DECISION MAKING IN TRACKING APPLICATIONS BY USING DEMPSTER-SHAFER THEORY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HASAN İHSAN TURHAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JULY 2014

Approval of the thesis:

DECISION MAKING IN TRACKING APPLICATIONS BY USING DEMPSTER-SHAFER THEORY

submitted by **HASAN İHSAN TURHAN** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Gönül Turhan Sayan
Head of Department, **Electrical and Electronics Engineering** ________________

Prof. Dr. Mübeccel Demirekler
Supervisor, **Electrical and Electronics Eng. Dept., METU** ________________

Examining Committee Members:

Prof. Dr. Mustafa Kuzuoğlu
Electrical and Electronics Engineering Department, METU ________________

Prof. Dr. Mübeccel Demirekler
Electrical and Electronics Engineering Department, METU ________________

Assoc. Prof. Dr. Umut Orguner
Electrical and Electronics Engineering Department, METU ________________

Assoc. Prof. Dr. Afşar Saranlı
Electrical and Electronics Engineering Department, METU ________________

Prof. Dr. Osman Parlaktuna
Electrical and Electronics Engineering Department, ESOGU ________________

Date: ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: HASAN İHSAN TURHAN

Signature :

ABSTRACT

DECISION MAKING IN TRACKING APPLICATIONS BY USING DEMPSTER-SHAFER THEORY

Turhan, Hasan İhsan
M.S., Department of Electrical and Electronics Engineering
Supervisor : Prof. Dr. Mübeccel Demirekler

July 2014, 107 pages

The aim of this thesis is to study attribute data fusion and decision making for targets tracked by a sensor network consisting of several radars. As an application deciding both target class and identity are studied. Since only partial information is available, Dempster-Shafer theory is used for this application to assign and combine probability masses. In this study, we focus on the problems of basic probability assignment and decision/data fusion.

Classification of air vehicles according to their type is studied using the kinematic features obtained while tracking. The probability masses are obtained from tracker data and prior information that belong to possible target types. Prior information is modeled as a Gaussian mixture probability density function, while tracker data is modeled as a single Gaussian. This new methodology is tested with real data and its performance is examined by comparing it with the most similar method existing in the literature.

Special to this type of air vehicle classification problem, a decision fusion approach is proposed that uses Bayesian formalism. The main difference of the proposed methodology from the existing methods is fusing the data before assigning the basic probabilities. Methodology is tested with real data and compared with the existing combination rules in the literature.

Target identification is the decision of whether a target is a friend, hostile or neutral. This decision is made by using IFF Mod-4 information, IFF Mod-3 information, restricted area breach information, air corridor usage information and human-eye identification information. These piece of information are converted into probability masses and combined by using Analytic Hierarchy Process Interrogation methods and Dempster-Shafer Theory. Methodology is tested by using artificial scenarios.

Keywords: Dempster-Shafer Theory, Belief Functions, Basic Probability (Mass) Assignment, Dempster-Shafer Reasoning, Conditioning, Combination Rules, Target Tracking, Decision Making, Target Classification, Target Identification, Analytic Hierarcy Process, group aggregation methods, pair-wise comparison, inconsistency check

ÖZ

HEDEF TAKİP UYGULAMALARINDA DEMPSTER-SHAFER TEORİSİ KULLANARAK KARAR VERME

Turhan, Hasan İhsan
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi : Prof. Dr. Mübeccel Demirekler

Temmuz 2014 , 107 sayfa

Bu tez çalışmasının amacı çeşitli radarların bağlı olduğu füzyon sistemlerindeki hedef takip uygulamalarında veri füzyonu ve karar vermedir. Uygulama olarak hedefin tipine ve niteliğine karar verilmesi gerçekleştirilmiştir. Hedefin tipinin ve niteliğinin belirlenmesinde, literatürde Demspter-Shafer Teorisi olarak bilinen inanç fonksiyonları kullanılır. Dempster-Shafer Teorisi olaylara olasılıklar atayabilmek için yeterli istatistiğin olmadığı durumlarda eldeki veriler doğrultusunda olasılık ağırlıklarının atanması ve sonrasında da gelen tüm bilgilerin birleştirilmesi esasına göre uygulanır. Bu tez çalışmasında olasılık ağırlık ataması ve karar verme/birleştirme üzerinde yoğunlaşılmıştır.

Hedef tipi tespiti izleme sırasında elde edilen kinematik veriler kullanılarak yapılır. Olasılık ağırlık ataması, izleyici çıktısı ve olası hedef tiplerine ait öncül (prior) bilgilerin birlikte kullanılması yoluyla elde edilmektedir. Öncül bilgi karma Gauss dağılımlı olasılık yoğunluk işlevleri, izleyici çıktısı olan ölçümler ise Gauss dağılımlı olasılık yoğunluk işlevleri olarak modellenir. Yöntem gerçek veriler kullanılarak test edilmiş ve literatürde önerilen yönteme en yakın yöntemler karşılaştırmak suretiyle performansı irdelenmiştir.

Hedef tipi tespiti konusunda Bayes yaklaşımından yararlanan, bu probleme özgü bir de birleştirme yöntemi önerilmiştir. Yöntemin literatürdeki diğer yöntemlerden te-

mel farkı olasılık ağırlıklarını atamadan önce gelen verileri birleştirmesidir. Yöntem gerçek veriler kullanılarak test edilmiş ve literatürdeki birleştirilme kurallarıyla karşılaştırılmak suretiyle irdelenmiştir.

Hedefin niteliğine karar verilmesi, hedefin dost, düşman ya da tarafsız olduğunun belirlenmesidir. Bu karar, IFF Mod-4, IFF Mod-3, yasaklı bölge ihlali bilgisi, hava koridoru kullanım bilgisi ve görsel teşhis bilgisi kullanılarak verilir. Bu kaynaklardan gelen bilgiler, çoklu karar verme yöntemlerinden biri olan Analitik Hiyerarşi Süreci (Analytic Hierarcy Process) sorgulama yöntemleri ve Dempster-Shafer Teorisi kullanılarak olasılık ağırlıklarına çevrilir ve birleştirilir. Yöntem yapay veriler kullanılarak test edilmiştir.

Anahtar Kelimeler: Dempster-Shafer Teorisi İnanç Fonksiyonları, Temel Olasılık (Ağırlık) Ataması, Dempster-Shafer Tümevarımı, Koşullandırma, Birleştirme Kuralları, Hedef Takibi, Karar Verme, Hedef Tipi Belirleme, Hedef Teşhisi Belirleme,Analitik Hierarşi Süreci, grup karar birleştirme yöntemleri, ikili karşılaştırma, tutarsızlık tespiti

To my family and fiancée!!

ACKNOWLEDGMENTS

I want to express my sincere gratitudes to my thesis supervisor Prof. Dr. Mübeccel Demirekler. This thesis work cannot be completed without her great encouragement and support. She was more than my supervisor during my thesis work. There is no measure for her precious mentoring.

I want to thank Yalçın Demirekler for his patience, support and joyful co-supervising during my thesis work. I am very pleased to meet such an optimistic and full of life person.

I want to thank ASELSAN A.Ş. for providing related information and real data sets to me to test the algorithms. Also I want to thank Prof. Dr. Mustafa Kuzuoğlu, Assoc. Prof. Dr. Umut Orguner and Sensor Fusion Project team for their great support during the project and this thesis work.

I want to thank Dr. Sedat Doğru for his valuable encouragement and support during all my master education.

Also I want to thank my family for their endless support and faith in me for my all life.

Finally, I want to thank my fiancée Gülce Ergün for her great encouragement and support. She did not lose her faith in me, even when there was nothing to do.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Decision making is a difficult problem for all applications. Decision making has to somehow map and fuse information that comes from different sources and at different time instances. There are several theories for mapping and data fusion: probability theory, possibility theory, and Dempster-Shafer theory. Every theory has its own advantages and disadvantages that can change according the availability of the information.

Major decision making problems in target tracking are:

- Measurement to track association
- Track to track association
- Target classification
- Target identification
- Model selection for robust tracking performance

Target classification and target identification are the main decision making problems in target tracking, because all other items benefit from classification and identification results. Hence, this thesis work focuses on target classification and target identification.

1.1 Target Classification

Target classification is an important problem, which is encountered in designing an efficient air defense system. Basically, targets are classified among predefined target types like helicopter, fighter, etc. Features used in classification are selected according to their discriminating powers and availability. In this study, the available information comes from a target tracker that tracks air vehicles. Tracker provides rich information about the state of the target, which is composed of velocity and the position vectors, as a Gaussian density at discrete time instants [5].

Classification of a target is generally done using kinematic features [3, 9, 11, 28, 30, 31]. Besides kinematic features, radar cross section or any other relevant information like electronic support measures [9, 11, 31] can also be used. Smets et al. [41] uses IMM mode probabilities beside the kinematic information. Caramicoli et al. [13] and Ristic and Smets [31] use kinematic features to derive some classification rules.

Most of the related work that exists in the literature uses Dempster-Shafer theory. However, some authors use some other statistical methods: Angelova et al. [3] use particle filter and mixture Kalman filter, Ristic et al. [30] use Bayesian match-filter, Mei et al. [28] and Brooks et al. [12] use Bayesian classifier, Azimi et al. [4] use wavelets and neural networks.

We also use kinematic information in our study and we develop a new methodology for assigning masses to classes in the Dempster-Shafer framework. The main distinction of the mass assignment proposed in this work is all the information, i.e., the probability density function of the state instead of its mean, provided by the tracker is utilized. The proposed algorithms assume that the prior probability density functions of all classes and the current measurement are known for some kinematic features of the target. Furthermore, as a sensible assumption, we assume that the prior probability density functions of the related kinematic features of all classes can be approximated by Gaussian mixtures. Assigning masses by using both prior and measurement probability density functions in Dempster-Shafer framework is one of the main contributions of this study.

1.2 Target Identification

Target identification is another important problem which is encountered when designing an efficient air defense system. Target identification can be considered as a classification problem and classes are defined as friend, foe (hostile) and neutral.

Generally interrogation of friend or foe (IFF) is used for target identification. But other information sources can also be used for this purpose. For example some countries use discriminating devices like electronic support measures that can identify the radar type from its signal. Hence target can be identified as friend or foe by help of this information. Other possible sources of information that give clues on the identity of the target are restricted area breach and air corridor usage.

It is difficult to map such all/nothing information into numbers for data fusion. Therefore Dempster-Shafer Theory is widely used for target identification. Bogler [11], Caramicoli et al. [13] and Ristic and Smets [31] use radar cross section, IFF, and electronic support measures. None of them use restricted area breach or air corridor usage for this purpose. Probability masses are assigned almost intuitively. We use the same information sources given in the literature, but contrary to the works in the literature we don't assign the probability masses either intuitively or coarsely. Instead, we develop a new methodology that uses Analytic Hierarchy Process and Dempster-Shafer Theory together for assigning masses to classes (identities).

1.3 Motivation

Target classification and identification provide very valuable information in the battle field, because they can be used for weapon engagement. Furthermore, class and identity information improve the performance of target trackers. With the help of class and identity information:

- More accurate models can be chosen for Interacting Multiple Model (IMM) filters.
- Multiple measurements can be assigned to related tracks in a more correct way.

- Tracks that belong to the same target can be associated in distributed multi-source target tracking systems.

The reasons motivating us to utilize Dempster-Shafer theory can be listed as follows:

- Easy way of representing uncertainty
- Availability of a conflict measure
- Suitable mass assignment and combination methods

Finally, our motivation was to develop algorithms that use all available information in order to provide more correct mapping for attribute data fusion.

CHAPTER 2

DEMPSTER-SHAFER THEORY

Belief Functions were introduced by Arthur Dempster in 1967 as "upper and lower probabilities"[16]. Later it was developed as a mathematical theory of evidence in 1976 by Glenn Shafer[38]. Today it is mostly known as Dempster-Shafer Theory. The theory starts with unknown statistics or evidence of events, and aims to reach a decision using these unknown and incomplete statistics.

The first step of making a decision is to convert the evidence into numbers. The second step is combining the evidence expressed as those numbers. From this aspect the theory reminds us probability theory. However, it differs from probability theory in the way of mapping evidence into the interval $[0,1]$, combining the evidence, and its ability to represent uncertainty.

One important difference between the two theories is the representation of uncertainty. Probability theory represents complete loss of information by assigning equal probability to each event. However, Dempster-Shafer theory assigns "1" to the universal set and "0" to all other sets.

Another fundemental difference is the probability assignment. In probability theory, probability function that maps the field elements into the interval $[0,1]$ has certain properties. One of the properties is the probability of the set that is the union of two disjoint sets is equal to sum of their probabilities. However, union sets represent the local uncertainty in Demspter-Shafer theory and masses/probabilities of these union sets are assigned in a different manner. Even if a set consists of two disjoint sets, mass/probability of that set is not equal to sum of their probabilities.

Besides these differences Demspter-Shafer theory has one more important difference in combination of evidence. It has its own combination methods as well as own probability assignment methods. Dempster-Shafer theory assumes that the available information may be inconsistent, and distinct from probability theory, introduces a new concept which is called conflict. Conflict arises when different information sources assign different probability masses to the events.

In this chapter, Dempster-Shafer Theory is introduced with its basics, and probability assignment and combination (conditioning) methods.

2.1 The Basics of the Dempster-Shafer Theory

Suppose that Θ is a finite set, and let 2^Θ denote all subsets of Θ.

$$\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$$

$$2^\Theta = \{\phi, \{\theta_1\}, \{\theta_2\}, ...\{\theta_n\}, \{\theta_1, \theta_2\}, ..., \{\theta_1, \theta_2, ..., \theta_n\}\}$$

where ϕ denotes the empty set.

