the highest probability. After many runs of the simulation we can calculate the average probability assigned to the true departure time as a function of the number of messages received and the percent of the runs in which the true departure time has the highest probability. We choose the latter as our main MOE. We also calculate the standard deviation of this MOE across the runs conducted.

4. Input Parameters

The input parameters of the simulation are as follows:

1. Number of possible departure times (cardinality of T) - n
2. Size of each message - k
3. The true departure time – t_d (without loss of generality it is fixed to be the latest possible departure time - n).
4. The true value of the probability that a message contains the true departure time – q_k. This parameter controls the message generation process and is not known to the operator. q_k can be interpreted as the "true reliability" of the informant.
5. The estimated probability that a message is true – q_e. q_e can be interpreted as the estimated reliability of the informant by the operator. The updating process requires an estimate of this probability, and we assume that the operator knows this estimation ahead of time. (It is an input to the simulation) We assume this value is constant but does not have to be equal to q_k – as happens when the operator does not estimate correctly the reliability of the source. This allows us to generate a stream of messages that differs from the Bayes' and Dempster-Shafer theory update assumptions.

6. Non-uniformity parameter NU – controls the non-uniformity behavior of the messages produced between 0 and 1.
7. Number of messages constructed in the simulation.
8. Number of runs for each set of parameters.

5. Design of Experiments

We have conducted multiple runs of different scenarios where the number of possible departure times is $n = 9\left(=\left|T\right|\right)$ and the size of the message is $k = 3$. The departure time is fixed to be $T_d = n = 9$.

We construct a full design with the parameter values stated in Table 2:

Parameter	Values
q_k	0.4,0.7,0.95
q_e	0.4,0.7,0.95
NU	0,0.2,0.4,1

Table 2. Parameter values.

Note that if the informant is clueless and thus chooses the departure times in the message totally randomly (i.e., the informant provides no useful information), q_k would equal to $\frac{k}{n} = \frac{1}{3}$. All the q_k used in the simulation (see Table 2) imply a “useful” informant that provides true messages with probability higher than that generated from a uniform distribution.

For each set of parameter values, a stream of 30 messages is created 100 times.

6. Simulation Results

For most of the input parameter values, the different methods produce similar results. Figures 7 and 8 show the results obtained for one such scenario, where:

$q_k = q_e = 0.7, NU = 0$. Figure 7 depicts the percent of the runs in which the true departure time has the highest probability, while Figure 8 shows the average probability assigned to the true departure time.

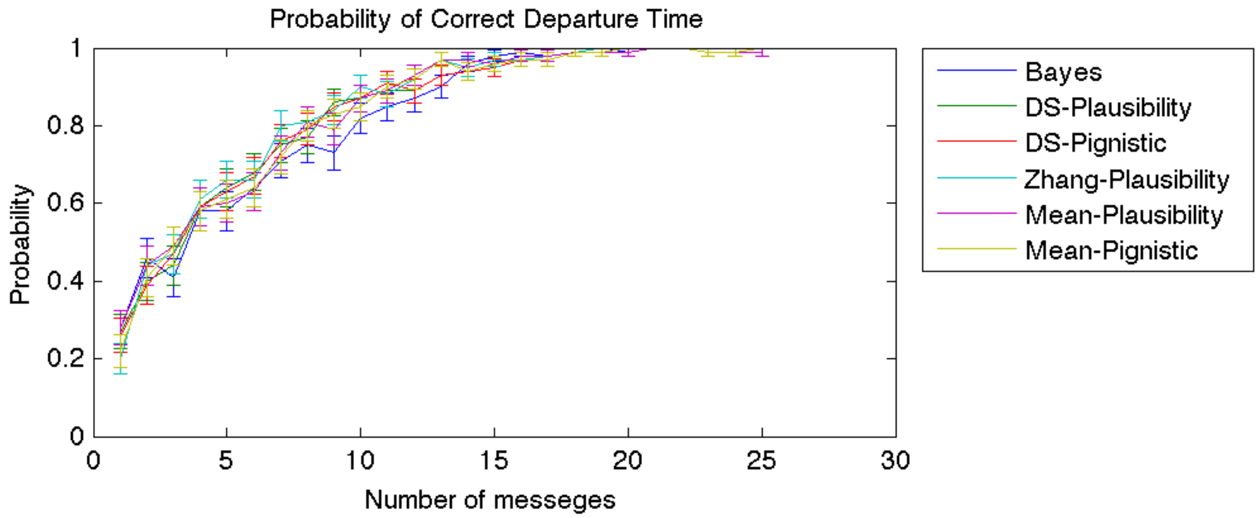

Figure 7. Probability of picking the correct departure time. 100 runs with parameters $q_k = q_e = 0.7, NU = 0$.

As one would expect, the probability of choosing the correct departure time increases with the number of messages received, for all methods. This occurs because the operator gains more useful information regarding the departure time. As the number of received messages increases, the probability that an incorrect departure time is included in more messages than the correct departure time decreases.

However, the probability assigned by the different methods to the correct departure time can differ significantly, as Figure 8 shows.

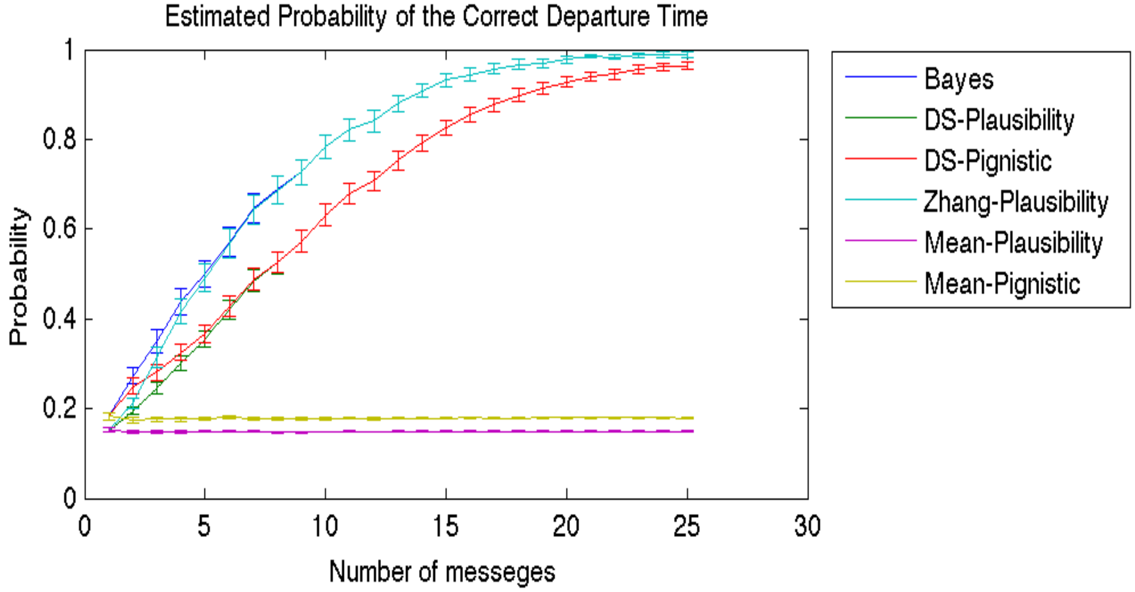

Figure 8. Probability assigned to the correct departure time. 100 runs with parameters $q_k = q_e = 0.7, NU = 0$.

Based on the results in Figure 8 we can divide the six methods into three groups. The probability assigned to the correct departure time by the Bayesian and Zhang methods increases at the fastest rate as a function of number of messages in Figure 8. The probability the two Dempster-Shafer combination rules (plausibility and pignistic) assign to the correct departure times increases at a slower rate with the number of messages, compared to the Bayesian update and the Zhang combination rule. The reason for that is that the two Dempster-Shafer combination rules do not take into account the sizes of the sets combined when assigning the mass of the intersection of those sets, as described in Chapter II. Interestingly, the mean combination methods (plausibility and pignistic) do not change the prior probability assigned to the correct departure time. Let us look deeper into this phenomenon by examining the pignistic-mean method.

A single message is true with probability q_k and false with probability $1-q_k$. The expected probability assigned to it, according to the pignistic transformation defined in (2.50), is $q_k \cdot \frac{q_k}{k} + (1-q_k) \cdot \frac{1-q_k}{n-k} \approx 0.18$ for the parameters' values $q_k = q_e = 0.7$. Now let us look at the situation after two messages, M_k^1 and M_k^2. The mass assignment that corresponds to those messages is: $m_1\left(M_k^1\right)=q_k, m_1\left(T-\left\{M_k^1\right\}\right)=1-q_k$ and $m_2\left(M_k^2\right)=q_k, m_2\left(T-\left\{M_k^2\right\}\right)=1-q_k$. According to the mean rule (2.47), the combined assignment would be:

$$m_{1,2}\left(M_k^1\right)=\frac{q_k}{2}, m_{1,2}\left(T-\left\{M_k^1\right\}\right)=\frac{1-q_k}{2}, m_{1,2}\left(M_k^2\right)=\frac{q_k}{2}, m_{1,2}\left(T-\left\{M_k^2\right\}\right)=\frac{1-q_k}{2}.$$

Now let us calculate the probability assigned to the correct departure time according to the pignistic transformation. Let us define $0 \le N_{in} \le 2$ as the number of messages t is included in, as in section A. For every possible value of N_{in} we calculate the probability of this $N_{in}=n_{in}$ and the probability assigned to time *t* given it is included in n_{in} messages. The explicit calculations are shown in the following table:

n_{in}	$P\left(N_{in}=n_{in}\right)$	$m_{1,2}\left(t \mid N_{in}=n_{in}\right)$
0	$\left(1-q_k\right)^2=0.09$	$\frac{1}{2}\cdot\frac{1-q_k}{n-k}+\frac{1}{2}\cdot\frac{1-q_k}{n-k}=\frac{1-q_k}{n-k}=0.05$
1	$2q_k\left(1-q_k\right)=0.42$	$\frac{1}{2}\cdot\frac{q_k}{k}+\frac{1}{2}\cdot\frac{1-q_k}{n-k}\approx 0.16$
2	$q_k^2=0.49$	$\frac{1}{2}\frac{q_k}{k}+\frac{1}{2}\cdot\frac{q_k}{k}=\frac{q_k}{k}\approx 0.23$

Table 3. Probability of *having* n_{in} messages that include *t* and the mass assigned to *t.*

From those three possibilities we can calculate the expected probability assigned to the correct time would be $E\left[m_{1,2}\left(t\right)\right]=\sum_{n_{in}=0,1,2} P\left(N_{in}=n_{in}\right)\cdot m_{1,2}\left(t \mid N_{in}=n_{in}\right)=0.18$. Doing the same calculation with an increasing number of messages yields similar results.