Belief functions are constructed over basic probability assignments (bpa). Basic probabilities, in other words probability masses, map the evidence into $[0, 1]$ interval. Basic probabilities are defined by using partial information that is available and assigned to the subsets of Θ. They are defined as follows:

$$m : 2^\Theta \rightarrow [0, 1] \tag{2.1}$$

The function m satisfies the following conditions:
(1) $m(\phi) = 0$
(2) $\sum\limits_{X \in 2^\Theta} m(X) = 1$

The belief function is defined over Θ as follows:

$$Bel : 2^\Theta \rightarrow [0, 1] \tag{2.2}$$

The function Bel satisfies the following conditions:

(1) $Bel(\phi) = 0$

(2) $Bel(\Theta) = 1$

(3) $Bel(A) = \sum\limits_{X \subseteq A} m(X)$

The theory defines belief functions as "lower probabilities". Besides belief functions, there are plausibility functions that can be defined as "upper probabilities"[16]. Plausibility function is also constructed using basic probability assignments. The construction is as follows:

$$Pl : 2^{\Theta} \rightarrow [0, 1] \tag{2.3}$$

The function Pl satisfies the following conditions:

(1) $Pl(\phi) = 0$

(2) $Pl(\Theta) = 1$

(3) $Pl(A) = \sum\limits_{X \cap A \neq \phi} m(X)$

(4) $Pl(A) = 1 - Bel(\bar{A})$

2.2 Basic Probability Assignment

Basic probability assignment is one of the main steps in any application of the Dempster-Shafer theory. The methods of basic probability assignment in the literature are numerous. At the early stages of the theory, probability masses are assigned according to the expert opinion in various applications. Afterwards, various models are used for this purpose. Yager [44] used belief functions as a fuzzy measure. Zhu et al. [47] used membership functions. Florea et al. [20] used membership values as probability masses. Römer et al. [32] used possibility and necessity measures of fuzzy logic theory for defining belief functions. Bloch [10], Jiang et al. [25], and Masson et al. [27] used distances to cluster centers for basic probability assignment. Utkin [15] used imprecise Dirichlet model. Bendjebbour et al. [8], Hagarat-Mascle et al. [23], Salzestein et al. [36], and Xu et al. [42] used probabilities for basic probability assignment.

In the remaining part of this section we classify the existing methods and give brief descriptions for each of them.

Clustering Based Basic Probability Assignment

This method is inspired from clustering algorithms. The basic idea behind clustering is grouping similar data for classification. Clustering algorithms can be mainly classified into two families: hierarchical partitioning and hard/fuzzy partitioning [27].

There are two types of hierarchical clustering methods: agglomerative and divisive. Agglomerative methods take each element of the data set as a cluster and merge similar ones by using a similarity measure. Divisive methods take the whole data set as one cluster, and divide it by using a dissimilarity measure between its elements.

Hard/fuzzy partitioning determines an initial cluster and modifies it by increasing the inter distance of clusters and decreasing the intra distance of a cluster. The distance is measured by some appropriate objective function.

In hard partitioning clusters are generated in such a way that each data point belongs to only own cluster. However, in fuzzy partitioning, an element can belong to more than one cluster with a different membership degree where the sum of the membership degrees is 1.

The evidential partitioning algorithm arises from fuzzy partitioning. Basic probability assignment can be made by using distance to cluster centers [27]. Distance measures can be different from algorithm to algorithm [10, 25] . The steps of clustering based basic probability assignment are as follows:

Training:

- Generate the clusters
- Find cluster centers by using an objective function. Objective function uses two main metrics: distances between clusters and distances between cluster elements within the same cluster.

Probability mass assignment:

- Find the distance of new inputs (possibly a measurement) to each cluster center.
- Use a function of the computed distance to assign masses to each cluster.

- Finally apply normalization.

Fuzzy Logic Based Basic Probability Assignment

Ordinary set theory claims that an element can belong to only one set; but, fuzzy set theory claims that a member of one set can also be a member of another set. Here membership function definition shows up.

Membership function is a continuous mapping, μ : feature set $\rightarrow [0, 1]$, which maps the data points to the closed interval $[0, 1]$. The value $\mu(x)$ shows the membership degree.

Like other theories, fuzzy set theory is also adapted to the Dempster-Shafer theory. The related part of fuzzy logic with belief functions is the membership functions. Membership functions are obtained by using logic-based rules. Mass assignment is done according to the membership values [47].

Probability Density Function Based Basic Probability Assignment

Mass assignment from a probability density function (pdf) is a relatively new concept. This concept can be used for systems that we have prior data. Prior data is used to generate the prior probability density functions of the events. Xu et al. [42] produce Gaussian densities using the training data and propose a new methodology for basic probability assignment. We explain the method of Xu et al. [42] on a generic example of $\Theta = \{A, B, C\}$. The probability density functions of the three classes are shown in Figure 2.1 for the feature 'x'.

It is assumed that a result of an experiment gives x as $x = 850$. 'x' is shown in Figure 2.1 by black dotted line. The likelihood values of each class is read from the prior probability density functions as shown in the Figure 2.1. Note that $p_A(850) > p_B(850) > p_C(850)$.

$p_A(850)$ is assigned to the set $\{A\}$ directly. So, $m(\{A\}) \propto p_A(850)$. $p_B(850)$ is assigned to the set $\{A, B\}$, since it is also under the pdf curve of A. So, the decision

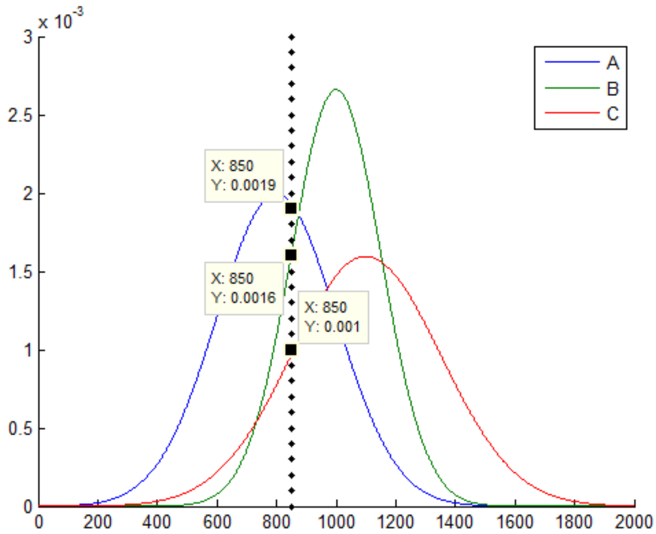


Figure 2.1: Probability density functions

is: it cannot be determined whether it is A or B, so $m(\{A, B\}) \propto p_B(850)$. $p_C(850)$ is assigned to the set $\{A, B, C\}$, since it is both under the pdf curve of A and B. So, the decision is: it cannot be determined whether it is A, B or C, so $m(\{A, B, C\}) \propto p_V(850)$. Finally all three values are normalized, so that their some is 1.

2.3 Combination of Evidence

In this section we give a brief summary of combination rules that exist in the literature. At the very beginning of this theory, Dempster combination rule has been used for combining evidence. But, later it is criticized about not giving reasonable results in case of conflict. This fact is first discovered and criticized by Zadeh [45]. This drawback has been tried to be solved by assigning some probability masses to the uncertainty [38]. However in applications it is observed that probability masses assigned to uncertainty drop quickly to zero after few combinations. Some effort is spent to overcome this problem and a set of combination rules are proposed. Before explaining these combination rules, we need to make the following definitions.

Conflict: Conflict means opposite opinions of decision makers on an event. If we examine conflict on the basis of the universal set and its subsets, we get two definitions for conflict, namely partial conflict and total conflict.

Let $m_1(.)$ and $m_2(.)$ be two probability mass function defined over 2^Θ.

Partial conflict: Partial conflict is the conflict among two subsets of Θ, and is defined as the multiplication of the masses assigned to the two (or more) sets that have empty intersection. The partial conflict among $m_1(.)$ and $m_2(.)$ for the sets A and B is defined as

$$k^{AB}{}_{12} = m_1(A).m_2(B), \text{ where } A, B \subseteq 2^\Theta \text{ and } A \cap B = \phi$$

Total conflict: Total conflict is equal to total probability masses that correspond to the empty set. The total conflict among $m_1(.)$ and $m_2(.)$ is defined as the sum of all partial conflicting masses.

$$k_{12} = \sum_{\substack{X_1, X_2 \subseteq 2^\Theta \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2)$$

VBA (Vacuous Belief Assignment): This occurs at total ignorance condition. This type of bpa does not contain any information. It gives all probability masses to the universal set and this means that the decision maker has total ignorance on the subject.

$$m_v(\phi) = 0$$
$$m_v(\Theta) = 1$$

Combination of two decisions can be considered as an operator acting on two basic probability assignment functions and generating a new one. Combination operator is denoted by the symbol $\oplus$. While examining the performance of different combination techniques we need the following definitions.

Commutativity: An operator is said to be commutative if,

$$m_1 \oplus m_2 = m_2 \oplus m_1$$

Associativity: An operator is said to be associative if,

$$(m_1 \oplus m_2) \oplus m_3 = m_1 \oplus (m_2 \oplus m_3)$$

Neutral Impact of VBA: As defined earlier VBA occurs at total ignorance condition. Therefore the VBA m_v should not affect the result.

$$(m_s \oplus m_v) = m_s$$

Coherence of the Combination Rules in all Possible Cases: Combination Rules should give reasonable results for any number of sources, any values of basic probability assignments (bpa) and any types of frames and models which may change or stay invariant over time. This rule is not precisely defined and it is somehow subjective. Decision on the satisfaction of the rule usually depends on expert opinion.

Reliability: Information sources can give false or imprecise information because of quality or trustability of the sources. So reliability of a source should be taken into account in combination rules.

Dempster's Combination Rule

Historically Dempster's combination rule is the first rule ever proposed [38]. All others are derived to overcome its shortages. Equation 2.4 gives the rule. In this equation the probability masses assigned by the first expert are denoted by the subscript 1 and the second expert by the subscript 2.

$$m_{DS}(\phi) = 0 \tag{2.4a}$$

$$m_{DS}(X) = \frac{\sum\limits_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1).m_2(X_2)}{1 - \sum\limits_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2)} \tag{2.4b}$$

✓ It satisfies the Commutativity property.

✓ It satisfies the Associativity property.

✓ It satisfies Neutral Impact of VBA.

× It is not coherent to all possible cases. (Zadeh's example [45])

× Reliability of the sources is not taken into consideration.

Example:

Let $m_1(.)$ and $m_2(.)$ be two probability mass function defined for $\Theta = \{\theta_1, \theta_2, \theta_3\}$.

$m_1(\{\theta_1\}) = 0.2$

$m_1(\{\theta_3\}) = 0.4$

$m_1(\{\theta_1, \theta_3\}) = 0.2$

$m_2(\{\theta_1\}) = 0.5$

$m_2(\{\theta_2\}) = 0.3$

$m_2(\{\theta_1, \theta_2\}) = 0.2$

Table 2.1: Unnormalized results of Dempster's combination rule

	$m_1(\{\theta_1\}) = 0.2$	$m_1(\{\theta_3\}) = 0.4$	$m_1(\{\theta_1, \theta_3\}) = 0.2$
$m_2(\{\theta_1\}) = 0.5$	$m_{12}(\{\theta_1\}) = 0.1$	$m_{12}(\{\phi\}) = 0.06$	$m_{12}(\{\theta_1\}) = 0.02$
$m_2(\{\theta_2\}) = 0.3$	$m_{12}(\{\phi\}) = 0.2$	$m_{12}(\{\phi\}) = 0.12$	$m_{12}(\{\phi\}) = 0.08$
$m_2(\{\theta_1, \theta_2\}) = 0.2$	$m_{12}(\{\theta_1\}) = 0.1$	$m_{12}(\{\phi\}) = 0.06$	$m_{12}(\{\theta_1\}) = 0.04$

Table 2.1 illustrates the un-normalized masses assigned by the Dempster combination rule. Each cell of this table, which contains a mass corresponding to the empty set, shows the partial conflict among the corresponding sets (events) and the total conflicting mass which is calculated as:

$$k_{12} = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2) = 0.52$$

The normalized masses, i.e., the result of combination, are obtained by normalizing the above given result. For this example the combination gives mass values as follows:

$m_{DS}(\{\theta_1\}) = 1$

Yager's Combination Rule

In historical perspective, the first rule after Dempster's combination rule is Yager's rule. It was proposed in 1987 by Yager. Yager's rule assigns the conflicting masses to the universal set [43].

$$m_Y(\phi) = 0 \tag{2.5a}$$

$$m_Y(X) = \sum_{\substack{X_1,X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1).m_2(X_2) \tag{2.5b}$$

$$m_Y(\Theta) = m_1(\Theta).m_2(\Theta) + \sum_{\substack{X_1,X_2 \in 2^\Theta \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2) \tag{2.5c}$$

✓ It satisfies the Commutativity property.

× It does satisfy the Associativity property.

✓ It satisfies Neutral Impact of VBA.

× It is not coherent to all possible cases. (Partial conflicts are not taken into consideration.)

× Reliability of the sources is not taken into consideration.

Dubois' and Parade's Combination Rule

After Yager, Dubois and Parade introduced a new combination rule in 1988. The new combination rule deals with the conflicting cases in a different manner. It assigns mass of conflict to the sets that are the union of conflicting sets [18].

$$m_{DB}(\phi) = 0 \tag{2.6a}$$

$$m_{DB}(X) = \sum_{\substack{X_1,X_2 \in 2^\Theta \\ X_1 \cap X_2 = X \\ X_1 \cap X_2 \neq \phi}} m_1(X_1).m_2(X_2) + \sum_{\substack{X_1,X_2 \in 2^\Theta \\ X_1 \cup X_2 = X \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2) \tag{2.6b}$$

✓ It satisfies the Commutativity property.

× It does not satisfy the Associativity property.

✓ It satisfies Neutral Impact of VBA.

× It is not coherent to all possible cases.

× Reliability of the sources is not taken into consideration.

Smets' Combination Rule

In 1994, Smets proposed the Transferable Belief Model. With this model conflicting masses are transferred to the empty set. Decisions based on the result of the combination are made among remaining alternatives by using the Pignistic Transformation. This modeling assumes open world and speculates that universal set is not really universal and there are possibilities other than the set 2^{Θ} [40].

$$m_S(\phi) = \sum_{\substack{X_1, X_2 \in 2^{\Theta} \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2) \tag{2.7a}$$

$$m_S(X) = \sum_{\substack{X_1, X_2 \in 2^{\Theta} \\ X_1 \cap X_2 = X}} m_1(X_1).m_2(X_2) \tag{2.7b}$$

✓ It satisfies the Commutativity property.

✓ It satisfies the Associativity property.

✓ It satisfies Neutral Impact of VBA.

✓ Open world assumption is done and this provides coherency to all possible cases.

× Reliability of the sources is not taken into consideration.

The transferable belief model concept uses pignistic probability while making decisions among combined evidences. Pignistic probability just a classical probability measure and it is derived from the mass function. It is denoted by $BetP$ and is defined as follows [40];

$$BetP(A) = \sum_{B \in 2^{\Theta}} m(B) \frac{|A \cap B|}{|B|} \tag{2.8}$$

Here, $|X|$ denotes cardinality of set X.

Weighted Average Operator

In 2003, Jøsang, Daniel and Vannoorenberghe proposed a new method which redistributes the conflicting masses to all sets in the core. The distribution is proportional to the masses of these sets [26].

$$m_{WAO}(\phi) = 0 \tag{2.9a}$$

$$m_{WAO}(X) = \sum_{\substack{X_1,X_2 \in 2^{\Theta} \\ X_1 \cap X_2 = X}} m_1(X_1).m_2(X_2) + w(x) \sum_{\substack{X_1,X_2 \in 2^{\Theta} \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2) \tag{2.9b}$$

$$w(x) = \frac{1}{2}(m_1(X) + m_2(X)) \tag{2.9c}$$

✓ It satisfies the Commutativity property.

✓ It satisfies the Associativity property.

× Conflicting masses are redistributed to all sets. This removes Neutral Impact of VBA.

× Redistribution of the conflict to all sets also removes coherency to all possible cases.

× Reliability of the sources is not taken into consideration.

Daniel's minC Rule

Again in 2003, Daniel proposed to redistribute the total conflicting masses to all subsets of Θ without considering whether they cause conflict or not, in order to retain minimum conflict [15].

✓ It satisfies Commutativity property.

✓ It satisfies Associativity property.

× Conflicting masses are redistributed to all sets. This removes both Neutral Impact of VBA

× Redistribution of the conflict to all sets also removes coherency to all possible cases.

× Reliability of the sources is not taken into consideration.

Principal Conflict Redistribution Rule

In 2004, Smarandache and Dezert examined all these rules and revealed their drawbacks. Then, they proposed a new partial conflict redistribution (PCR) rule [39].

$$m_{PCR1}(\phi) = 0 \tag{2.10a}$$

$$m_{PCR1}(X) = \sum_{\substack{X_1,X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1).m_2(X_2) + \frac{c_{12}(X)}{d_{12}} \sum_{\substack{X_1,X_2 \in 2^\Theta \\ X_1 \cap X_2 = \phi}} m_1(X_1).m_2(X_2) \tag{2.10b}$$

$$c_{12}(X) = m_1(X) + m_2(X) \tag{2.10c}$$

$$d_{12} = \sum_{X \in 2^\Theta} c_{12}(X) \tag{2.10d}$$

✓ It satisfies Commutativity property.

✓ It satisfies Associativity property.

× Conflicting masses are redistributed to all sets. This removes Neutral Impact of VBA

× Redistribution of the conflict to all sets also removes coherency to all possible cases.

× Reliability of the sources is not taken into consideration.

Later in 2006, they proposed 5 new PCR rules, which also provide Neutral Impact of VBA and coherency to all possible cases [17]. Only the last one is given here.

$$m_{PCR6}(\phi) = 0 \tag{2.11a}$$

$$\begin{aligned} m_{PCR6}(X) = &\sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2=X}} m_1(X_1).m_2(X_2) \\ &+ \sum_{\substack{Y\in 2^\Theta \\ c(X_1\cap X_2)=\phi}} \frac{m_1(X)^2.m_2(Y)}{m_1(X)+m_2(Y)} + \frac{m_2(X)^2.m_1(Y)}{m_2(X)+m_1(Y)} \end{aligned} \tag{2.11b}$$

✓ It satisfies Commutativity property.

✓ It satisfies Associativity property.

✓ It provides Neutral Impact of VBA. Because conflicting masses are redistributed to only conflicting sets.

✓ It provides coherency to all possible cases because of the same reason.

× Reliability of the sources is not taken into consideration.

Florea's Combination Rule

In 2009, Florea criticized all proposed combination rules about not considering reliability of sources. Then, he proposed a new method, which takes into account reliability of sources [21].

$$m_{Flo}(\phi) = 0 \tag{2.12a}$$

$$\begin{aligned} m_{Flo}(X) = &\frac{1}{1-k-k^2} \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2=X}} m_1(X_1).m_2(X_2) \\ &+ \frac{k}{1-k-k^2} \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cup X_2=X}} m_1(X_1).m_2(X_2) \end{aligned} \tag{2.12b}$$

$k = 0 \rightarrow$ Both sources are coherent and reliable, Conjunctive Combination Rule.

$k = 1 \rightarrow$ Total Conflict. Disjunctive Combination Rule.

$k = 0.5 \rightarrow$ Unknown Case.

- ✓ It satisfies Commutativity property.
- × It does not satisfy Associativity property.
- ✓ It provides Neutral Impact of VBA.
- ✓ Reliability of sources is taken into consideration.
- × But, reliability cannot be known for all cases. Therefore coherency to all possible cases is an open problem.

CHAPTER 3

ANALYTIC HIERARCHY PROCESS

The analytic hierarchy process (AHP) is a very popular method for the multi-criteria decision making problem. AHP was introduced by Saaty in 1980 [37]. AHP is suitable for both individual and group decision making. Group decision making suitability was studied by Dyer and Foreman in 1992 [22]. The main idea behind AHP is defining a common hierarchy of criteria and pair-wise comparison of those criteria. The pair-wise comparison provides ranking of each criterion for selecting the optimum.

AHP is a powerful tool for multi-criteria decision making; however, AHP has some problems like determining the weights of the criteria and aggregating those weights on group decisions. In 1983 Belton and Gear examined the Saaty's method and criticized its weight determination [7]. According Belton and Gear if one alternative is taken out, the order of the other alternatives changes, in other words the best alternative is not the best alternative anymore. Dyer also criticized Saaty's method in 1990 for the same reason [19]. Saaty replied to the criticism in 1983 [35] and 1990 [34]; but these replies were not enough.

Another criticism to AHP is about group decision. AHP assumes that groups are homogeneous. According to Zahir this assumption does not hold in large groups [46]. Large groups do not lead to a unique decision among available alternatives. Even if they are allowed to share their ideas with each other as expressed in [46], some cluster of decision ranking forms. People in large groups split into subgroups while making a decision and this situation is the second problem in AHP.

One solution can be consensus voting that is proposed by Saaty [33]. In this approach, the problem should be well defined and understood clearly from the decision makers. However, this takes a long time. Even if groups are not homogeneous, individual judgments are combined somehow in group decision making problems.

In 1983, Aczel and Saaty proposed the geometric mean method that gives same importance to the judgments of all group members [1]. After that Basak and Saaty proposed a weighted geometric mean method giving much attention to some dominant group members in 1993 [6]. In 1994, Ramanathan and Ganesh evaluated the geometric mean method and the weighted arithmetic mean method. They criticized the geometric mean method about not satisfying the "Pareto optimality axiom" [29]. Pareto optimality axiom says that a solution is the optimum point if it is impossible to make any improvement without making any regression. On the other hand arithmetic mean method satisfies social axioms.

3.1 Hierarchical Model on Decision

The goal of the AHP is to choose the best alternative. For this purpose the first step of the AHP is building a model for hierarchy, in order to decide as shown in Figure 3.1

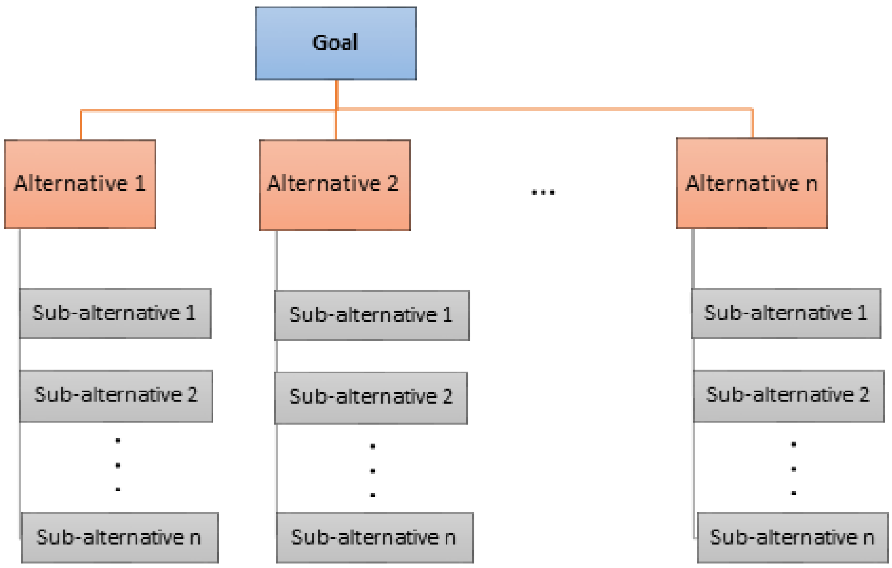


Figure 3.1: Hierarchical Model for AHP

For example the goal may be selecting a new car and the alternatives for selecting a car may be fuel consumption, price, size, etc. Sub-alternatives may be car models.

3.2 Pair-wise Comparison

The key idea of AHP for choosing the best alternative is pair-wise comparison. All alternatives are evaluated in some hierarchy level according to Table 3.1.

Table 3.1: Evaluation table for AHP

Importance Level	Explanation
1	Equally important
2	
3	More important
4	
5	Much more important
6	
7	Very much more important
8	
9	Absolutely important

Every importance level shows how much the first alternative is important than the second alternative. 2, 4, 6, 8 are the intermediate values. One can write $\frac{1}{3}$, if one thinks that the second alternative is more important than the first alternative. As an example by using Table 3.1 pair-wise comparison matrices are built up as given in Table 3.2.

Table 3.2: Example pair-wise comparison matrix A

	Alternative 1	Alternative 2	...	Alternative n
Alternative 1	1	$\frac{1}{3}$	...	4
Alternative 2	3	1	...	6
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Alternative n	$\frac{1}{4}$	$\frac{1}{6}$	...	1

3.2.1 Pair-wise Comparison Matrix Evaluation

Once the pair-wise comparison matrix, i.e. the matrix A, is obtained, weights of the alternatives are obtained. The following example is given to illustrate the methods of computation of weights.

Example: Suppose that matrix A is

$$A = \begin{bmatrix} 1 & 2 & \frac{1}{2} \\ \frac{1}{2} & 1 & 2 \\ 2 & \frac{1}{2} & 1 \end{bmatrix}$$

The rows of this matrix should be the ratio of the weights as

$$\tilde{A} = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \frac{w_1}{w_3} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \frac{w_2}{w_3} \\ \frac{w_3}{w_1} & \frac{w_2}{w_2} & \frac{w_3}{w_3} \end{bmatrix}$$

Observation 1: For this example it is not possible to find a set w_1, w_2, w_3 that gives the matrix A. That is obvious from the interpretation of the first and the second rows of A. First row indicates that $w_3 < w_1 < w_2$. On the contrary the second row indicates that $w_1 < w_2 < w_3$. The contradiction between the two rows indicates the non-existence of a solution. Any matrix that has rank greater than 1 has no solution.

Observation 2: The rank of the matrix $\tilde{A}$ is 1 since all rows are proportional to each other, so its 2 out of 3 eigenvalues are 0. The remaining eigenvalue is 3, and its corresponding eigenvector is the weight vector since $\tilde{A}w = 3w$ where $w = [w_1, w_2, w_3]^T$. It is easy to see that this statement is correct for an arbitrary 'consistent' $\tilde{A}$ matrix of any size. The following 3 algorithms that are mostly used to compute the weights are based on these observations.

Eigenvector Method

This method is proposed by Saaty [35]. If A corresponds to a pair-wise comparison matrix and w corresponds to a weight vector.

$$\begin{pmatrix} 1 & \frac{w_1}{w_2} & \dots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \dots & \frac{w_2}{w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \dots & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} nw_1 \\ nw_2 \\ \vdots \\ nw_n \end{pmatrix}$$

$$A.w = n.w$$

For such a positive matrix, n corresponds to the maximum eigenvalue. Then the weight vector is selected as the eigenvector that corresponds to the maximum eigenvalue.

$$A.w = \lambda_{max}.w \tag{3.1}$$

λ_{max} is the maximum eigenvalue of A and w is the corresponding eigenvector.

Least Squares Method

If $A = [a_{ij}]_{NxN}$ corresponds to a pair-wise comparison matrix and $w = [w_i]_{1xN}{}^T$ corresponds to a weight vector, then the least squares method is defined as follows [2]:

$$min \sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij} - \frac{w_i}{w_j})^2 \tag{3.2}$$
$$\sum_{i=1}^{n} w_i = 1$$
$$w_i > 0 \text{ for } i = 1, ..., n$$

Logarithmic Least Squares Method

If $A = [a_{ij}]_{NxN}$ corresponds to a pair-wise comparison matrix and $w = [w_i]_{1xN}{}^T$ corresponds to a weight vector, then the logarithmic least squares method is defined as follows [2]:

$$min \sum_{i<j} \sum_{j=1}^{n} [\ln a_{ij} - \ln(\frac{w_i}{w_j})]^2$$

$$\prod_{i=1}^{n} w_i = 1$$
$$w_i > 0 \text{ for } i = 1, ..., n$$

The logarithmic least squares method can be solved analytically. Let $x_i = \ln(w_i)$ and $y_i = \ln(a_{ij})$. With this definition the optimization problem is converted to a quadratic function of the form given below.

$$\sum_{i<j} \sum_{j=1}^{n} [y_{ij} - x_j + x_i]^2$$

The derivative of the above expression is equated to 0 to find the optimum point:

$$nx_i - \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_{ij} \; i = 1, ..., n,$$

under the constraint

$$\sum_{j=1}^{n} x_j = 0$$

The optimum point is obtained as

$$x_i = \frac{1}{n} \sum_{j=1}^{n} y_{ij} \; i = 1, ..., n,$$

So the weights can be expressed in the following form [14]:

$$w_i = (\prod_{j=1}^{n} a_{ij})^{\frac{1}{n}} \tag{3.3}$$

3.2.2 Aggregation Methods

In previous sections, individual decision making methodology of AHP is introduced. This section introduces how the individual decisions of group members' are aggregated.

Geometric Mean Method

Suppose we have n different weights from n different group members:

Then the group decision on weights is obtained as follows [1]:

$$w_j = (\prod_{i=1}^{n} w_{j_i})^{\frac{1}{n}} \tag{3.4}$$

Weighted Arithmetic Mean Method

If $w = [w_i]^T$ corresponds to a weight vector and $a = [a_j]^T$ corresponds to weights of individual judgments, final weights are obtained as follows [29]:

$$w_i^{aggregated} = \sum_{j=1}^{m} a_j w_i \tag{3.5}$$

3.2.3 Consistency Check

The final step of the AHP is the consistency check. Consistency check is made over a consistency ratio (CR), which can be defined as the reliability measure for the answers given to pair-wise comparisons.

It is assumed that the number of weights n corresponds to the maximum eigenvalue λ_{max} of the pair-wise comparison matrix. The judgments become more consistent as λ_{max} gets closer to n. So, the difference between λ_{max} and n can be used as a consistency measure. Instead of using $\lambda_{max} - n$ directly, another measure called consistency index (CI), which was proposed by Saaty [37], is used. Consistency ratio is defined using consistency index as follows:

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{3.6}$$

$$CR = \frac{CI}{R} \tag{3.7}$$

R is is chosen from Table 3.3 according to the number of weights.