Although the probability of the correct departure time derived by the two mean methods (pignistic and plausibility) is much lower than in the other methods, the probabilities assigned to the incorrect times are $\frac{1-0.18}{n-1}\approx 0.10$ (when they are uniform).

Thus, as Figure 7 illustrates, using the mean Dempster-Shafer methods yield the correct departure time in most cases. We conclude that although the mean methods assign incorrect probabilities to the departure times, they still point to the departure time with the highest probability as accurately as the other methods.

For most of the input combination of Table 3, the methods produced nearly identical results. However, there are two sets of input values that yielded non-trivial differences between the methods: 1) When true reliability value is low—$q_k = q_e = 0.4, NU \geq 0$. For these input values the Bayesian, mean-pignistic, and Zhang-Plausibility methods had the highest probability of picking the correct departure time, and 2) when the estimated reliability is lower than the true one $q_k = 0.7, 0.95, q_e = 0.4, NU = 0$. For these input values, the combination methods differ, with Bayes' and mean-pignistic methods performing the best, followed by the Zhang-Plausibility.

The input values that yield the greatest difficulties for the update process are those with low informant reliability and some non-uniformity of the generated messages, or those cases where the estimated reliability is less than the true reliability. Unexpectedly, Bayesian update proves to be robust and performs near the top over all scenarios, even when Bayes' assumptions do not hold. Mean-pignistic performed nearly as well as Bayes, with Zhang-plausibility performing slightly worse. The Dempster-Shafer combination rule and mean plausibility rule do not perform as well.

As an example for one of cases where the methods differ, let us look closely at the results for input values of $q_k = q_e = 0.4, NU = 0.4$, in Figure 9:

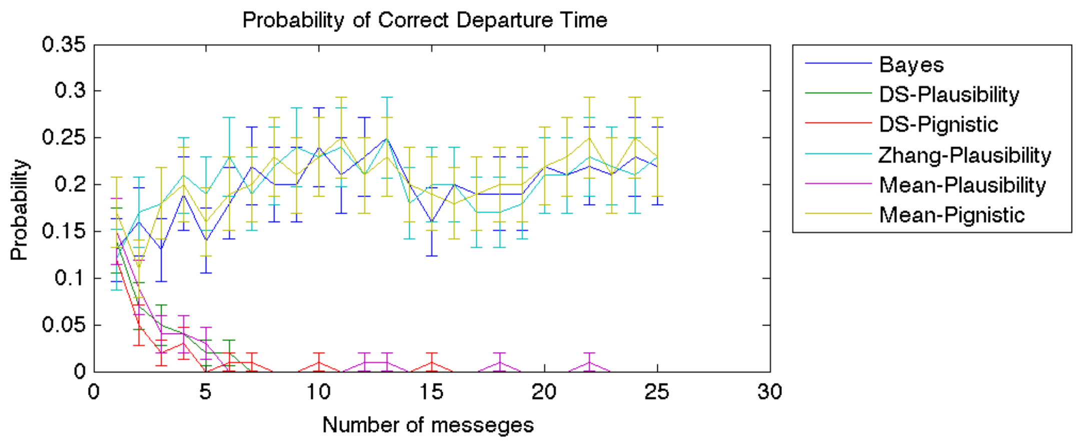

Figure 9. Probability of picking the correct departure time. 100 runs with parameters $q_k = q_e = 0.4, NU = 0.4$.

Bayes, Zhang and mean-pignistic methods clearly outperform the other methods. The probability of choosing the right departure time is increasing slowly for these cases, but is actually decreasing for the other methods. The intuition behind this phenomenon is the disregarding of the set sizes (as discussed in this section) and the bias caused by messages generated non-uniformly towards wrong departure times that mislead the Dempster-Shafer combination rules. Figure 10 emphasizes this point by showing how the probability of the true departure time changes as the number of messages increases:

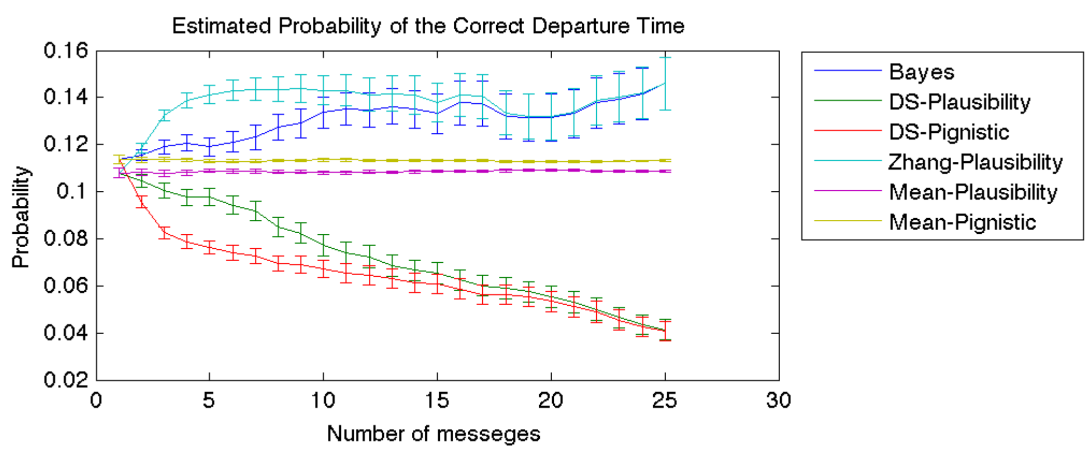

Figure 10. Probability assigned to the correct departure time. 100 runs with parameters $q_k = q_e = 0.4, NU = 0.4$.

It is not surprising that the Dempster-Shafer combination rule performs poorly since it does not take into account the size of the combined sets. Interestingly, mean-pignistic method performs better than mean-plausibility method, implying that the pignistic transformation is more appropriate in our context. This is consistent with the conclusion found in (Smets, 2002) that the pignistic transformation is proper whenever translating beliefs to decisions is required.

It is not trivial that the probability of choosing the correct departure time decreases for some of the methods. This can be explained by the non-uniformity parameter that causes messages that are biased toward incorrect departure times, and by the difficulty to cope with low reliability values.

D. DISCUSSION

In this chapter we compared the characteristics and performance of the updating methods described in Chapter II. The Bayes' method is mathematically rigorous but requires a number of assumptions not needed for the Dempster-Shafer methods. However, there are several ways to implement Dempster-Shafer update, and it is not clear in advance which implementation would be most appropriate for a given scenario.

We have developed a simulation and used it to compare the different updating methods under different conditions. Our analysis reveals that even when the assumptions of the Bayes' update process are violated, that is, if the messages provided by the informant are not constructed uniformly, it still manages to yield the best results. The Dempster-Shafer methods did not perform better than Bayes' update method even though they do not explicitly assume uniformity. Amongst the Dempster-Shafer methods, Zhang's and the mean combination rules perform better. Amongst the transformations from Belief to probabilities, the pignistic transformation was found to be more appropriate in our scenario because the probability assigned is used for decision-making, and not just for representing the measure of belief.

All the methods performed poorly when the reliability of the informant is low, or mistaken to be low, and there is non-uniformity in the way he produces messages.

IV. EXTENSIONS

In the previous chapters we assume a single vessel that can only travel on one route. In this chapter we extend the model to include multiple routes and multiple vessels.

A. SINGLE VESSEL AND MULTIPLE NON-INTERSECTING ROUTES

Let us consider the case where the vessel may use one of multiple routes that do not intersect. Recall in Chapter II that we assume the speed of the target is fixed and known and therefore its location at any time could be derived from the departure time. Now, the location is determined by the departure time and route.

Let $R = \{r_1, ..., r_v\}$ be a set of non-intersecting routes. Thus, the tuple $w \in W \subseteq T \times R$ is sufficient to describe the target location at any given time. We also redefine n as the number of possible combinations of departure time and route, $n = |W|$

We assume that an observation gives us information regarding a combination of departure time and a route ("The Radar has detected a vessel leaving at 8 on the route that is close to the coast"). A message from the informant relates to a subset of the possible departure time and route combinations, for example: "The vessel leaves at 8 or 10 am on route 1 or 2," or "The vessel will be on route 3, departure time is unknown."

Let w' be a two-dimensional parameter denoting the departure time and route. Now we can perform either the Bayesian or Dempster-Shafer theory updates as in Chapter II. For instance, the Bayesian update function after a positive observation w' regarding departure time and route would be (as in Equation (2.5)):

$$f_{update+}(w) = \begin{cases} \dfrac{1 - P_{f-}}{(1 - P_{f-}) \cdot f(w') + P_{f+} \cdot (1 - f(w'))} & w = w' \\ \dfrac{P_{f+}}{(1 - P_{f-}) \cdot f(w') + P_{f+} \cdot (1 - f(w'))} & w \neq w' \end{cases} \tag{4.1}$$

This allows us to calculate the probability that the target has departed at time t from route r, for every combination of departure time and route $w=(t,r)$.

In order to update messages, we must make similar assumptions to the ones made in Chapter II. We assume here that all the true messages of the same size are equally likely, where "size" refers to the number of tuples of departure time and routes combinations in the message. We also assume that the probability that a message is true is determined by the size k.

Following Equation (2.14), the update after a message that is true with probability q_k is

$$f_{update}(w)=\begin{cases}\dfrac{\frac{q_k}{k}}{\sum_{w'\in M} f(w')\cdot\frac{q_k}{k}+\sum_{w'\notin M} f(w')\cdot\frac{1-q_k}{n-k}} & w\in M\\[2ex] \dfrac{\frac{1-q_k}{n-k}}{\sum_{w'\in M} f(w')\cdot\frac{q_k}{k}+\sum_{w'\notin M} f(w')\cdot\frac{1-q_k}{n-k}} & w\notin M\end{cases} \tag{4.2}$$

Applying Dempster-Shafer theory methods to this case is also straightforward. If p is the probability of a true observation regarding departure time and route *w'* , the corresponding mass assignment is:

$$\begin{aligned} m(w') &= p\\ m(W-\{w'\}) &= 1-p \end{aligned} \tag{4.3}$$

and the mass assignment for a message *M* that includes *k* departure times and routes is:

$$\begin{aligned} m(M) &= q_k\\ m(W-M) &= 1-q_k \end{aligned} \tag{4.4}$$

Now that we have assigned masses to the subsets W, the combination rules and probability transformations discussed in Chapter II can be applied exactly in the same manner.