Table 3.3: Consistency Table [33]

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.48	1.56	1.57	1.59

The R values given in Table 3.3 are the expected value of the maximum eigenvalue for random reciprocal positive matrices. Consistency ratio CR should be smaller than 0.1 for a consistent judgment. The following example is given to clarify the consistency check concept.

Example: Consider the given pair-wise comparison matrix A

$$A = \begin{bmatrix} 1 & 2 & \frac{1}{3} \\ \frac{1}{2} & 1 & \frac{1}{4} \\ 3 & 4 & 1 \end{bmatrix}$$

The maximum eigenvalue of this matrix is $\lambda_{max} = 3.0183$, then

$CI = \dfrac{3.0183 - 3}{3 - 1} = 0.0091$

and CR is:

$CR = \dfrac{0.0091}{0.58} = 0.0158$

The consistency ratio $CR = 0.0158 < 0.1$, so this pair-wise comparison matrix is consistent.

CHAPTER 4

A NOVEL METHODOLOGY FOR TARGET CLASSIFICATION BASED ON DEMPSTER-SHAFER THEORY

In all classification problems, features are selected according to their discriminating powers and availability. In this study, the available information comes from a radar target tracker that tracks air vehicles. Tracker provides rich information about the state of the target, which is composed of the velocity and the position vectors [5]. Tracker provides the probability density function of the state as a Gaussian density at discrete time instants [5].

Most of the related work that exists in the literature uses Dempster-Shafer theory for target classification. Classification of a target is made using kinematic features or radar cross section or any other relevant information like electronic support measures [9, 11, 31]. Caramicoli et al. [13] and Ristic and Smets [31] use kinematic features to derive some classification rules.

We also use kinematic information in our study and we develop a new methodology for assigning masses to classes. The main distinction of the mass assignment method proposed in this work is to use all available information, i.e., the probability density function of the state, instead of only its mean. The algorithm assumes that the prior probability density functions of all classes and the current measurement are known for some kinematic features of the target. Furthermore, as a sensible assumption, we assume that prior probability density functions of the related kinematic features of all classes can be approximated by Gaussian mixtures. Assigning masses using both prior and measurement probability density functions in Dempster-Shafer framework is the one of the main contributions of this study.

4.1 Dempster-Shafer Framework from Tracking Perspective

In tracking problems, the tracker output is the only information that we can get. Sensors give the state vector of the target with some uncertainty as its probability density function, which is Gaussian. The state vector usually consists of the position and the velocity of the target. The problem is to classify the target as one of the predefined types by using this information. Our first approach is to convert this information to probability masses and combine them, as the information is collected in time. Our application is the classification of air vehicles, so we use the speed and the altitude of the aircraft as discriminating features. The proposed method assigns masses to classes in a novel way. The combination of the masses is done by using Dempster's rule of combination. Our second approach is to combine information first, and then turn them into probability masses.

4.1.1 Basic Probability Assignment

Speed and altitude of an air vehicle are selected as discriminating features for classification. It is assumed that these two variables are independent, hence estimating individual densities is sufficient. The application classifies air targets into four classes, which are bomber and surveillance plane, helicopter, fighter, and unmanned air vehicle.

The prior information about the features for different classes has been collected mainly from Jane's book [24] and the internet. Nominal and maximum speeds, and altitudes of the above defined air vehicle types are used to generate prior probability density functions as mixtures of Gaussians. For each class prior speed and altitude are represented by N Gaussians, where N is selected according to the available data. The probability density functions are obtained by applying the kernel smoothing method to the collected data. The resultant prior probability density functions are given in Section 4.1.1.3.

A tracker gives the (Gaussian) probability density function of the velocity and the position of the target at each time instant, which is considered as the 'measurement'. Altitude, which is one of the features, is already part of the state vector. Hence its probability density function is available. The probability density function of the speed on the other hand should be calculated from the velocity vector. Speed is defined as:

$$s_k = \sqrt{v_x^2 + v_y^2 + v_z^2} \tag{4.1}$$

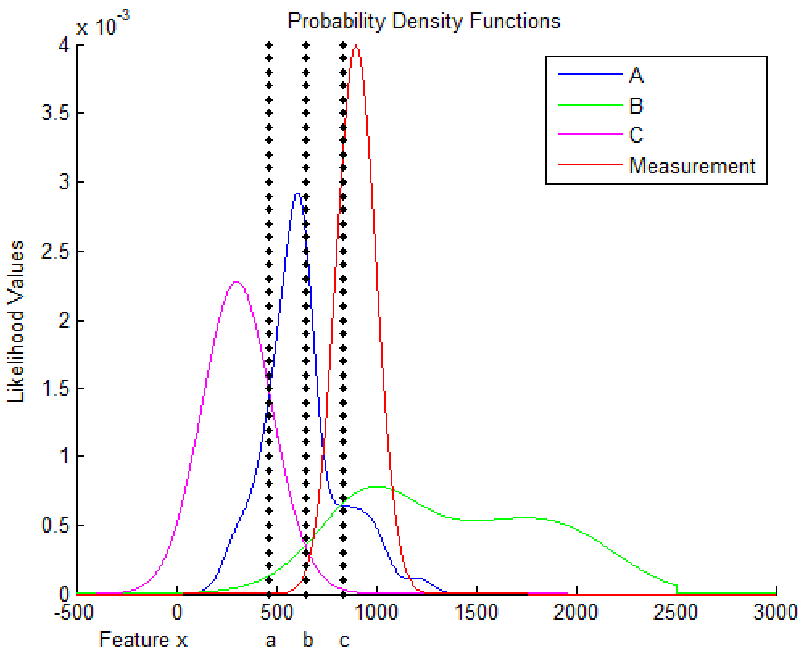
The probability density function of s_k can be approximated as a non-central chi square distribution with 3 degrees of freedom. In this work we obtained the probability density function of the speed from the given Gaussian distribution of the velocity and Equation 4.1 using Monte Carlo methods.

4.1.1.1 The Novel Basic Probability Assignment Method

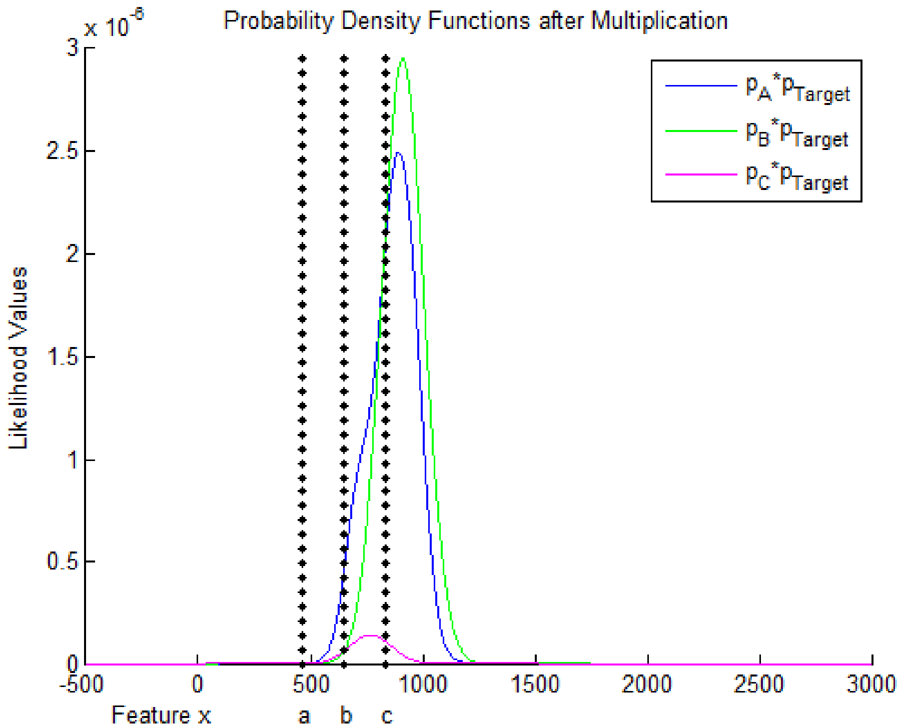
We describe the mass assignment algorithm for a three-class case. Generalization of the algorithm to any number of classes is trivial. Since the number of classes is three, the universal set contains three elements denoted by A, B and C, that is

$$\Theta = \{A, B, C\}$$

Figure 4.1a gives typical prior and measurement probability density functions. The points a, b and c are the equal likelihood points of the corresponding classes and they are used for mass assignments. In order to obtain masses, the measurement probability density function is multiplied by the prior probability density function of each class. The resultant curves are the unnormalized posterior probability density functions, which are shown in Figure 4.1b. The masses are selected to be proportional to the areas under the curves over some selected intervals after multiplication. Formal definitions of the associated masses for a three-class problem are given below.



(a) Prior and measurement probability density functions



(b) Unnormalized posterior probability density functions

Figure 4.1: Probability Density Functions

Let $p_A(x)$, $p_B(x)$ and $p_C(x)$ be prior probability density functions of the A, B and C. Let $p(x)$ be the probability density function of the measurement. Then the masses assigned to the sets are:

$$m(\{A\}) = \int\limits_{x \in S_A} \frac{p_A(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \tag{4.2a}$$

$$S_A = \{x | p_A > p_B \text{ and } p_A > p_C\} \tag{4.2b}$$

$$m(\{B\}) = \int\limits_{x \in S_B} \frac{p_B(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \tag{4.2c}$$

$$S_B = \{x | p_B > p_A \text{ and } p_B > p_C\} \tag{4.2d}$$

$$m(\{C\}) = \int\limits_{x \in S_C} \frac{p_C(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \tag{4.2e}$$

$$S_C = \{x | p_C > p_A \text{ and } p_C > p_B\} \tag{4.2f}$$

$$\begin{aligned} m(\{A, C\}) = &\int\limits_{x \in S_{AC}} \frac{p_A(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \\ + &\int\limits_{x \in S_{CA}} \frac{p_C(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \end{aligned} \tag{4.2g}$$

$$S_{AC} = \{x | p_A < p_C \text{ and } p_A > p_B\} \text{ and } S_{CA} = \{x | p_C < p_A \text{ and } p_C > p_B\} \tag{4.2h}$$

$$\begin{aligned} m(\{A, B\}) = &\int\limits_{x \in S_{AB}} \frac{p_A(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \\ + &\int\limits_{x \in S_{BA}} \frac{p_B(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \end{aligned} \tag{4.2i}$$

$$S_{AB} = \{x | p_A < p_B \text{ and } p_A > p_C\} \text{ and } S_{BA} = \{x | p_B < p_A \text{ and } p_B > p_C\} \tag{4.2j}$$

$$\begin{aligned} m(\{B, C\}) = &\int\limits_{x \in S_{BC}} \frac{p_B(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \\ + &\int\limits_{x \in S_{CB}} \frac{p_C(x)p(x)}{p_A(x)p(x) + p_B(x)p(x) + p_C(x)p(x)} dx \end{aligned} \tag{4.2k}$$

$$S_{BC} = \{x | p_B < p_C \text{ and } p_B > p_A\} \text{ and } S_{CB} = \{x | p_C < p_B \text{ and } p_C > p_A\} \tag{4.2l}$$

$$m(\{B,C\}) = \int\limits_{x \in S_{ABC}} \frac{p_A(x)p(x)}{p_A(x)p(x)+p_B(x)p(x)+p_C(x)p(x)}dx + \int\limits_{x \in S_{BAC}} \frac{p_B(x)p(x)}{p_A(x)p(x)+p_B(x)p(x)+p_C(x)p(x)}dx + \int\limits_{x \in S_{CAB}} \frac{p_C(x)p(x)}{p_A(x)p(x)+p_B(x)p(x)+p_C(x)p(x)}dx \tag{4.2m}$$

$S_{ABC} = \{x|p_A < p_B$ and $p_A < p_C\}$, $S_{BAC} = \{x|p_B < p_A$ and $p_B < p_C\}$ and $S_{CAB} = \{x|p_C < p_A$ and $p_C < p_B\}$

(4.2n)

Note that the procedure formulated above gives the normalized masses, so their sum is unity.

4.1.1.2 Analysis of the Proposed Basic Probability Assignment Method

The analysis of the new methodology is done by comparing it with the basic probability assignment method of Xu et al. [42], which is the most similar method to ours. The following example is introduced both to demonstrate the performance of the new method and to compare it with the method of [42]. A two class artificial scenario is generated for this purpose. Since Xu et al. [42] use Gaussian densities, we selected the prior probability density functions of the two classes and the measurement as Gaussian.

Example

Consider a two-class classification problem. The available information is the prior probability density functions (pdf) of both of the classes and the measurement.

Case 1: Assume that the prior and the measurement probability density functions are as illustrated in Figure 4.2, and are given as:

Prior pdf for class A : $N(x; 1000, 300^2)$
Prior pdf for class B : $N(x; 2000, 300^2)$

pdf of the measurement: $N(x; 1513, 100^2)$

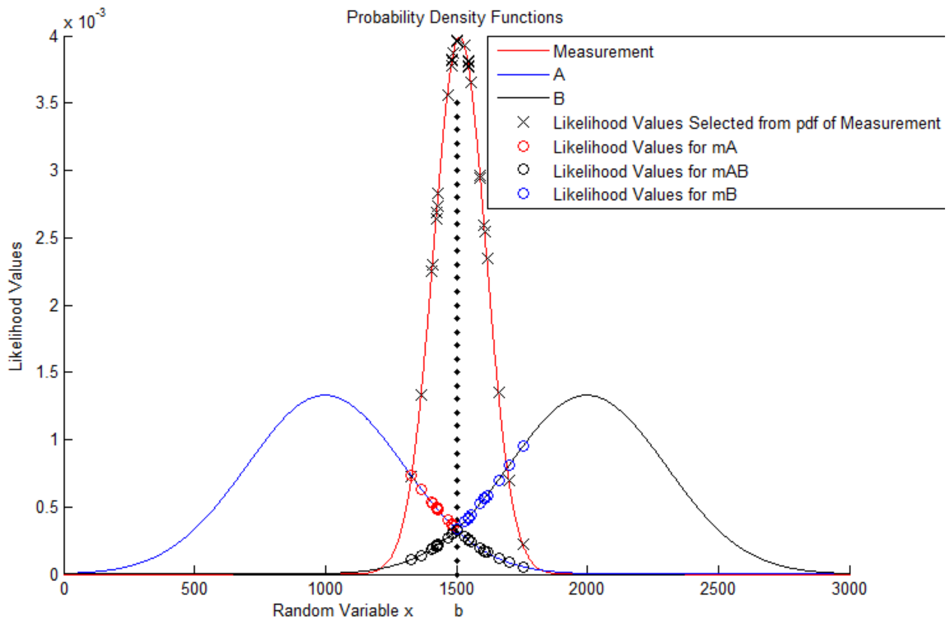


Figure 4.2: Prior and measurement probability density functions

Under these conditions the proposed method assigns the following masses:

$m(\{A\}) = 0.3062$

$m(\{B\}) = 0.3961$

$m(\{A, B\}) = 0.2977$

Case 2: Measurement probability density function is changed to $N(x; 1700, 100^2)$ as shown in Figure 4.3. The new probability masses are as follows.

$m(\{A\}) = 0.0101$

$m(\{B\}) = 0.8741$

$m(\{A, B\}) = 0.1158$

Comparison of the two results shows that when the data is more informative about class identity as in the second case, the mass assignments change accordingly.

To compare the output of the proposed method with Xu et al. [42], we assume that the actual measurement is some x drawn from the measurement probability density

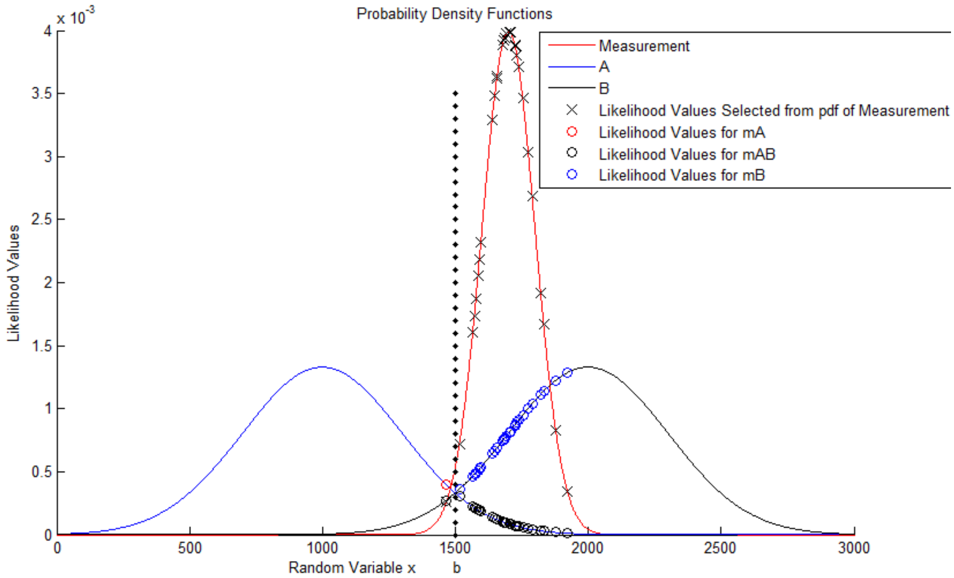


Figure 4.3: Prior and measurement probability density functions

function. Xu et al. [42] assign the masses according to the actual measurement as given below.

$$m(.) = \begin{cases} m(\{A\}) = \alpha p_A(x) \text{ and } m(\{A,B\}) = \alpha p_B(x), & \text{if } p_A(x) > p_B(x) \\ m(\{B\}) = \alpha p_B(x) \text{ and } m(\{A,B\}) = \alpha p_A(x), & \text{if } p_B(x) > p_A(x) \end{cases} \tag{4.3}$$

Note that in this formulation the uncertainty of the measurement is not used. Tables 4.1 and 4.2 are generated from masses that are assigned by drawing 30 samples from the given measurement probability density function as shown in Figures 4.2 and 4.3 according to Xu et al. [42]. The mass values for these 30 samples are illustrated in Figures 4.4 and 4.5 for Case 1 and Case 2 respectively. Figures indicate that the mass values vary according to the selected sample.

Table 4.1: Mean and Standard Deviation Values of the method of Xu et al. [42] for the first case

	Mean	Standard Deviation
$m(\{A\})$	0.2907	0.3453
$m(\{B\})$	0.3908	0.3618
$m(\{A,B\})$	0.3186	0.1232

Table 4.2: Mean and Standard Deviation Values of the method of Xu et al. [42] for the second case

	Mean	Standard Deviation
$m(\{A\})$	0.0198	0.1083
$m(\{B\})$	0.8305	0.1920
$m(\{A,B\})$	0.1497	0.1208

From the results, it can be concluded that the mean values of probability masses that are assigned with the method of Xu et al. [42] coincide with the probability masses that are assigned with the proposed method. However, standard deviations for the 30 samples that are obtained for Case 1 are large enough to take into consideration. Considering the fast convergence in some combination methods, this may create a problem. The proposed method can handle such uncertain data and gives reasonable results compared to the other method. This is because our method uses whole information (the probability density function of the measurement), whereas the other method uses one realization of the measurement.

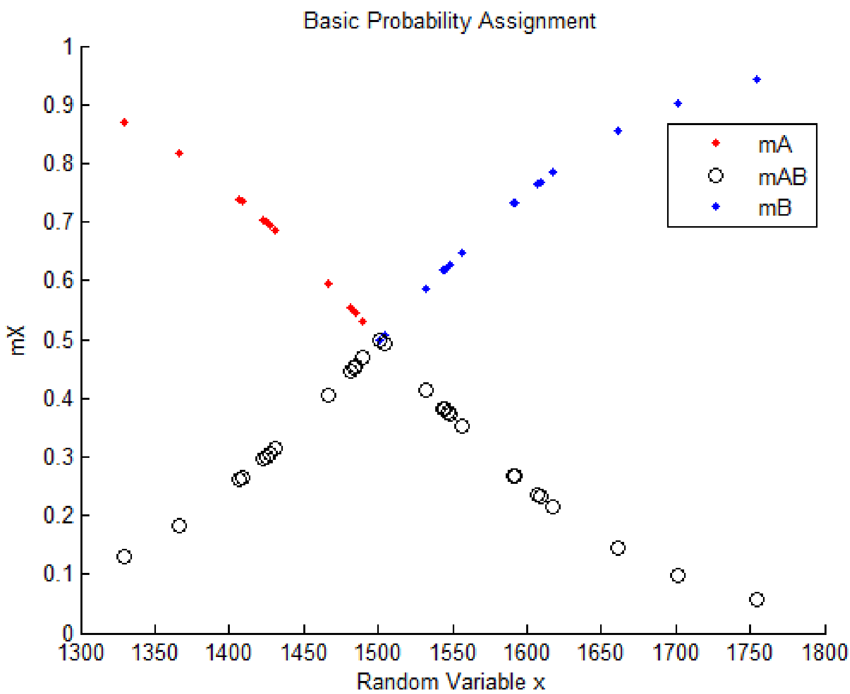


Figure 4.4: Basic probability assignments for the randomly selected 30 samples for the first case

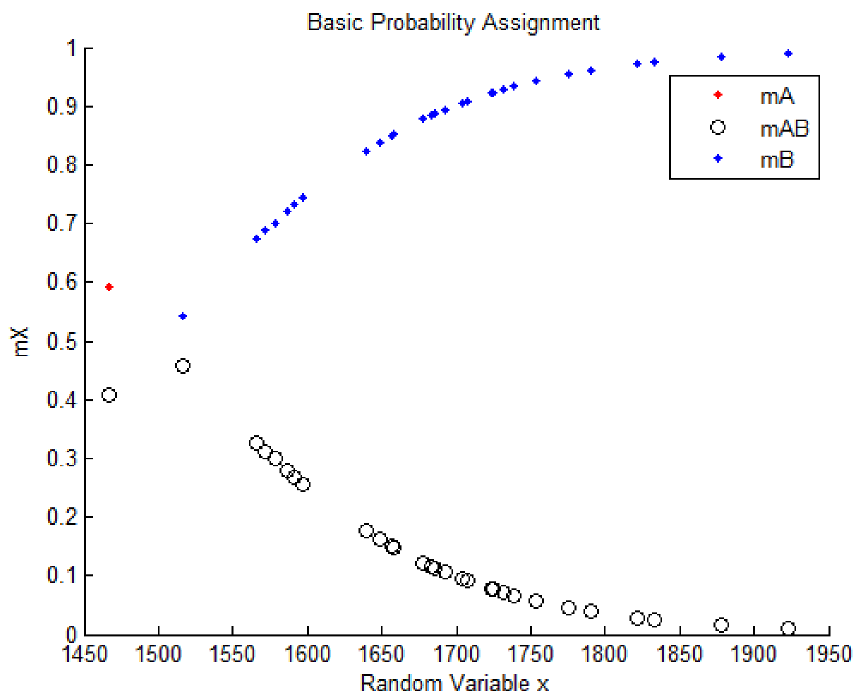


Figure 4.5: Basic probability assignments for the randomly selected 30 samples for the second case

4.1.1.3 Experimental Results

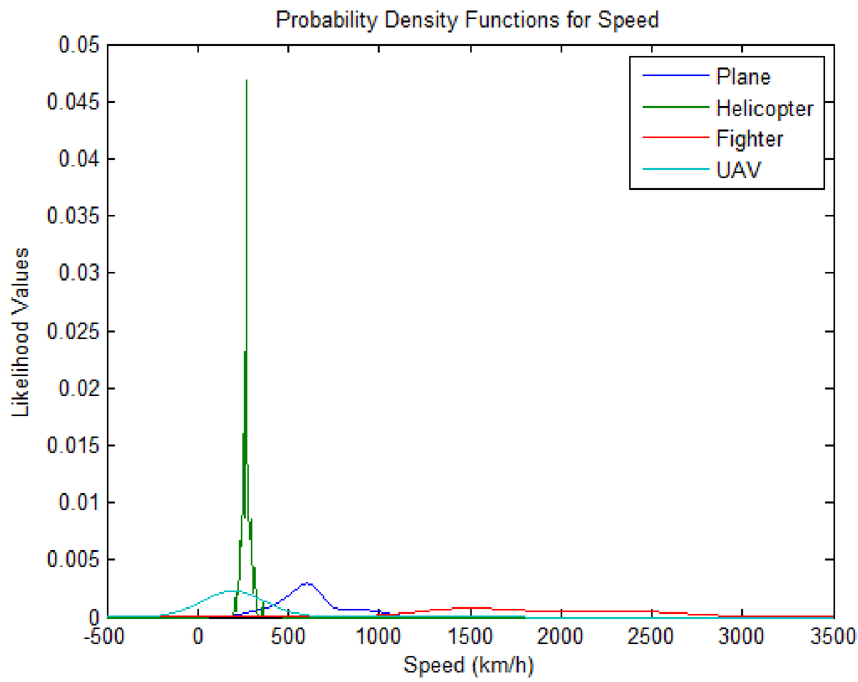
We have conducted two target tracking experiments to examine the proposed method and to compare proposed method with the one that uses only the mean value. The second approach is similar to the approach of Xu et al. [42]. The experiments use real data.

Four air vehicle types are to be classified: Bomber and surveillance planes (P), fighter planes (F), helicopters (H), and unmanned air vehicles (U). In other words the universal set consists of four elements.

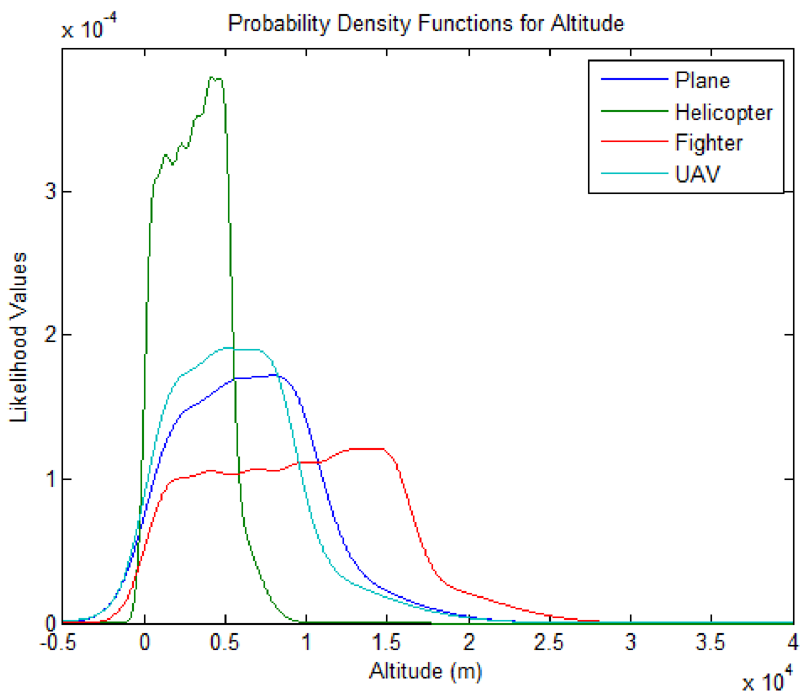
$$\Theta = \{P, F, H, U\}$$

The features used for classification are the speed and the altitude of the air vehicle. To apply the method, first the prior speed and altitude information of the classes, which are collected from open sources, are transformed into probability density functions. These functions are illustrated in Figure 4.6. Mass assignments are done using the tracker outputs, which are probability density functions.

In the first experiment we used fighter data. In the second experiment we used helicopter data. The trajectories and the tracker outputs are given in Figures 4.7 and 4.8 for both of the experiments.