Messages already apply to multiple values in the basic model discussed in Chapter II, and therefore this extension is not relevant for updating messages - the update mechanism for the messages is exactly as in Chapter II. In the following sections we will discuss only the application of extensions to the update process of observations.

B. MULTIPLE ROUTES WITH INTERSECTIONS

Intersecting routes do not affect the updating mechanism regarding the messages because the time and the route aspects of the problem decouple. However, if routes can intersect an observation may apply to more than one route. Let us assume that we receive an observation O regarding an intersection point that may apply to k tuples of routes and departure times. An observation may apply to more than one departure time if, for instance, one of the routes' departure points is further from the intersection point.

Figure 11 depicts an example to such a scenario. In this example, an observation is made at 9:00 in the intersection of multiple possible routes. This observation can be applied to the tuples (departure time is 6:00, "Northwest-Southeast route"), (departure time is 7:00, "middle route") and (departure time is 6:00, "Northeast-Southwest route").

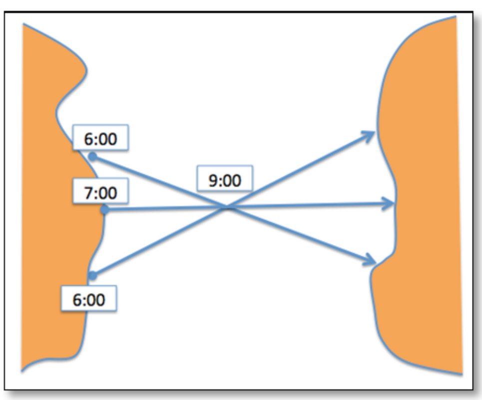

Figure 11. Intersecting routes.

Let us define an observation as a subset of the possible tuples that it relates to $O \subseteq W$. We define the false negative and false positive probabilities of the sensor as in Chapter II; we denote the positive observation, meaning "there is something that corresponds to locations and departure times O" as O_+ and a negative observation, "The sensor did not recognize anything as belonging to O_-." The False positive error is defined as $P\left(O_+ \mid w' \notin O\right) = P_{f+}$ and the false negative error as $P\left(O_- \mid w' \in O\right) = P_{f-}$. It is reasonable to assume that the errors do not depend on the departure time and route and therefore the updated probability of *w'* following a positive observation O_+ is:

$$f_{update}\left(w\right) = \begin{cases} \dfrac{\dfrac{1-P_{f-}}{k}}{\sum\limits_{w' \in O} f\left(w'\right) \cdot \dfrac{\left(1-P_{f-}\right)}{k} + \sum\limits_{w' \notin O} f\left(w'\right) \cdot \dfrac{P_{f+}}{n-k}} & w \in O \\ \dfrac{\dfrac{P_{f+}}{n-k}}{\sum\limits_{w' \in O} f\left(w'\right) \cdot \dfrac{\left(1-P_{f-}\right)}{k} + \sum\limits_{w' \notin O} f\left(w'\right) \cdot \dfrac{P_{f+}}{n-k}} & w \notin O \end{cases} \tag{4.5}$$

The derivation of (4.5) follows the same derivation as Equation (2.5) and is similar to case of receiving a message that contains multiple departure times as in Equation (2.14), but with one difference: while for messages we had to assume some uniformity among the messages produced, here the fact that the errors are independent is sufficient.

If the vessel can switch routes at the intersection points, as shown in Figure 12, the situation is slightly more complicated:

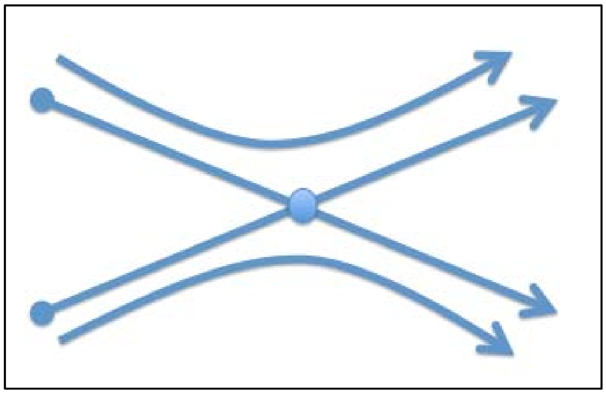

Figure 12. Intersecting routes with possibility of switching routes.

In this case we can formulate the problem as consisting of four routes, two of which overlap at the departure and arrival points, and all four intersect in the middle point. Similarly, wherever there is an intersection of r routes in an intersection point, we can define r^2 distinct routes. After this small alteration we can apply the update process defined in (4.5) to solve this case as well,

C. MULTIPLE TARGETS

Dealing with multiple targets is a much more complicated topic. However, when the total number of targets in the area of interest is known, a similar scheme to the one presented in sections A and B of this chapter can be applied. We assume that there is only one possible route, as in Chapter II.

Let us assume that there are u targets $Z = \{z_1, ..., z_u\}$. The set $w = \{t, z\}$ specifies that target z departed at time t. We redefine our space of interest as $W \subseteq T \times Z$, $|W| = n$, which describes every possible vessel's departure time and identity.

An observation can relate to a specific subset of departure times and vessels. If the sensor can recognize vessels, the observation will include only a single one, if it can classify them into different categories than the observation may relate to a subset of the vessels.

Assuming that the sensor's false positive and false negative probabilities do not depend on the type of the target, the update function following a positive observation O_+ is:

$$f_{update}(w)=\begin{cases}\dfrac{\frac{1-P_{f-}}{k}}{\sum\limits_{w'\in O} f(w')\cdot\frac{(1-P_{f-})}{k}+\sum\limits_{w'\notin O} f(w')\cdot\frac{P_{f+}}{n-k}} & w\in O\\ \dfrac{\frac{P_{f+}}{n-k}}{\sum\limits_{w'\in O} f(w')\cdot\frac{(1-P_{f-})}{k}+\sum\limits_{w'\notin O} f(w')\cdot\frac{P_{f+}}{n-k}} & w\notin O\end{cases} \tag{4.6}$$

Exactly as in (4.5)

If the sensor characteristics depend on the type of the vessel (which is reasonable, since bigger vessels are easier to detect, and debris are more likely to be misclassified as small vessels), the situation is more complicated. Now the false positive error is a function of the target $P(O_+ \mid w\notin O, z)=P_{f+}(z)$ and likewise the false negative error is $P(O_- \mid w\in O, z)=P_{f-}(z)$.

The update process can be formulated as:

$$f_{update}(w=(t,z))=\begin{cases}\dfrac{\frac{1-P_{f-}(z)}{k}}{\sum\limits_{w'=(t',z')\in O} f(w')\cdot\frac{(1-P_{f-}(z'))}{k}+\sum\limits_{w'=(t',z')\notin O} f(w')\cdot\frac{P_{f+}(z')}{n-k}} & w\in O\\ \dfrac{\frac{P_{f+}(z)}{n-k}}{\sum\limits_{w'=(t',z')\in O} f(w')\cdot\frac{(1-P_{f-}(z'))}{k}+\sum\limits_{w'=(t',z')\notin O} f(w')\cdot\frac{P_{f+}(z')}{n-k}} & w\notin O\end{cases} \tag{4.7}$$

This scheme works nicely for a known number of targets, however this Bayesian update process is straightforward to implement in cases where the number of vessels is not known or changes during the scenario.

D. DIFFERENT VELOCITIES

Let us again consider the case of a single target and a single route but multiple fixed velocities. Let us define a set of possible velocities $V = \{v_1, ..., v_h\}$ and discuss the probability distribution of $\{t, v\} = w$ with our space of interest being $W \subseteq T \times V$ and $|W| = n$.

An observation that was taken at time t regarding a certain location – say at a distance d along the route needs to be translated to the corresponding subset of W: all the departure time and velocity combinations that would bring the vessel to the location of the observation at the time it was taken. The subset that corresponds to this observation can be constructed by:

$$O = \{w' = (t', v') : (t - t') \cdot v' = d\} \tag{4.8}$$

Now that we have established the subset of W that the observation refers to, we can continue with the update process as in the previous sections.

E. SUMMARY

In this chapter we have proposed several extensions to the basic model developed in Chapter II to accommodate for multiple routes (with and without intersections), multiple targets and multiple velocities. The probability updating process for both observations and messages is very similar to the processes discussed in Chapters II and III and requires only small changes to accommodate those extensions. Therefore we expect that the results found in Chapter III extend to these cases as well. Accommodating multiple extensions simultaneously requires more bookkeeping and notation, but is straightforward using the techniques described in this chapter.

Known correlations between the parameters of the model, such as the velocity and the type of vessel, can be incorporated into the prior distribution. The prior can also account for intelligence regarding the likelihood of the possible routes.

THIS PAGE INTENTIONALLY LEFT BLANK

V. ASSESSING THE INFORMANT'S RELIABILITY

In previous chapters we assumed that the reliability q_k of an informant is known. In reality this will not be known. As the operator receives more information from an informant, he should update his estimate of the informant's reliability. In this chapter we develop a method of estimating and updating the reliability parameter q_k, together with the distribution of the departure time T_d. A key factor in updating the reliability is whether the operator is able to verify the truthfulness of the informant's previous messages. In section A we examine the situation where the operator is able to verify the truth of the message, and in section B we examine the more challenging situation where the operator does not have this capability.

A. MESSAGE TRUTH CAN BE VERIFIED

We assume that the operator does not know the exact reliability parameter q_k, but has some prior knowledge about the probability distribution of it. Recall that q_k is the probability that a message of size *k* contains the true departure time. Equivalently we can interpret it as the long run fraction of messages that are true/correct. Over time, the operator observes whether messages of a certain size are true and uses this information to update the probability distribution of q_k. This situation is similar to the problem of flipping a biased coin with unknown probability for obtaining a Head and updating that probability over time based on the observed outcome of the flips. We treat the reliability parameter as a random variable Q_k that receives the values $0 \le q_k \le 1$. Let us define X as a random variable denoting whether the message received is verified to be true - $X = 1$ or untrue - $X = 0$. The probability of receiving a true or false message is (by definition):

$$P\left(X = x \mid q_k\right) = q_k^x \left(1 - q_k\right)^{1-x} \tag{5.1}$$

We can use the expression in Equation (5.1) to formulate an update for the distribution of Q_k based on the veracity of the most recent message

$$P\left(Q_k = q_k \mid X = x\right) = \frac{P\left(X = x \mid q_k\right)\cdot P_{prior}\left(Q_k = q_k\right)}{P\left(X = x\right)} = \frac{q_k^x\left(1-q_k\right)^{1-x}\cdot P_{prior}\left(Q_k = q_k\right)}{P\left(X = x\right)} \tag{5.2}$$

where $P\left(X = x\right)$ is the probability the operator verifies the message as true.