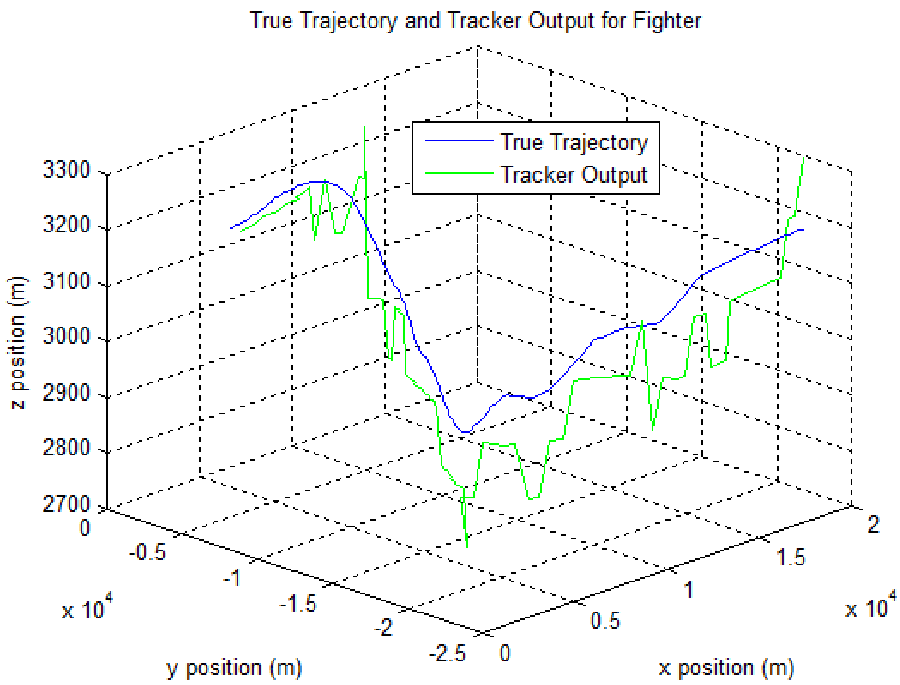


(a) Speed

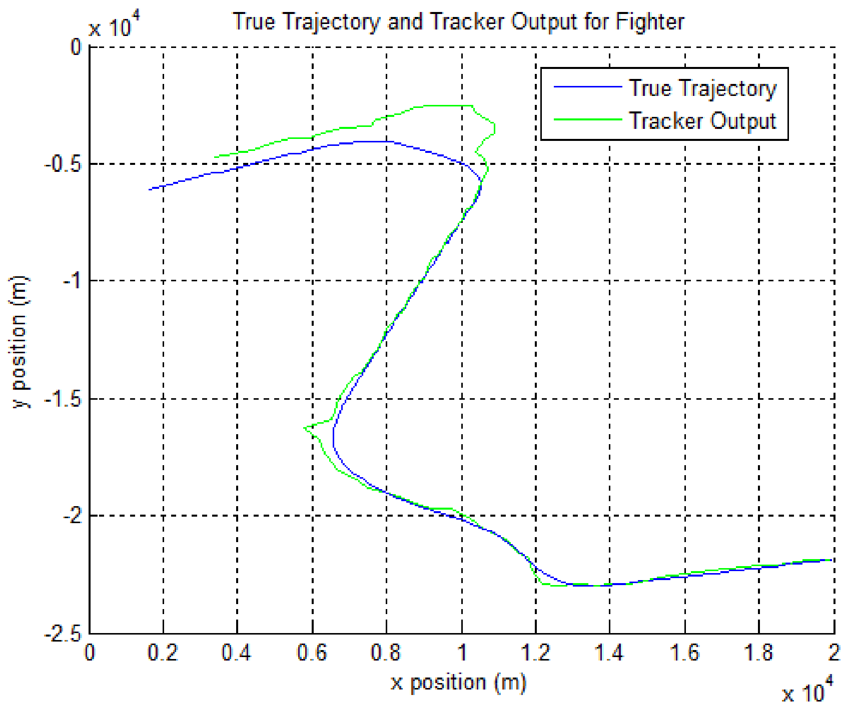


(b) Altitude

Figure 4.6: Prior probability density functions

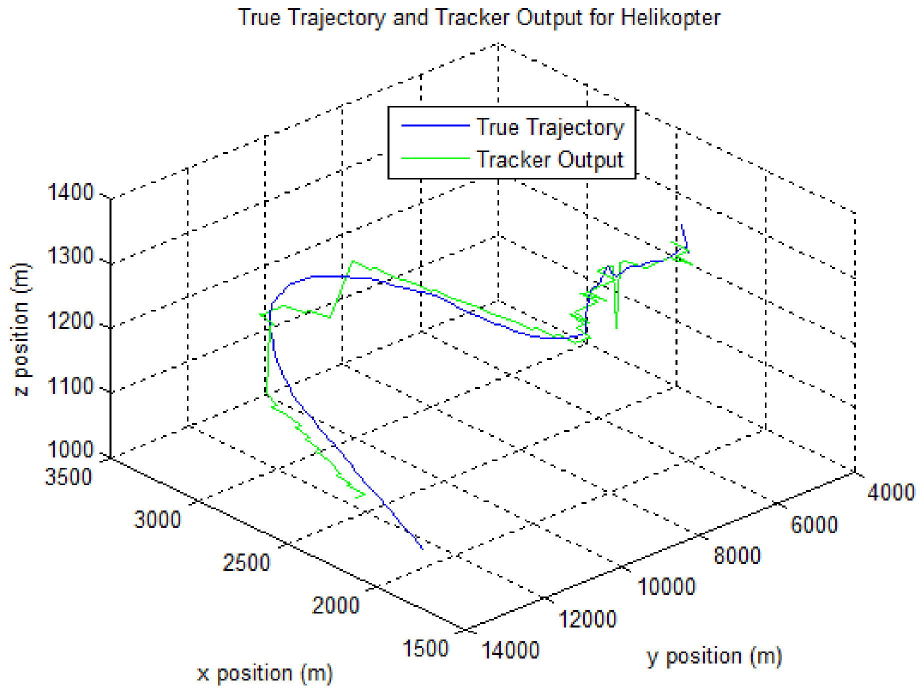


(a) 3D view

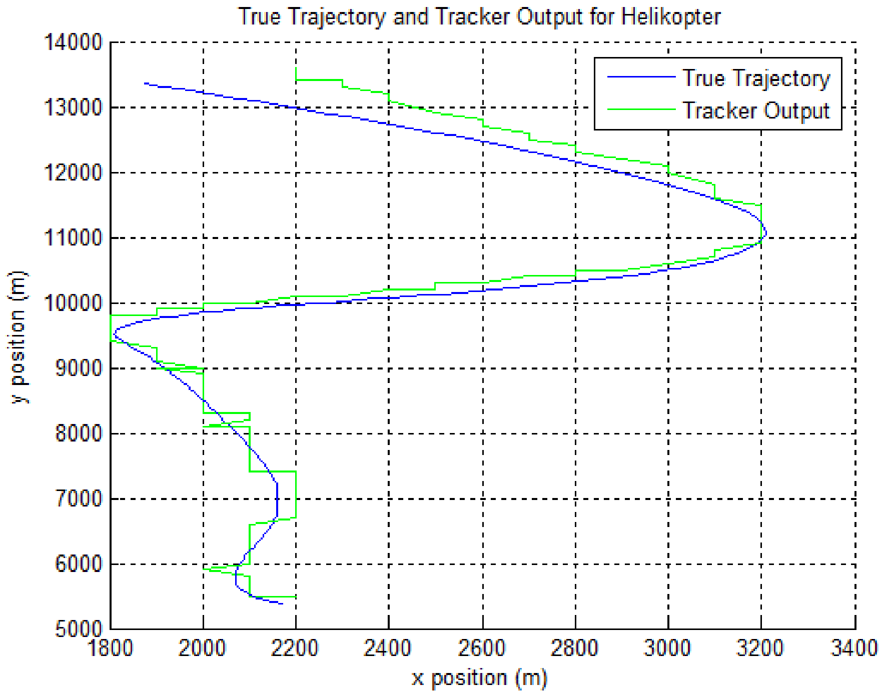


(b) x-y plane view

Figure 4.7: True trajectory and the tracker output for Fighter



(a) 3D view



(b) x-y plane view

Figure 4.8: True trajectory and the tracker output for Helicopter

Experiment 1

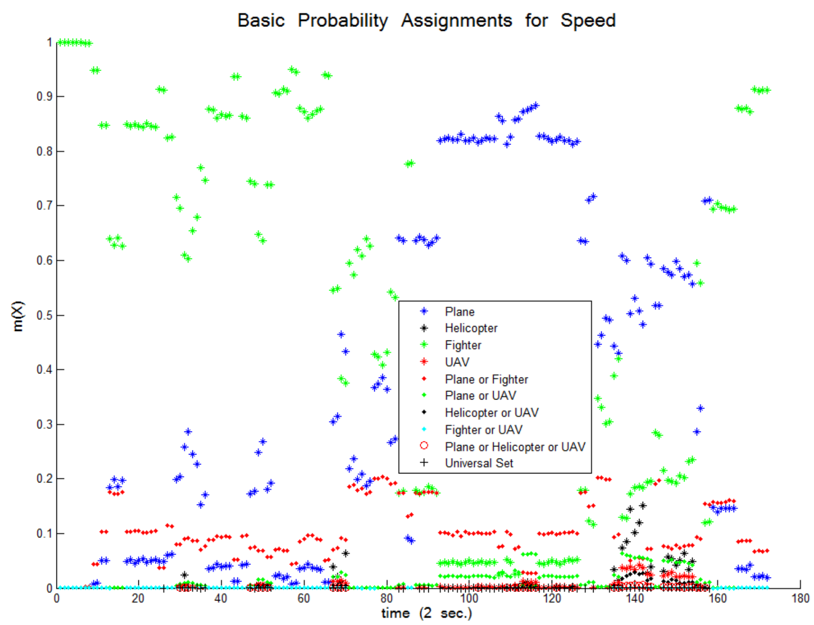
In the first experiment a fighter is tracked. Its route and tracker output are shown in Figure 4.7. The air vehicle begins its journey at high speeds that only a fighter can achieve. So, the proposed algorithm that uses speed classifies the target as fighter. However towards the end of the scenario it slows down. So, proposed algorithm assigns high probability masses to set $\{P\}$ and some probability masses to the set $\{F, P\}$ towards the end of the scenario. The instantaneous mass assignments due to speed are given in Figure 4.9a.

The target flies at low altitudes. So, the algorithm that uses altitude assigns high probability masses to the set $\{H\}$ and some probability masses to the sets $\{H, U\}$, $\{P, H, U\}$ and Θ.

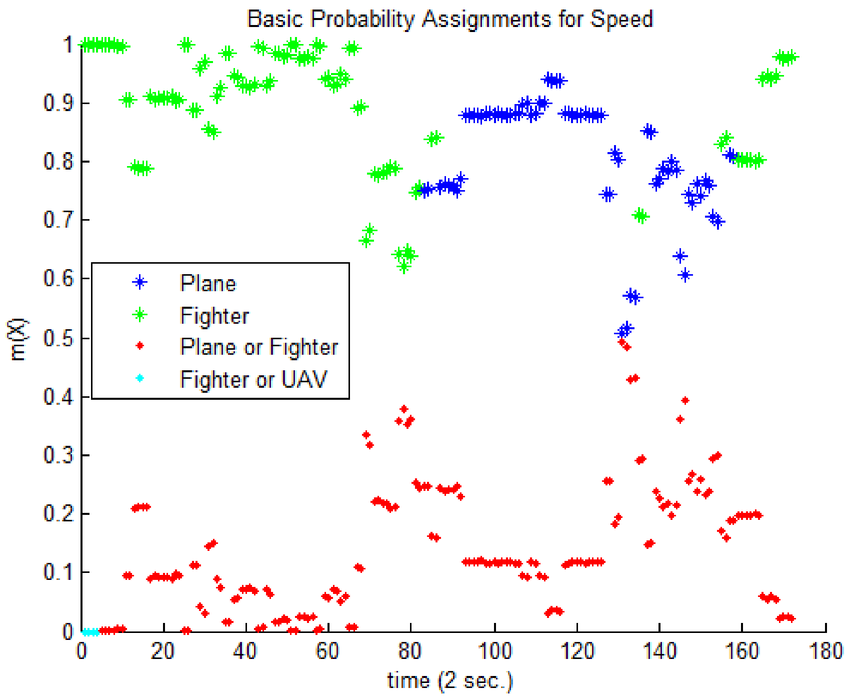
The instantaneous masses are combined in time. As seen in Figure 4.14a the combination of probability masses according to speed assigns high probability masses to the fighter up to 134 seconds and then switches to plane. This undesired effect can be eliminated by constraining the decision as: once the decision is a fighter then it cannot be switched to any other type. The result of the constraining the decision can be seen from Figure 4.15. The idea behind constraining the decision is using limits of the physical capabilities of the air vehicles. For example none of the air vehicles can reach the speed of a fighter. Once the target is classified as a fighter with a probability greater than 0.98, the class of the air vehicle cannot be changed.

We have compared the results of the proposed method with the method that uses only the mean (or the estimate) of the features and it is similar to the one given in Xu et al. [42]. The results of instantaneous mass assignments for both proposed method and method proposed in Xu et al. [42] are given in Figures 4.9, 4.10 and 4.11. The results of combined mass assignments in time for both proposed method and method proposed in Xu et al. [42] are given in Figures 4.12, 4.13 and 4.14.

Comparison of the two methods shows that the method that uses only mean values is more certain about its decisions while the proposed method is more conservative. That is an expected result since the proposed one uses the uncertainty as well. The certainty is good when the models are good, as in this example, but some conservativeness is necessary when the estimates deviate from the true value as the next example shows.

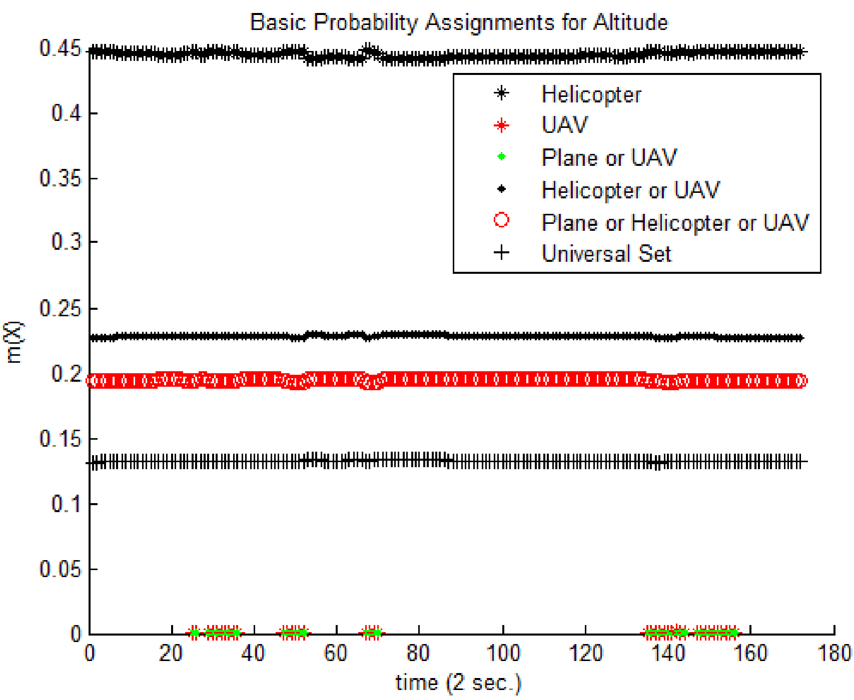


(a) Proposed method

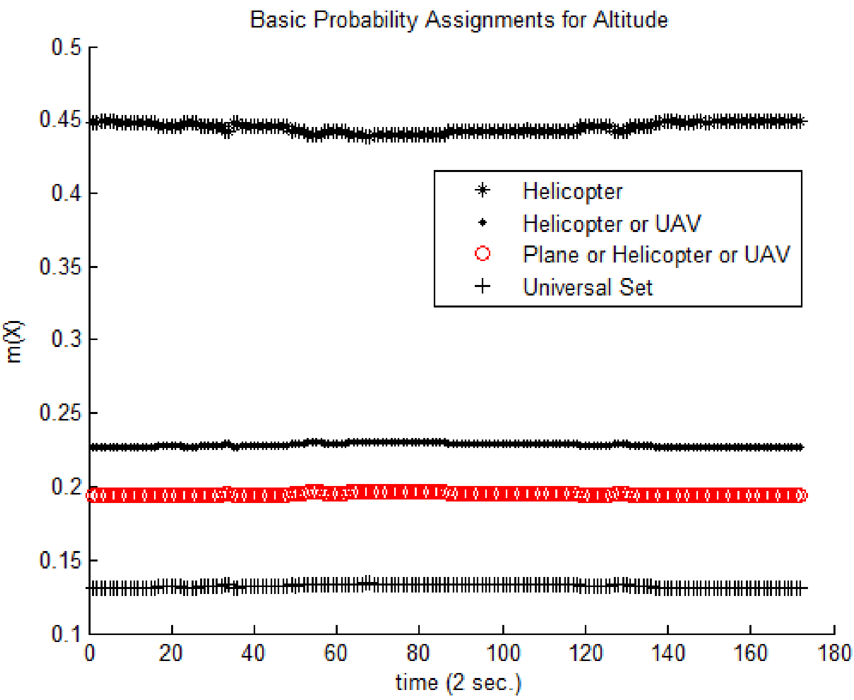


(b) Method that uses the mean

Figure 4.9: Instantaneous mass assignment for speed

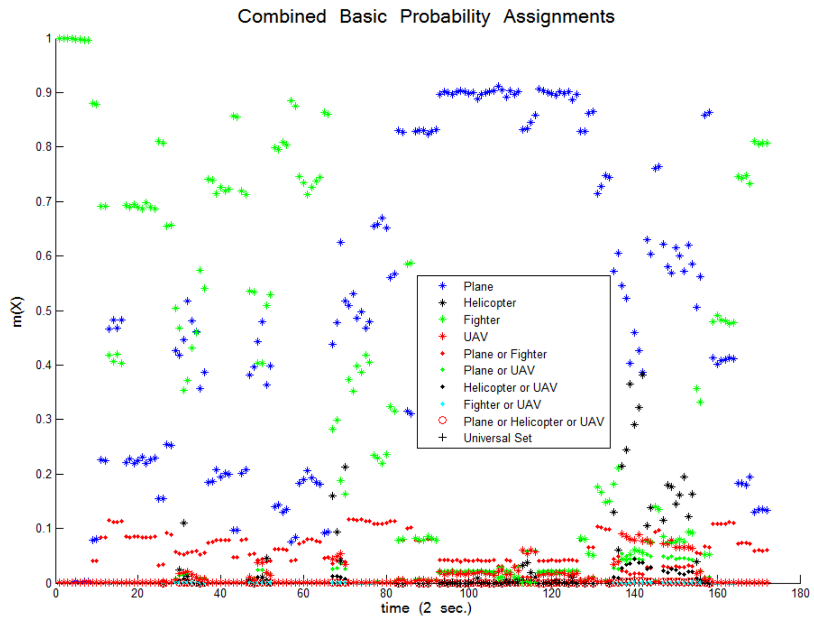


(a) Proposed method

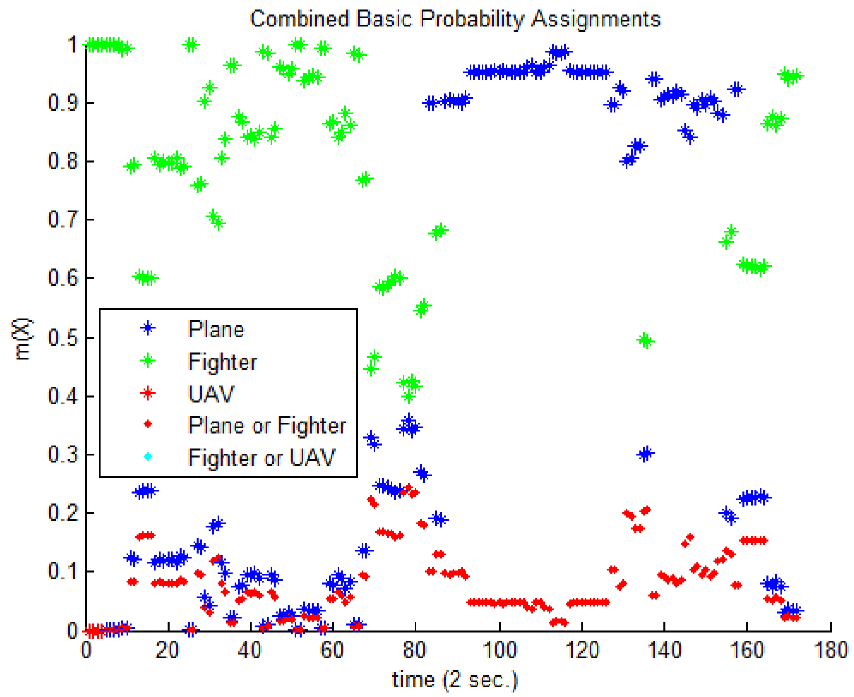


(b) Method that uses the mean

Figure 4.10: Instantaneous mass assignment for altitude

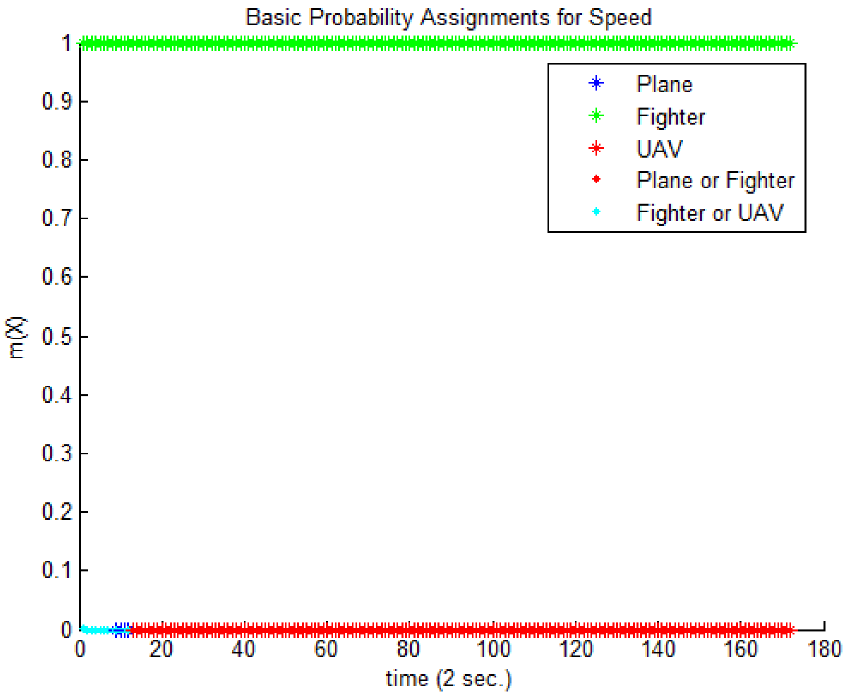


(a) Proposed method

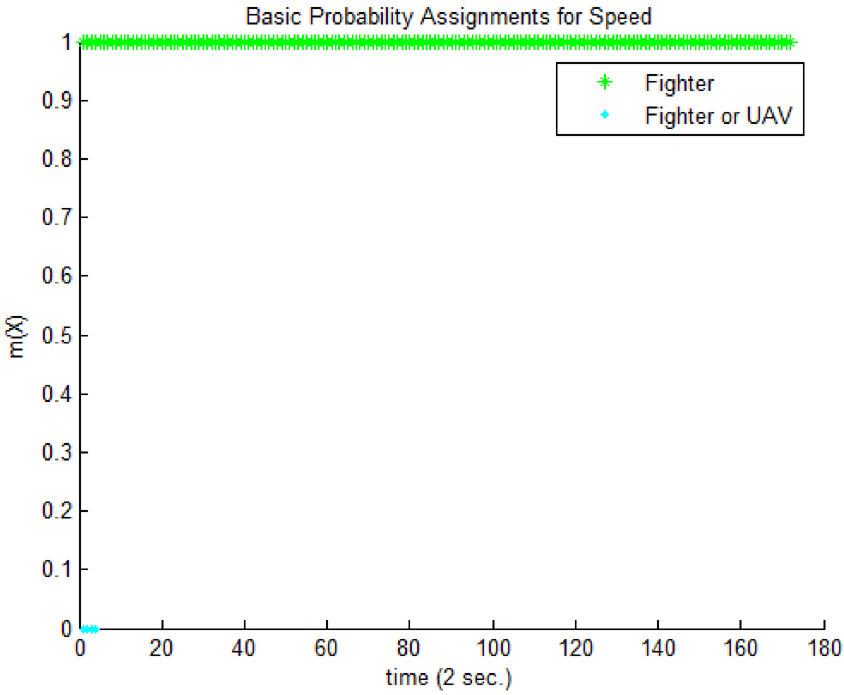


(b) Method that uses the mean

Figure 4.11: Instantaneous combined masses of features (speed and altitude)