We would like the posterior distribution of Q_k to be from the same family as its prior, so that the calculations are tractable. In Bayesian terminology, this characteristic is called a "conjugate prior." The Beta distribution satisfies this condition for the case of Bernoulli trials and therefore it is common to use the Beta distribution as the prior (see Berger, 1993). The Beta distribution has two parameters α,β, with mean $E\left[Q_k\right] = \frac{\alpha}{\alpha+\beta}$, and probability distribution function:

$$f_{Q_k}^{prior}\left(q_k\right) = \frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta-1}}{B\left(\alpha,\beta\right)} \tag{5.3}$$

where the denominator of (5.3) is the Beta function, which is defined as $B\left(x,y\right) = \int_0^1 t^{x-1}\left(1-t\right)^{y-1}dt$. By substituting the prior (5.3) into the Bayes' theorem (5.2) we obtain the posterior distribution of Q_k:

$$f_{Q_k}\left(q_k \mid x\right) = \frac{P\left(x \mid q_k\right) f_{Q_k}\left(q_k\right)}{P\left(x\right)} = \frac{q_k^x\left(1-q_k\right)^{1-x}\cdot q_k^{\alpha-1}\left(1-q_k\right)^{\beta-1}}{B\left(\alpha,\beta\right)\cdot P\left(x\right)} = \frac{q_k^{\alpha-1+x}\left(1-q_k\right)^{\beta-x}}{B\left(\alpha,\beta\right)\cdot P\left(x\right)} \tag{5.4}$$

And since $\cdot P\left(x\right)$ is merely a normalizing term, we obtain:

$$f_{Q_k}\left(q_k \mid x\right) \sim Beta\left(\alpha+x,\beta+1-x\right) \tag{5.5}$$

As desired, the posterior distribution of Q_k also has a Beta distribution but with different parameters. As seen from equation (5.5) a true message ($\chi = 1$) will increase the α parameter of the distribution by 1 while a false message will increase the β parameter by 1.

Let us limit the current discussion to a single informant, providing messages of a specific size k and reliability Q_k. Although q_k is now the value of a random variable we can apply the same reasoning as in Chapter II in conjunction with the law of total expectation to generalize the calculation of receiving a specific message M_k in Equation (2.11):

$$P\left(M_k \mid T_d = t\right) = \begin{cases} \int_{q_k} q_k \cdot \frac{1}{\binom{n-1}{k-1}} f\left(q_k\right) dq_k = \frac{1}{\binom{n-1}{k-1}} E\left[q_k\right] & t \in M \\ \int_{q_k} \left(1-q_k\right) \cdot \frac{1}{\binom{n-1}{k}} f\left(q_k\right) dq_k = \frac{1}{\binom{n-1}{k}} \left(1 - E\left[q_k\right]\right) & t \notin M \end{cases} \quad (5.6)$$

As in Chapter II, we assume that all true messages of a certain size are equally likely, and all false ones are also equally likely. When a new piece of intelligence is received, the departure time distribution can be updated using the same method as in Chapter II, while replacing q_k with $E\left[Q_k\right]$ which equals $\frac{\alpha}{\alpha+\beta}$ in the case of the Beta distribution. Since the departure time update is the same as in Chapter II, in this section we focus on the update of Q_k.

Let us discuss the values we should assign to the parameters α, β in the prior distribution. If we have no prior knowledge about the reliability of the informant, a common uninformative conjugate prior for q_k is $Beta\left(1/2, 1/2\right)$. This is known as Jeffrey's prior and includes the minimal information regarding the distribution of Q_k among all possible priors (Berger, 1993). However, if we do have a prior estimate of the reliability we can use it to set the prior. The mean reliability is given by $\frac{\alpha}{\alpha+\beta} = E\left[Q_k\right]$.

The "weight" of the prior can be controlled by the values of α and β - the larger α and β, the less the posterior distribution will change with new evidence. Let us look at two cases: in the first case $\alpha = \beta = 1$ and in the second case, $\alpha = \beta = 10$. In both cases $E\left[Q_k\right] = \frac{1}{2}$, however after a single true message that has been verified, the estimated reliability of the informant in the first case would be $E\left[Q_k\right] = \frac{\alpha+1}{\alpha+1+\beta} = \frac{2}{3}$ while in the second case $E\left[Q_k\right] = \frac{\alpha+1}{\alpha+1+\beta} = \frac{11}{21}$.

We assume that the larger the message size *k* the larger $E\left[Q_k\right]$ will be since there are more possibilities for the message to contain the true departure time. We would also expect $E\left[Q_k\right] \geq \frac{k}{n}$ for a "useful" informant that provides correct messages with probability higher than that generated uniformly random.

In practice the verification can occur in many ways, for example later intelligence may confirm without any doubt the location of the vessel, or the perhaps by the capture and interrogation of drug smugglers. Once the true departure time of the vessel is determined, the truthfulness of the messages received can determined as well, and can be used to update the reliability of the informant, as described in this section. The updated reliability can then be used to better estimate the departure time of other vessels based on messages from the same informant.

B. UNVERIFIED MESSAGES

In many situations, the operator cannot verify the informant's information. However, we still want to use the new information provided by the informant to update both the departure time distribution and informant's reliability. For instance, if we receive a message that contradicts everything we know so far, it is more likely that the informant is incorrect, and therefore his reliability should be updated downwards. Unlike the update procedure in section A of this chapter where we update the time departure distribution and the reliability sequentially, here we update the time departure distribution and the reliability parameter simultaneously.

1. The Bayesian Update Process

The main idea behind the simultaneous update process is to define a joint probability distribution for the time departure T_d and reliability Q_k, denoted by $f_{T_d,Q_k}\left(t_d,q_k\right)$. This is a mixed joint distribution, since the time parameter is a discrete random variable while the reliability parameter is continuous. As in section A of this chapter, the random reliability parameter Q_k will take value q_k.

The update of the joint distribution of the true departure time T_d and reliability parameter Q_k when a new message M_k is received can be calculated by applying Bayes' theorem:

$$f_{T_d,Q_k}^{new}\left(t_d,q_k\right)=\frac{P\left(M_k \mid T_d=t_d,Q_k=q_k\right)\cdot f_{T_d,Q_k}^{prior}\left(t_d,q_k\right)}{P\left(M_k\right)} \tag{5.7}$$

In order to evaluate (5.7) we need to define the prior $f_{T_d,Q_k}^{prior}\left(t_d,q_k\right)$ and the update function $f_{T_d,Q_k}^{update}\left(t_d,q_k\right)=\dfrac{P\left(M_k \mid T_d=t_d,Q_k=q_k\right)}{P\left(M_k\right)}$.

The prior $f_{T_d,Q_k}^{prior}\left(t_d,q_k\right)$, is combined from two parts: (1) departure time, and (2) the reliability. We assume that these two components are independent for the prior. For the departure time, the prior can include any information, however without further knowledge we assume it to be uniform among the departure times: $T_d \sim U\left[T\right]$. For the reliability part we assume that a prior from the Beta family as in section A is appropriate and so $Q_k \sim Beta\left(\alpha,\beta\right)$. The independence between T_d and Q_k holds only for the prior, but not for the updated distribution. The final expression of the joint distribution prior is:

$$f_{T_d,Q_k}^{prior}\left(t,q_k\right) = P\left(T_d = t\right) \cdot f_{Q_k}\left(q_k\right) = \frac{1}{n} \cdot \frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta-1}}{B\left(\alpha,\beta\right)} \tag{5.8}$$

Next let us derive the update function $f_{T_d,Q_k}^{update}\left(t_d,q_k\right) = \frac{P\left(M_k \mid T_d = t_d, Q_k = q_k\right)}{P\left(M_k\right)}$.

First we need to calculate $P\left(M_k \mid T_d = t, Q_k = q_k\right)$. As in Chapter II, Equation (2.11), the probability of receiving a message M_k given the true departure time T_d and reliability parameter Q_k is:

$$P\left(M_k \mid T_d = t, Q_k = q_k\right) = \begin{cases} \dfrac{q_k}{\binom{n-1}{k-1}} & t \in M_k \\ \dfrac{1-q_k}{\binom{n-1}{k}} & t \notin M_k \end{cases} \tag{5.9}$$

By defining an indicator function $I\left(s \in M_k\right) = \begin{cases} 1 & s \in M_k \\ 0 & s \in M_k \end{cases}$ we can combine the two cases in (5.9):

$$P\left(M_k \mid T_d = t, Q_k = q_k\right) = \left(\frac{q_k}{\binom{n-1}{k-1}}\right)^{I\left(t \in M_k\right)} \left(\frac{1-q_k}{\binom{n-1}{k}}\right)^{1-I\left(t \in M_k\right)} \tag{5.10}$$

Notice that there are two ways for the probability of receiving a certain message M_k in (5.10) to be large. Either the informant is considered reliable (high Q_k) and the message includes the true departure time, or the message does not include the true departure time and the informant is considered unreliable (low Q_k)

The denominator of the update function $P\left(M_k\right)$ is the normalizing term and can be calculated by integrating over T_d and Q_k:

$$P\left(M_k\right)=\sum_{t\in T}\int_0^1 f_{T_d,Q_k}\left(t,q_k\right)\cdot P\left(M_k \mid T_d=t,Q_k=q_k\right)dq_k \tag{5.11}$$

which in our case for the prior translates to:

$$\begin{aligned}&=\sum_{t\in M_k}\int_0^1 \frac{1}{n}\cdot\frac{1}{\binom{n-1}{k-1}}\cdot\frac{q_k^{\alpha}\left(1-q_k\right)^{\beta-1}}{B\left(\alpha,\beta\right)}dq_k+\sum_{t\notin M_k}\int_0^1 \frac{1}{n}\cdot\frac{1}{\binom{n-1}{k}}\cdot\frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta}}{B\left(\alpha,\beta\right)}dq_k\\&=\frac{k}{n\binom{n-1}{k-1}}\cdot\frac{B\left(\alpha+1,\beta\right)}{B\left(\alpha,\beta\right)}+\frac{n-k}{n\binom{n-1}{k}}\cdot\frac{B\left(\alpha,\beta+1\right)}{B\left(\alpha,\beta\right)}=\frac{k!\left(n-k\right)!}{n!}=\frac{1}{\binom{n}{k}}\end{aligned} \tag{5.12}$$

This simple expression for $P\left(M_k\right)$ – the probability that the informant provides the message M_k – before the arrival of any new information is not surprising given the uniform prior and uniform generation of messages. There are $\binom{n}{k}$ possible messages that can be generated by picking k departure times from n possibilities and so the probability of picking one of them uniformly is $\frac{1}{\binom{n}{k}}$.

Now that we have both the prior and the update function we can evaluate the new joint distribution after receiving a message:

$$f^{update}_{T_d,Q_k}\left(t,q_k\right)=\frac{P\left(M_k \mid T_d=t,Q_k=q_k\right)}{P\left(M_k\right)}=n\cdot\left(\frac{q_k}{k}\right)^{I\left(t\in M\right)}\left(\frac{1-q_k}{n-k}\right)^{1-I\left(t\in M\right)} \tag{5.13}$$

Combining the prior Equation (5.8) and the update function (5.13), the updated joint distribution is:

$$
\begin{aligned}
f_{T_d,Q_k}^{new}\left(t_d,q_k\right) &= n\cdot\left(\frac{q_k}{k}\right)^{I\left(s\in M_k\right)}\left(\frac{1-q_k}{n-k}\right)^{1-I\left(s\in M_k\right)}\cdot\frac{1}{n}\cdot\frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta-1}}{B\left(\alpha,\beta\right)}\\
&=\left(\frac{1}{k}\right)^{I\left(s\in M_k\right)}\left(\frac{1}{n-k}\right)^{1-I\left(s\in M_k\right)}\frac{q_k^{\alpha-1+I\left(s\in M_k\right)}\left(1-q_k\right)^{\beta-I\left(s\in M_k\right)}}{B\left(\alpha,\beta\right)}
\end{aligned}
\tag{5.14}
$$

A convenient way to think about this joint distribution is by envisioning n different reliability parameters $Q_{k,i}$, one for each possible departure time. Each of those variables has a Beta distribution with parameters α_i and β_i. To update the joint probability distribution we increase α_i by 1 for departure times that are included in the message and increase β_i by 1 for departure times that are not included in the message. This increases the expected $Q_{k,i}$ for departure times that are included in messages, and decreases the expected $Q_{k,i}$ of departure times that are not included.

2. The Marginal Distributions after a Single Message

In order to gain more insight on the influence of a message on the distribution of T_d and Q_k, we calculate the marginal distributions after a single message.

a. The Marginal Distribution of T_d

The general expression for calculating a marginal distribution of a joint distribution is $P\left(T_d=t\right)=\int_0^1 f_{Q_k|T_d=t}\left(t,q_k\right)dq_k$. Applying it to the joint distribution defined in (5.14) yields:

$$P\left(T_d = t\right) = \frac{1}{B\left(\alpha,\beta\right)} \cdot \begin{cases} \frac{1}{k} \cdot \int_0^1 q_k^{\alpha} \cdot \left(1-q_k\right)^{\beta-1} dq_k = \frac{B\left(\alpha+1,\beta\right)}{k} & t \in M_k \\ \frac{B\left(\alpha,\beta+1\right)}{n-k} & t \notin M_k \end{cases} \tag{5.15}$$

But we can simplify this expression since the Beta function has the property $\frac{B\left(\alpha+1,\beta\right)}{B\left(\alpha,\beta\right)} = \frac{\alpha}{\alpha+\beta}$ so the marginal distribution of the departure time is simply:

$$P\left(T_d = t\right) = \begin{cases} \frac{1}{k} \cdot \frac{\alpha}{\alpha+\beta} & t \in M_k \\ \frac{1}{n-k} \cdot \frac{\beta}{\alpha+\beta} & t \notin M_k \end{cases} \tag{5.16}$$

Just like in Chapter II, Equation (2.15), the probability of a certain departure time, given a message M_k depends on the size of the message k, and the expected value of the reliability of the message $\frac{\alpha}{\alpha+\beta} = E\left[Q_k\right]$. In order for the marginal distribution of the times that are in the message to increase, we require that $\frac{1}{k} \cdot \frac{\alpha}{\alpha+\beta} > \frac{1}{n-k} \cdot \frac{\beta}{\alpha+\beta}$, or $\frac{\alpha}{\alpha+\beta} > \frac{k}{n}$. We can interpret this condition as stating the informant must be "useful," that is, one that has a higher probability to generate a true message than if one were to pick a message uniformly at random. The probability that the message contains the true value is equal to the expected reliability:

$$P\left(T_d \in M_k\right) = k \cdot \frac{1}{k} \cdot \frac{\alpha}{\alpha+\beta} = \frac{\alpha}{\alpha+\beta} = E\left[Q_k\right] \tag{5.17}$$

which is exactly the desired quality of the reliability notion.

b. *The Marginal Distribution of* Q_k

The general formula for the marginal distribution over a discrete random variable is $f_{Q_k}\left(q_k\right)=\sum_{t\in T} f_{Q_k|T_d=t}\left(q_k,t\right)$ which in our case translates to:

$$f_{Q_k}\left(q_k\right)=\frac{q_k^{\alpha}\left(1-q_k\right)^{\beta-1}}{B\left(\alpha,\beta\right)}+\frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta}}{B\left(\alpha,\beta\right)}=\frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta-1}}{B\left(\alpha,\beta\right)} \tag{5.18}$$

The marginal distribution of Q_k after a single message remains the same. This result is intuitive. Because the prior distribution of the departure time is uniform, we cannot utilize that distribution to gain information about whether the first message is true or false (or in other words, all the possible messages are as likely to be received) and therefore the estimated reliability remains unchanged. We cannot draw any information about the informant from his message because we did not have any specific information about the departure times. If the prior distribution for the departure time was not uniform, the marginal distribution of Q_k would change.

The mean of the reliability is the same as before this message:

$$E\left[Q_k\right]=\int_0^1 f_{Q_k}\left(q_k\right)q_k dq_k=\frac{B\left(\alpha+1,\beta\right)}{B\left(\alpha,\beta\right)}=\frac{\alpha}{\alpha+\beta} \tag{5.19}$$

In order to make this model clearer, let us look at a numeric example, following the example in Chapter II, section D.4. We will assume that there are only four possible departure times: $T=\{t_1,t_2,t_3,t_4\}$, of which the true departure time is *t_1*. We receive a message of size *k=2*, $M_2=\{t_1,t_2\}$ and the reliability our informant for messages of size 2 is $Q_2 \sim Beta\left(\alpha=9,\beta=1\right)$ (and $E\left[Q_2\right]=0.9$). The prior probability distribution is assumed to be uniform: $f_{prior}\left(t_1\right)=f_{prior}\left(t_2\right)=f_{prior}\left(t_3\right)=f_{prior}\left(t_4\right)=0.25$.

The joint distribution, as developed in Equation (5.14) in this example is calculated to be:

$$f_{T_d,Q_2}^{new}\left(t,q_2\right)=\begin{cases}\frac{1}{2}\cdot\frac{q_2^9\cdot\left(1-q_2\right)^0}{B\left(9,1\right)} & T_d=t_1,t_2\\ \frac{1}{2}\cdot\frac{q_2^8\cdot\left(1-q_2\right)^1}{B\left(9,1\right)} & T_d=t_3,t_4\end{cases} \tag{5.20}$$

The updated marginal distribution will be:

$$\begin{aligned}f\left(T_d=t_1\right)=f\left(T_d=t_2\right)&=\int_0^1\frac{1}{2}\cdot\frac{q_2^9\cdot\left(1-q_2\right)^0}{B\left(9,1\right)}dq_2=\frac{1}{2}\frac{1/10}{1/9}=0.45\\ f\left(T_d=t_3\right)=f\left(T_d=t_4\right)&=\int_0^1\frac{1}{2}\cdot\frac{q_2^8\cdot\left(1-q_2\right)^1}{B\left(9,1\right)}dq_2=\frac{1}{2}\frac{\left(1/9-1/10\right)}{1/9}=0.05\end{aligned} \tag{5.21}$$

Which is in full agreement with the results in (2.16).

The marginal distribution of Q_k is now:

$$f_{Q_k}\left(q_k\right)=\frac{q_k^9\left(1-q_k\right)^0}{B\left(9,1\right)}+\frac{q_k^8\left(1-q_k\right)^1}{B\left(9,1\right)}=\frac{q_k^8\left(1-q_k\right)^0}{B\left(9,1\right)} \tag{5.22}$$

which means that $Q_k\sim Beta\left(9,1\right)$ as before the message.

In the following section, we will discuss an example where the marginal distribution of Q_k does change after the update.

3. The Joint Distribution after Two Identical Messages

We will now examine how the joint distribution evolves after two identical messages. By applying the update process described in (5.7) twice, with the new prior being the resulting posterior distribution after the first update (5.14) we obtain the expression for the distribution after two identical messages M_k with C as the normalizing constant:

$$f^{new}_{T_d,Q_k}\left(t_d,q_k\right)=C\cdot\left[\left(\frac{q_k}{k}\right)^{I(s\in M_k)}\left(\frac{1-q_k}{n-k}\right)^{1-I(s\in M_k)}\right]^2\cdot\frac{1}{n}\cdot\frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta-1}}{B(\alpha,\beta)} \tag{5.23}$$

By inserting all the constants into the normalizing parameter and renaming it, we obtain:

$$f^{new}_{T_d,Q_k}\left(t_d,q_k\right)=C'\cdot\left(\frac{1}{k}\right)^{2\cdot I(s\in M_k)}\left(\frac{1}{n-k}\right)^{2(1-I(s\in M_k))}q_k^{\alpha-1+2\cdot I(s\in M_k)}\left(1-q_k\right)^{\beta-1+2\cdot(1-I(s\in M_k))} \tag{5.24}$$

We calculate the normalizing term C' by integrating the joint density to 1:

$$\begin{aligned}
&\sum_t\int_0^1 f^{new}_{T_d,Q_k}\left(t_d,q_k\right)dq_k=1\\
&C'\cdot\left[k\cdot\frac{B(\alpha+2,\beta)}{k^2}+(n-k)\frac{B(\alpha,\beta+2)}{(n-k)^2}\right]=1\\
&C'=\frac{1}{\frac{B(\alpha+2,\beta)}{k}+\frac{B(\alpha,\beta+2)}{n-k}}
\end{aligned} \tag{5.25}$$

And now we can calculate the updated distribution in (5.24). Unlike in section 2b of this chapter, the marginal distribution of Q_k now changes after the second message. We compute $f_{Q_k}\left(q_k\right)=\sum_{t\in T}f_{Q_k|T_d=t}\left(q_k,t\right)$ using the new distribution after two messages obtained in (5.24)

$$f_{Q_k}\left(q_k\right)=C'\cdot\left[\frac{q_k^{\alpha+1}\left(1-q_k\right)^{\beta-1}}{k}+\frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta+1}}{n-k}\right] \tag{5.26}$$

We can see that the marginal distribution of Q_k that was a single Beta distribution before the update became a mixture of two Beta distributions after the update. The mean of Q_k after two messages is:

$$
\begin{aligned}
E\left[Q_k\right] &= \int_0^1 f_{Q_k}\left(q_k\right) q_k \, dq_k \\
&= C' \cdot \int_0^1 \left[\frac{q_k^{\alpha+1}\left(1-q_k\right)^{\beta-1}}{k} + \frac{q_k^{\alpha-1}\left(1-q_k\right)^{\beta+1}}{n-k} \right] q_k \, dq_k \\
&= \frac{\alpha}{\alpha+\beta+2} + \frac{2}{\alpha+\beta+2} \cdot \frac{\frac{\alpha\left(\alpha+1\right)}{k}}{\frac{\alpha\left(\alpha+1\right)}{k} + \frac{\beta\left(\beta+1\right)}{n-k}}
\end{aligned}
\tag{5.27}
$$

For the case in which the parameters satisfy $\frac{\alpha+1}{\alpha+\beta+2} > \frac{k}{n}$ the mean of the updated marginal distribution is bigger than before the update:

$$
\begin{aligned}
&E\left[Q_k\right] = \frac{\alpha}{\alpha+\beta+2} + \frac{2}{\alpha+\beta+2} \cdot \frac{\frac{\alpha\left(\alpha+1\right)}{k}}{\frac{\alpha\left(\alpha+1\right)}{k} + \frac{\beta\left(\beta+1\right)}{n-k}} > \\
&\frac{\alpha}{\alpha+\beta+2} + \frac{2 \cdot \frac{\alpha}{\alpha+\beta}}{\alpha+\beta+2} = \frac{\frac{\alpha^2+\alpha\beta+2\alpha}{\alpha+\beta+2}}{\alpha+\beta} = \frac{\alpha}{\alpha+\beta}
\end{aligned}
\tag{5.28}
$$

This is a reminiscent of the requirement of $\frac{\alpha}{\alpha+\beta} > \frac{k}{n}$ for the informant to be considered useful from the section 2.b. Note that in cases when the informant is not "useful" (meaning, his mean reliability is lower than what it would be if he had picked departure times randomly), the updated $E\left[Q_k\right]$ might decrease. This occurs because if we believe the informant is very unreliable then we believe that the departure times that he stated in the message are less likely, and in this light, we downgrade his reliability even further.

4. Visualization of the Update Process

We develop a small simulation that visualizes the joint distribution update process. In the following example $n=9, k=3$ and the initial Q_k parameters value are $\alpha=4, \beta=1$ leading to $E\left[Q_k\right]=0.8$.

The joint distribution before any messages arrive is depicted in Figure 13:

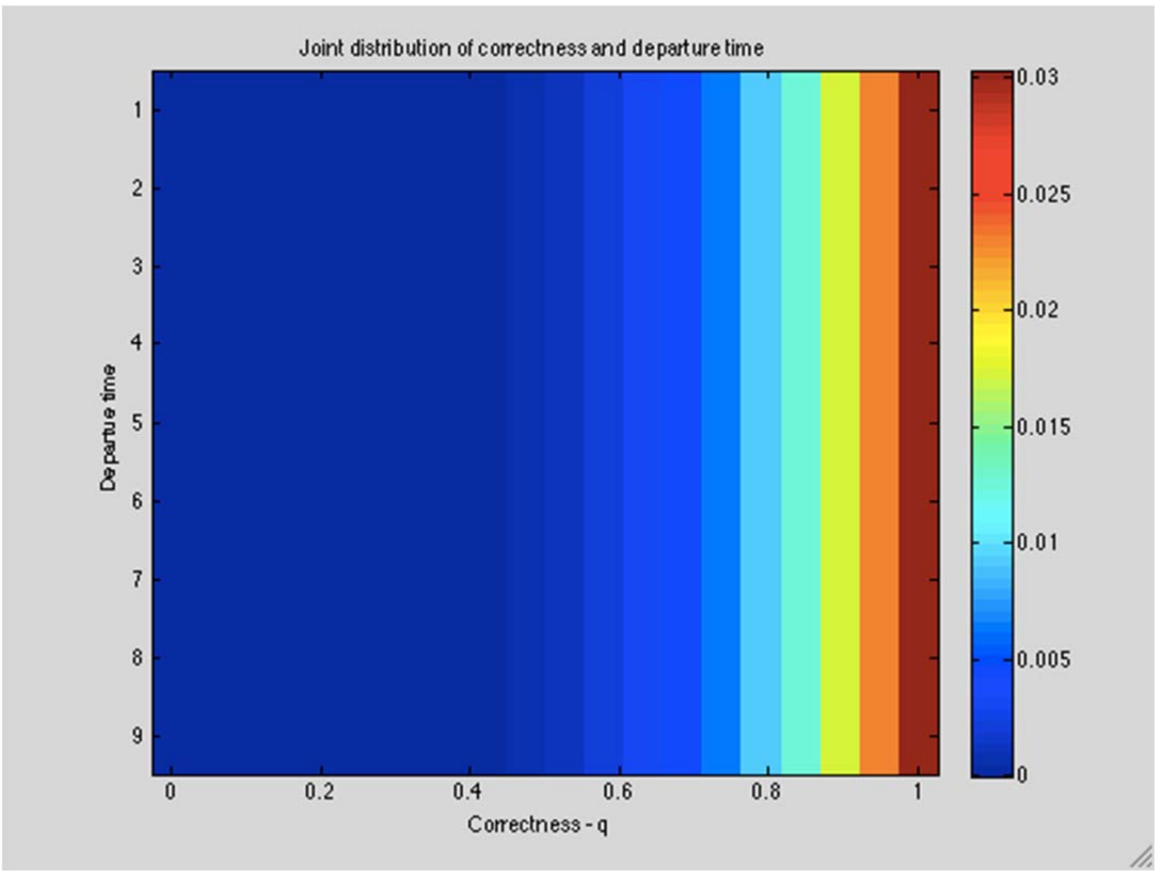

Figure 13. Joint prior distribution.

As we can see the variables are independent. The departure time is uniform, and the higher values of Q_k are more likely. Now let us examine what happens after receiving a single message $M_k=\{T_4, T_5, T_9\}$. The joint distribution changes to:

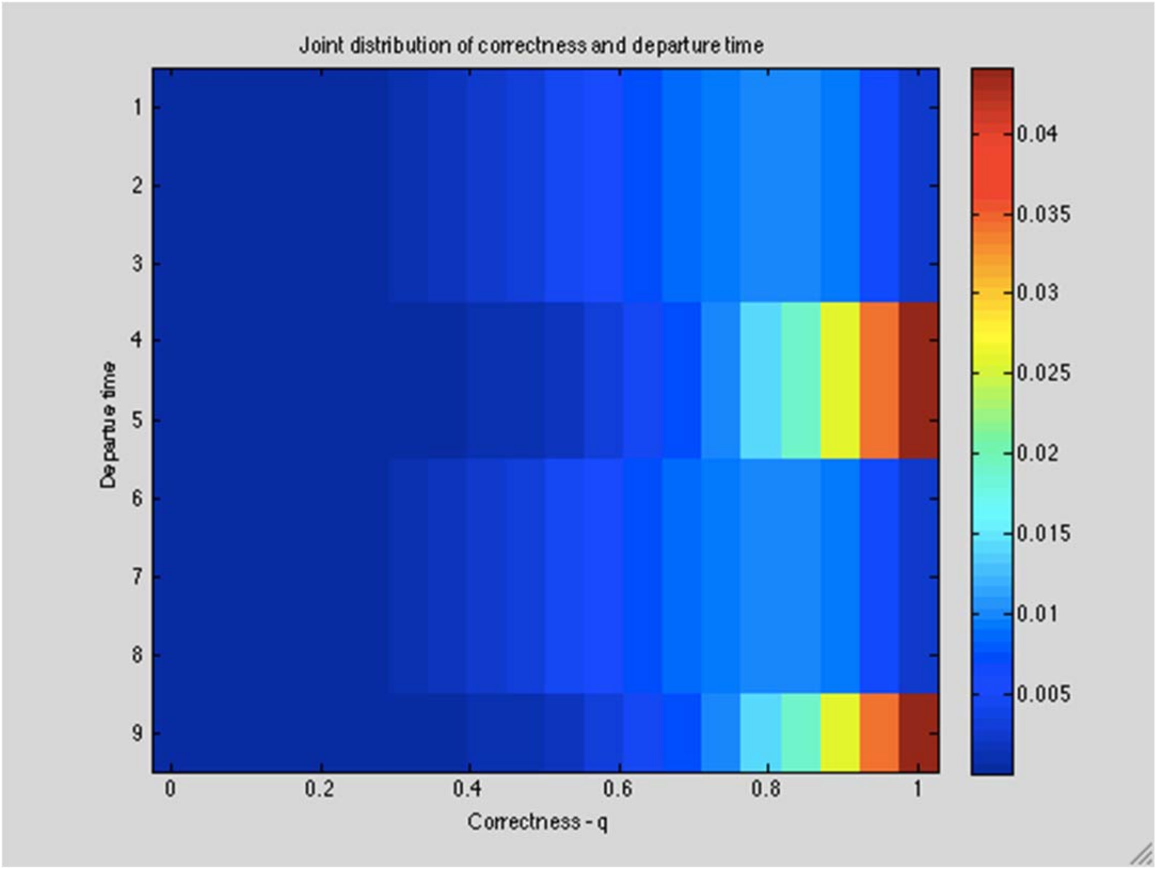

Figure 14. Joint distribution after a single message.

This implies that either the true departure time is one of $\{T_4, T_5, T_9\}$ and the reliability of the informant is rather high, or the true departure time is not one of the above, and the informant reliability is low. If we receive another message $M_k = \{T_3, T_7, T_9\}$, the distribution update can be seen in Figure 15:

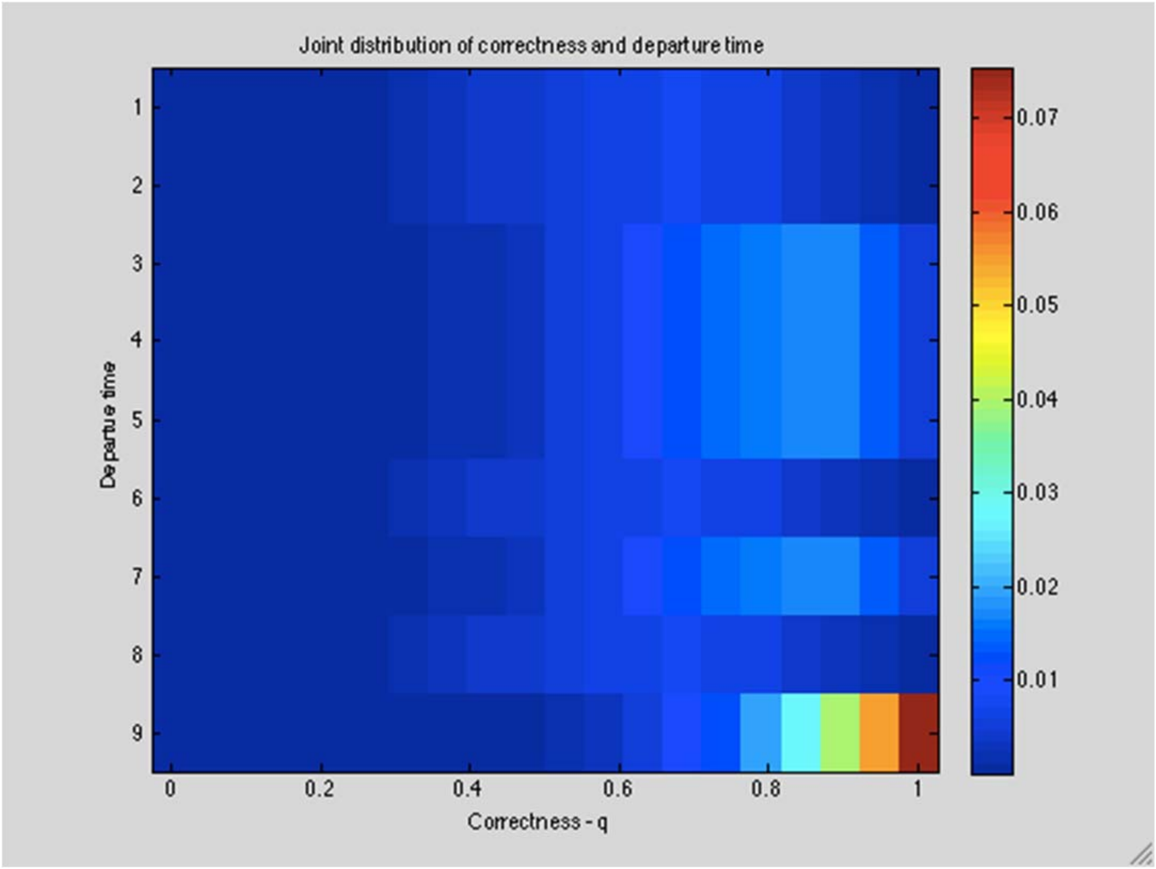

Figure 15. Joint distribution after a two messages.

It is very likely that the true departure time is T_9, with a very reliable informant. However, there is still some probability that the true departure time is not T_9 and the reliability is lower.

5. Change of T_d and Q_k with Number Messages

It is also interesting to examine the distribution of T_d and the mean of Q_k after multiple messages. In the following example, $n = 2, k = 1$ and the initial Q_k parameters value are $\alpha = 3, \beta = 2$ leading to $E\left[Q_k\right] = 0.6$.

Let us assume we receive 10 identical messages: $M_k = \left\{T_1\right\}$. The distribution of T_d as function of number of messages is plotted in Figure 16:

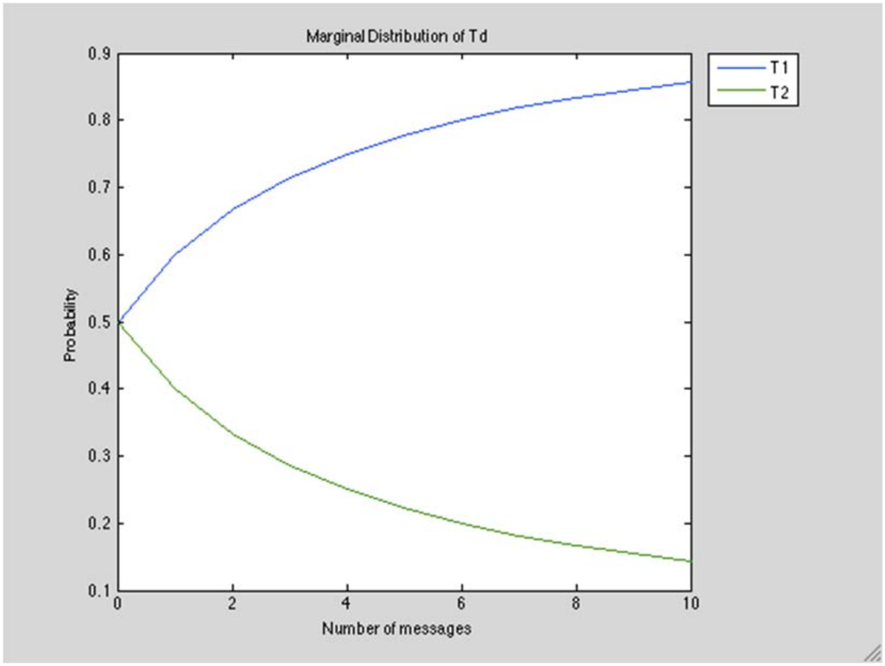

Figure 16. Marginal distribution of T_d.

As expected, the probability of T_1 increases with the number of messages, while the probability of T_2 decreases.

Figure 17 shows how the mean of Q_k changes with the number of messages received:

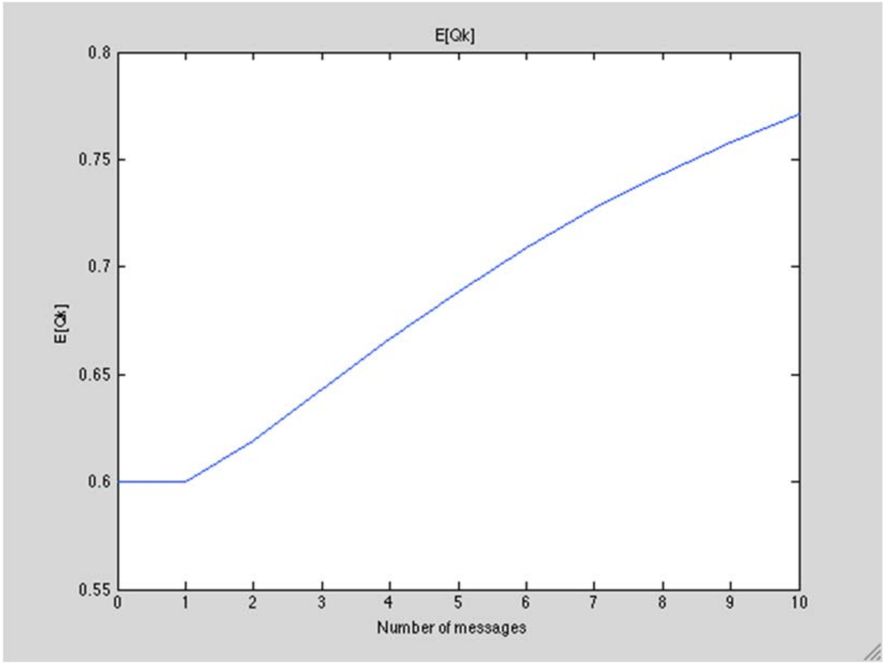

Figure 17. Mean of Q_k.

As more messages arrive, the mean of Q_k increases. The repeated messages confirm both the departure time (as shown in Figure 16) and the informant's reliability. Note that the first message does not change the mean of Q_k.

Now to a slightly more interesting example: let us assume that $n = 2, k = 1$ and the initial Q_k parameters value are $\alpha = 2, \beta = 3$ leading to $E\left[Q_k\right] = 0.4$. The mean of Q_k now is lower than the uniform probability of picking T_1 at random, and therefore the informant is more likely to provide incorrect information. Again, we assume we receive 10 identical messages: $M_k = \{T_1\}$

The marginal distributions in this case are:

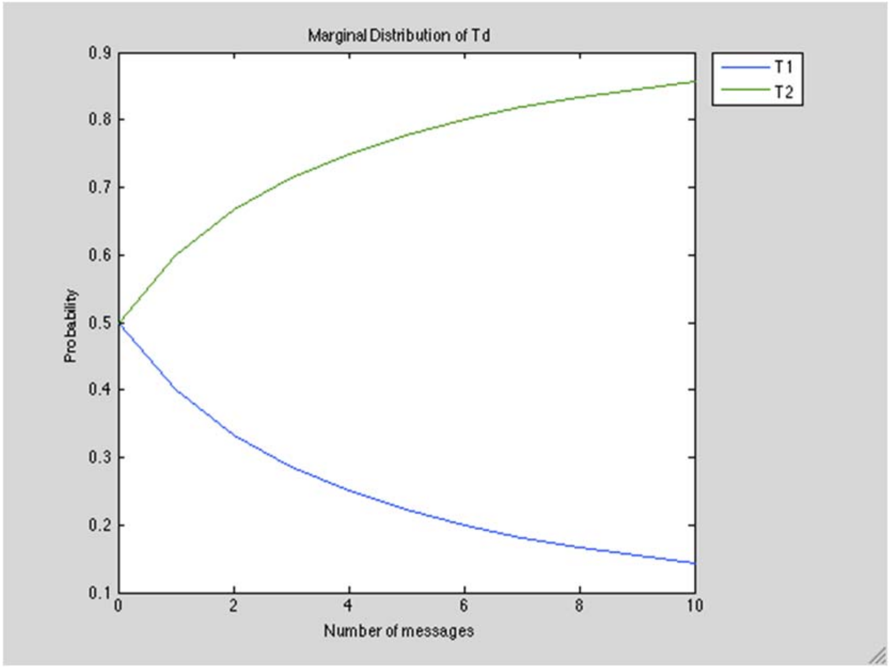

Figure 18. Marginal distribution of T_d.

Because the initial reliability is so low, we believe the informant is misleading us and thus the probability for T_1 decreases and the probability of T_2 increases.

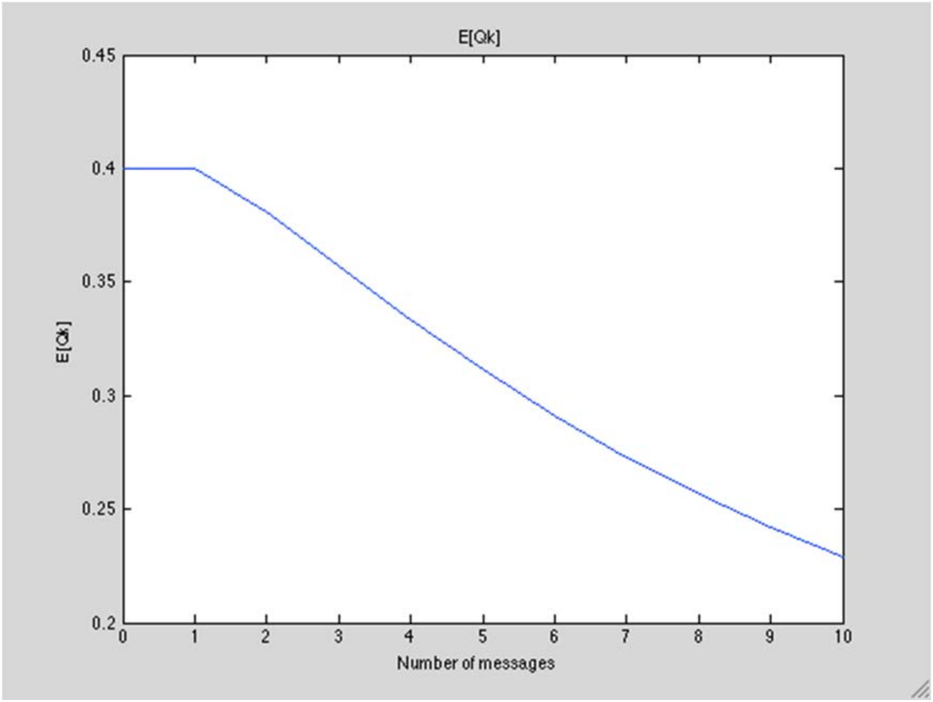

Figure 19. Mean of Q_k.

The mean of Q_k is decreasing with the number of messages. This result is also quite intuitive: since the probability of T_1 is decreasing, informant's repeated messages that include T_1 are considered less true and the estimated reliability of the informant is downgraded.

C. DISCUSSION

We have proposed a scheme for simultaneous update of the reliability and the estimation of the departure time. This scheme includes a mixed joint distribution of a discrete (T_d) and a continuous (Q_k) random variables that updates in a Bayesian fashion. As we saw in Chapter III, even if the assumptions are not fulfilled, the performance of the Bayesian method is satisfactory.

The scheme proposed makes use of conjugate Beta functions ensuring simple calculations that are easy to implement while maintaining flexibility to specify the mean reliability of an informant and the "strength" of our estimation regarding this reliability.

Although the reliability is unknown, and the true departure time is also unknown we can still estimate both of them and improve our estimate as we receive more intelligence. The scheme takes full use of the information provided by the informant to efficiently update the informant's reliability and the vessel's location simultaneously.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. SUMMARY, CONCLUSIONS, AND FUTURE WORK

A. DATA FUSION

In this thesis we formulate a model to assist the Joint Interagency Task Force South in its efforts to fight drug traffickers originating from South America. The main problem addressed in this thesis is how to combine different sources of intelligence into a coherent picture to effectively estimate the location of drug smugglers. In the initial model, we focus on determining the departure time of a smuggler. In later chapters we develop methods to estimate the route the smuggler travels, the vessel type, and velocity. The main contribution of this thesis is developing models to fuse information from two different types of intelligence sources, namely sensor-based sources and human-based sources, into a coherent intelligence picture. We update this picture as new information arrives.

The main model we explore is the Bayesian model, which is quite intuitive, mathematically rigorous and elegant. However, this method requires assumptions regarding the underlying probability distributions related to the intelligence gathered. Those assumptions are usually difficult to justify in practice since their validation requires gathering large amounts of data.

We compare the Bayesian model to a different type of intelligence updating mechanism, the Dempster-Shafer method. We examine several ways to implement Dempster-Shafer theory and compare those methods to Bayes' theory both qualitatively and quantitatively. The quantitative comparison is done using a simulation across multiple possible values of an informant's reliability and ways in which the messages are created.

We found that even when the assumptions of the Bayes' update process are violated, it still manages to yield the best results in the scenarios examined. It specifies the correct departure time a larger fraction of the time than the other methods. All the updating methods perform poorly when the reliability of the informant is low or is mistaken to be low, and there is non-uniformity in the way he produces messages.

B. UPDATING THE INFORMANTS RELIABILITY

A major contribution of this thesis is a Bayesian model that allows the operator to assess the reliability of the informant and update the vessels location simultaneously. We can do this when the informants' messages can be verified and when they cannot. The informant's reliability-departure time joint distribution model, described in detail in Chapter V, allows estimating both the location of the vessel and the reliability of the informant together and updating the estimate as more intelligence is received. Even though neither the true departure time of the vessel nor the reliability of the informant are known initially.

C. FUTURE WORK

This thesis suggests multiple models for updating the operator's perception as he receives more intelligence and sets a framework for the comparison and evaluation of those data fusion models. However, the research on the models developed can be extended in the following ways:

1. Extending the Model

In Chapter IV, multiple extensions to the basic model were suggested, but in order to encompass real-life situations, one may extend the model even further.

Possible extensions of interest are:

- Accounting for variable velocities of the vessels.
- Accounting for the case where the number of vessels in the theater of operations changes over time.
- Evaluating the probability that the informant delivers a message of a certain size l_k. In our analysis we assumed that this probability is known, but in fact it can be evaluated as more information arrives in a similar fashion to the one used in Chapter V to estimate the reliability q_k.

2. More Extensive Comparison

Although we have shown that the Bayes' model preforms best in the scenarios examined, the intelligence community may still benefit from more exhaustive comparison between the models.

A worthwhile direction to improve the comparison conducted in this thesis is by examining the models with other streams of messages, created in different ways than described in Chapter III. Examining the probability of specifying the correct departure time after receiving messages of different sizes from multiple informants with different reliabilities may also be of interest

Lastly, comparing the computational complexity of the update methods directly by computing the time required to perform the computations of different update methods.

3. Real Data

Inputting the models with real data may increase immensely the insights we can gain from the models and allow us to compare them more effectively. Such real data may relate to 1) prior knowledge about the vessels departure times, velocities and routes, 2) the characteristics of sensors, namely the false positive error P_{f+} and the false negative error P_{f-} and 3) the characteristics of the informants such as their reliability and most interestingly – the way in which they produce their messages, and what types of mistakes they tend to make.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Bayes, Rev. T. (1763). "An essay toward solving a problem in the doctrine of chances," *Phil. Trans. Roy. Soc*., 370–418.

Berger, J. O. (1993). *Statistical decision theory and bayesian analysis* (2nd ed.). New York: Springer-Verlag.

Cobb, R. B., & Shenoy P.P. (2003). "On Transforming Belief Function Models to Probability Models," University of Kansas School of Business Working Paper No. 293, July 2003, Revised February 2004, Lawrence, KS.

Chen, Z. (2003). *Bayesian filtering: From Kalman filters to particle filters and beyond.* Technical report, McMaster University.

Dempster, A. P. (1967). *Upper and lower probabilities induced by a multi-valued mapping.* Ft. Belvoir: Defense Technical Information Center.

Elmenreich W. (2002). *An introduction to sensor fusion.* Institut f¨ur Technische Informatik Vienna University of Technology, Austria

Hall, D. L., & Llinas J. (2001). *Handbook of Multisensor Data Fusion*. New York: CRC. Print.

JIATF-S (2013). Joint Interagency Task Force-South, Retrieved from http://www.jiatfs.southcom.mil/index.aspx

Koks, D., & Challa, S. (2003). *An introduction to bayesian and dempster-shafer data fusion* (No. DSTO-TR-1436). Defence Science and Tech Org.

Korb, K. B. & Nicholson, A. E. (2011). *Bayesian Artificial Intelligence* (2nd ed.). Boca Raton, FL: CRC Press.

Lincoln, T. L., & Parker, R. D. (1967). "Medical diagnosis using Bayes theorem." *Health Services Research, 2* (1), 34–45.

Martin, A. (2009). "Implementing general belief function framework with a practical codification for low complexity." In Smarandache, Florentin, *Advances and Applications of DSmT for Information Fusion Collected Works*. (pp. 217–273) United States: American Research Press.

Morelande, M. R., Kreucher, C. M., & Kastella, K. (2007). "*A Bayesian approach to multiple target detection and tracking," IEEE Transactions on Signal Processing, 55*(5).

NDIC (2011). The National Drug Intelligence Center. “National drug threat assessment 2011.” Retrieved from: http://www.justice.gov/archive/ndic /pubs44/44849/44849p.pdf

Palter, S. (2009, April, 6). Stratford Bungles Central American Drug Trade Analysis. Retrieved June 01 2013, from “Bellum - a project of the Stanford Review” Web Site: http://bellum.stanfordreview.org/?p=1114

Ratcliffe, J. (2000). “Implementing and integrating crime mapping into a police intelligence environment.” *International Journal of Police Science and Management*, *2*(4), 313–323.

Ross, S. M., & Ross, S. M. (2010). *Student's Manual To Accompany Introduction to Probability Models*, (10th ed.). Amsterdam: Elsevier/Academic Press.

Sentz K. (2002). *Combination of evidence in Dempster-Shafer theory.* Sandia Report SAND2002-0835.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press.

Smets, P. (1990).” Constructing the pignistic probability function in a context of uncertainty.” In R. D. Shachter, T. S. Levitt, J. F. Lemmer, & L. N. Kanal, *Uncertainty in Artificial Intelligence* (pp. 29–40). Amsterdam.

Smets, P. (2002). “Decision making in a context where uncertainty is represented by belief functions.” In R. P. Srivastava, & T. J. Mock, *Decisions, Belief Functions in Business* (pp. 316–332). Heidelberg: Springer.

Steinberg, N. A., Bowman, C. L. & White, F. E. (1999). “Revisions to the JDL data fusion model,” *Sensor Fusion: Architectures, Algorithms, and Applications III 3719*(1), 430–441

Sun Tzu (1971). *Art of war.* Oxford, UK: Oxford University Press.

United States Department of the Army. (2007). *Joint Intelligence: Field manual 2-0,* Washington, DC: Headquarters, Dept. of the Army.

White, F. E. (1991). *Data Fusion Lexicon*. Ft. Belvoir: Defense Technical Information Center.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California