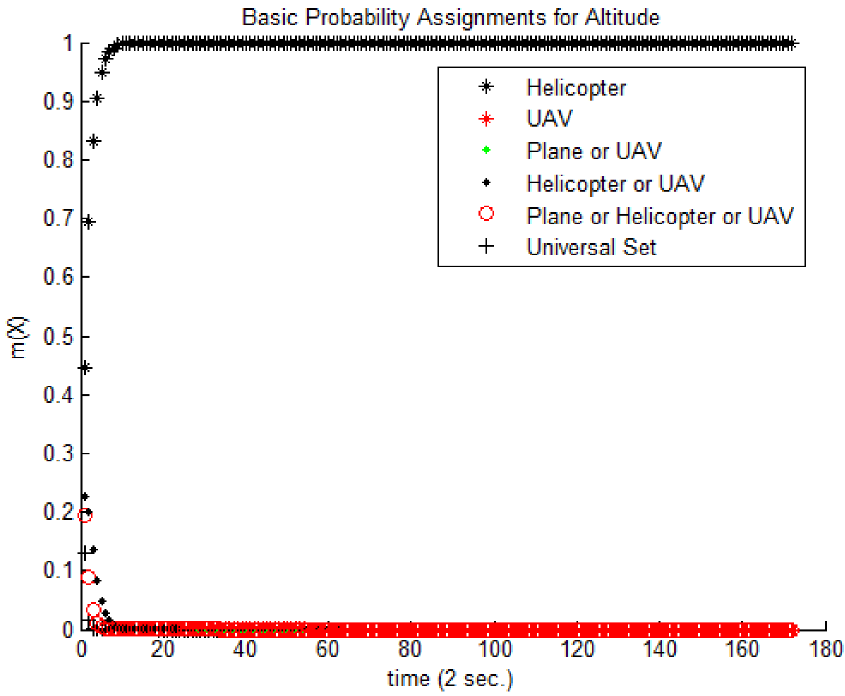


(a) Proposed method

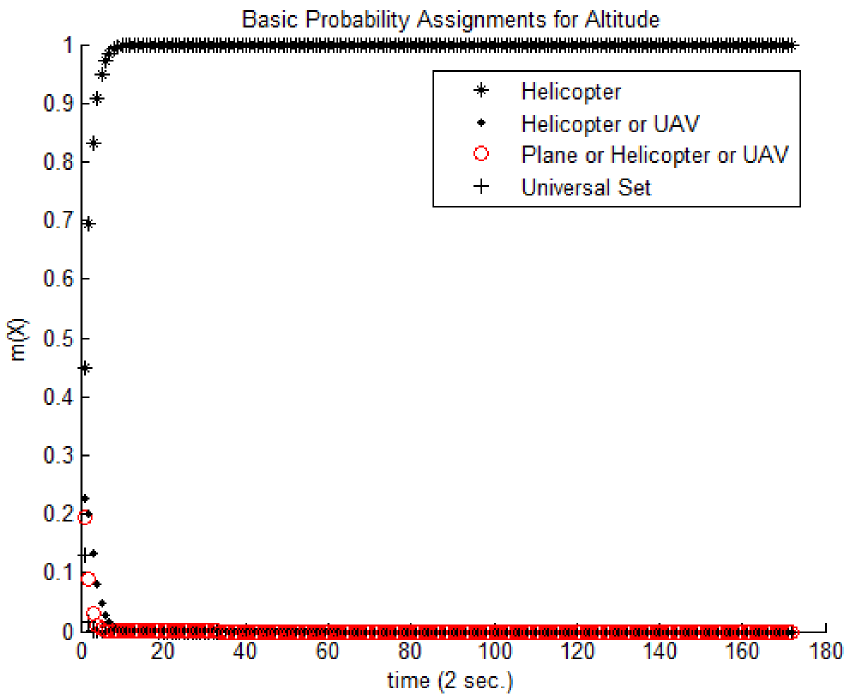


(b) Method that uses the mean

Figure 4.12: Combined masses from initial time to final time for speed

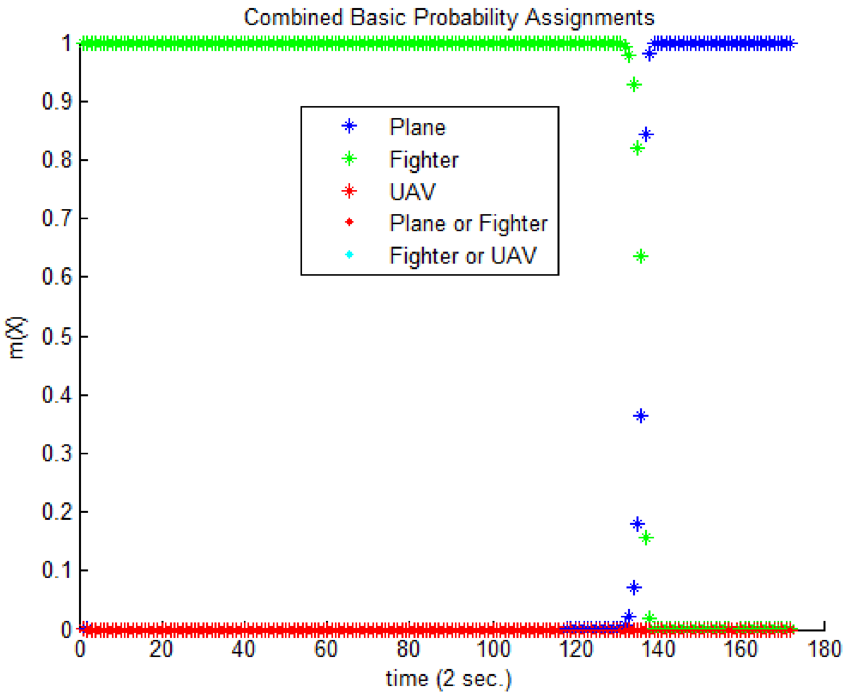


(a) Proposed method

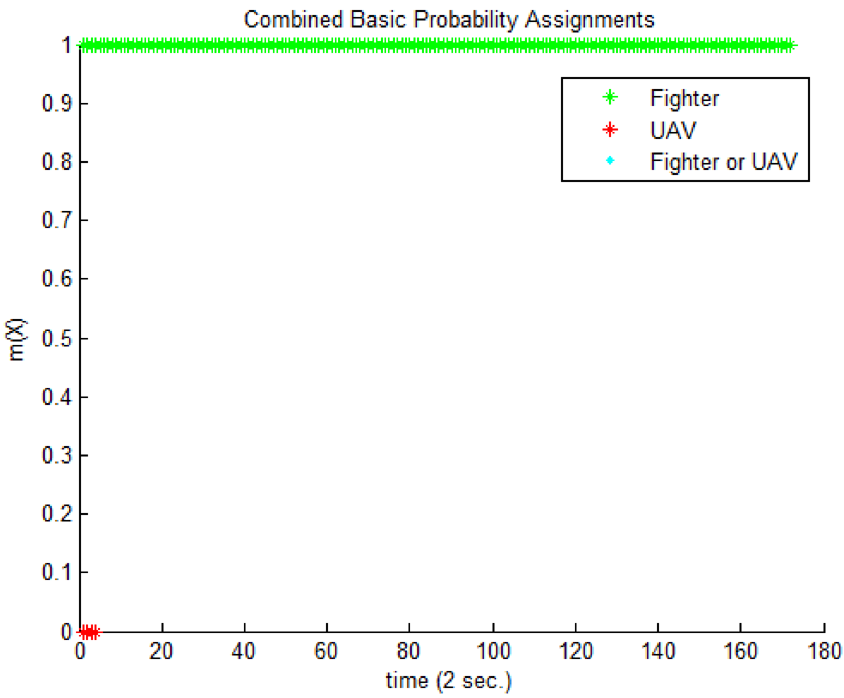


(b) Method that uses the mean

Figure 4.13: Combined masses from initial time to final time for altitude



(a) Proposed method



(b) Method that uses the mean

Figure 4.14: Combined masses from initial time to final time of features (speed and altitude)

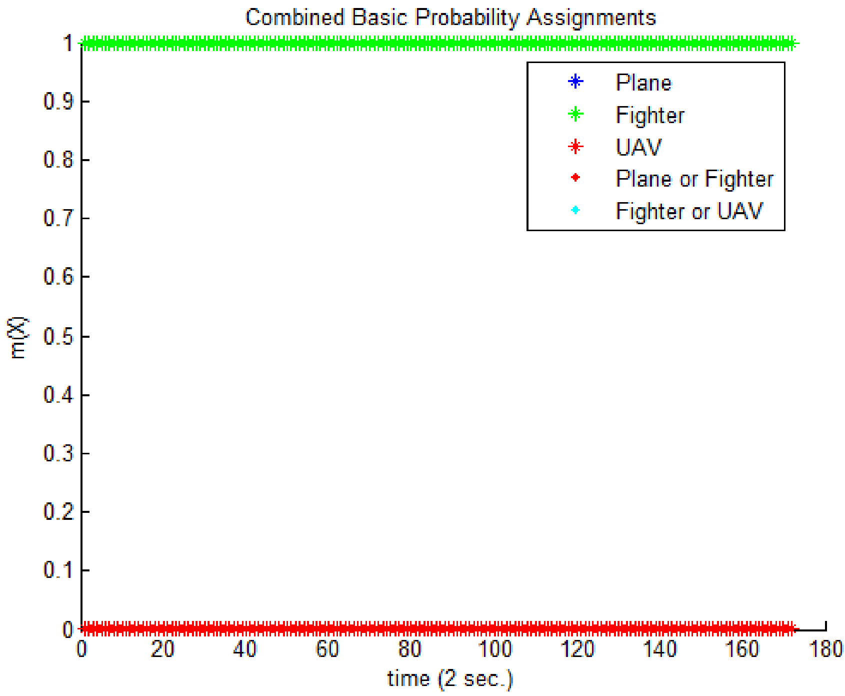


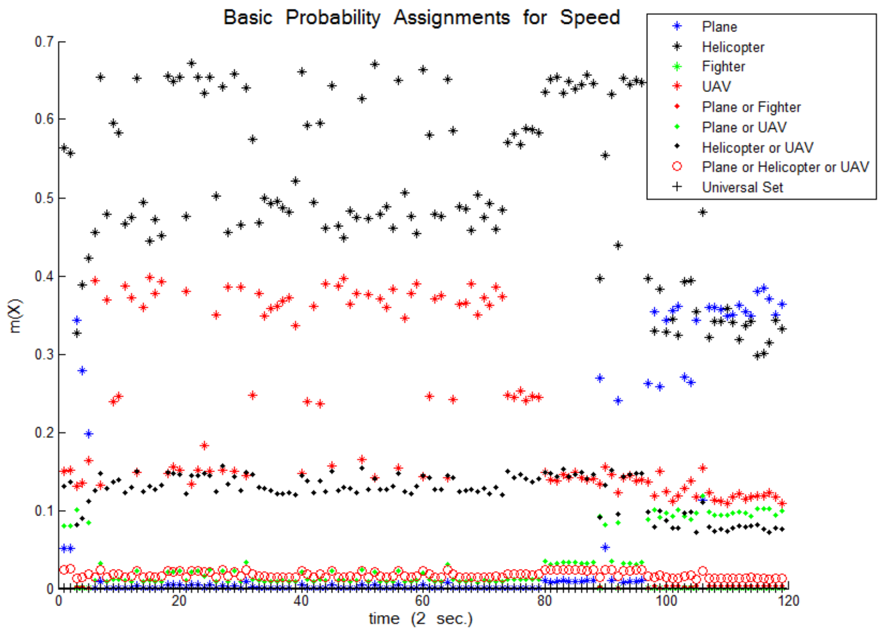
Figure 4.15: Thresholded combined masses from initial time to final time of features (speed and altitude)

Experiment 2

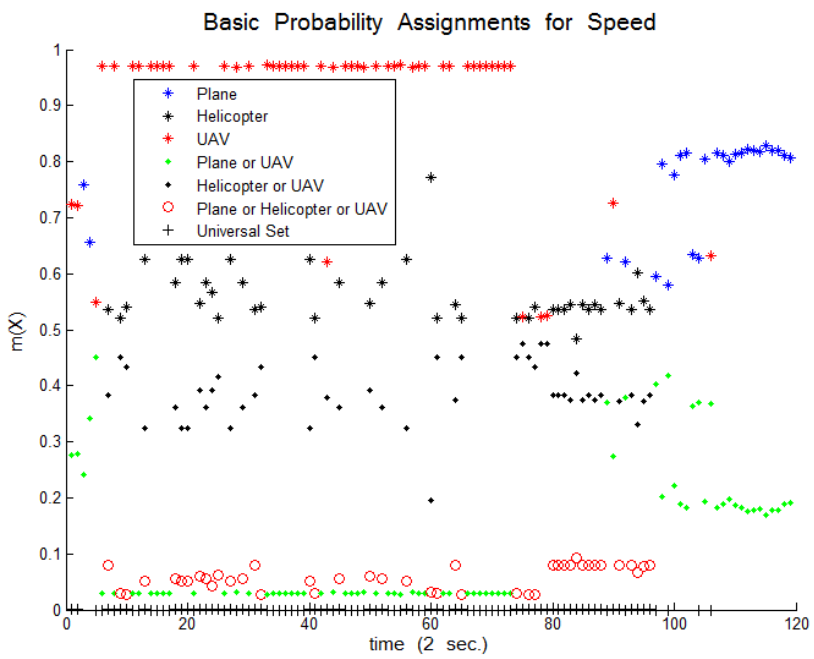
In the second experiment a helicopter is tracked. Its route and tracker output are shown in Figure 4.8. From the beginning to the end of the trajectory, the proposed algorithm assigns high probability masses to the set $\{H\}$ and some probability masses to other sets. For an ultimate decision, first, the instantaneous masses obtained from the speed and the altitude are combined, then these instantaneous decisions from the initial time to the final time are combined. The target is classified as a helicopter at all times after the 8^{th} second with a mass greater than 0.98.

To make a comparison, we also applied the method of Xu et al. [42] to the same data by considering the measurements to be the mean value of the features that are provided by the tracker. The results of instantaneous and combined mass assignments for both the proposed method and the method proposed in Xu et al. [42] are given in Figures 4.16, 4.17 4.18, 4.19, 4.20 and 4.21.

The method that uses only the mean value makes wrong assignments in quite a number of instances since the mean alone may give low likelihood value due to the unsmooth nature of the speed of the helicopter. This results in a wrong decision.

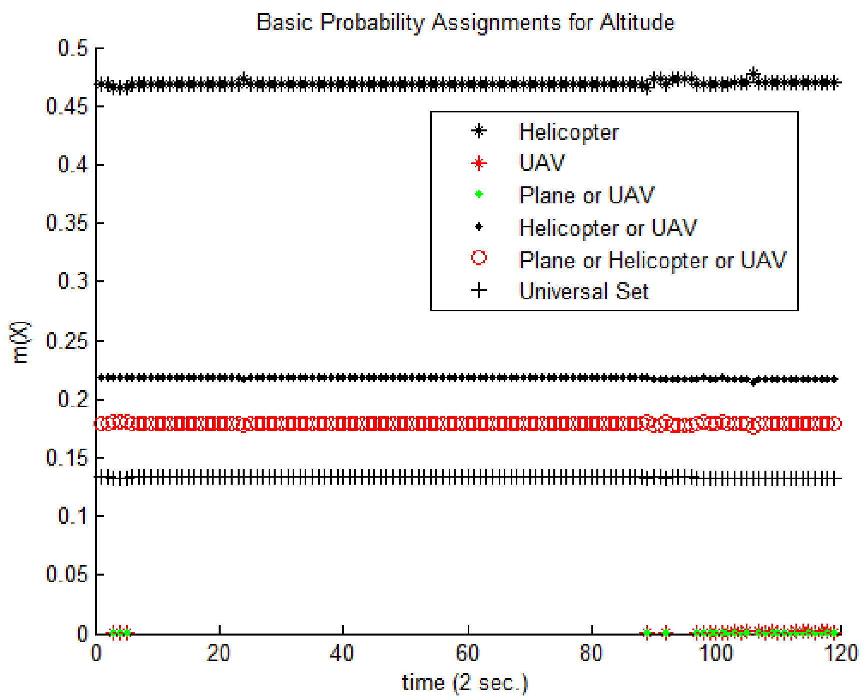


(a) Proposed method

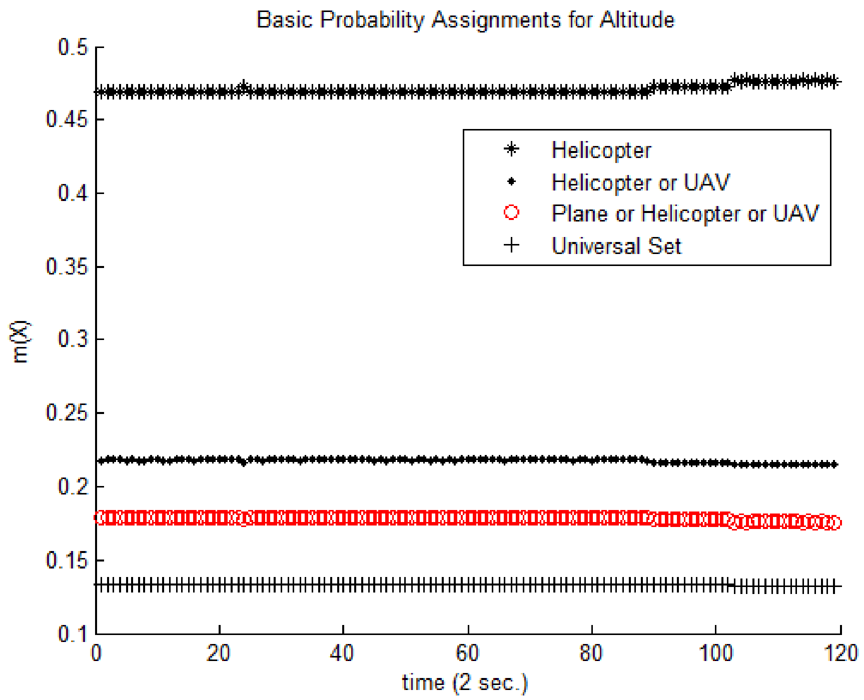


(b) Method that uses the mean

Figure 4.16: Instantaneous mass assignment for speed

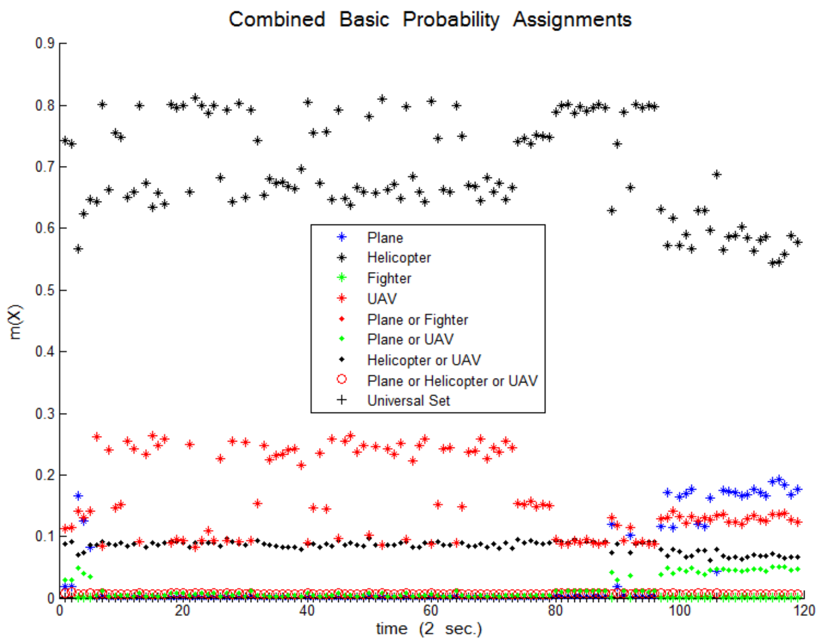


(a) Proposed method

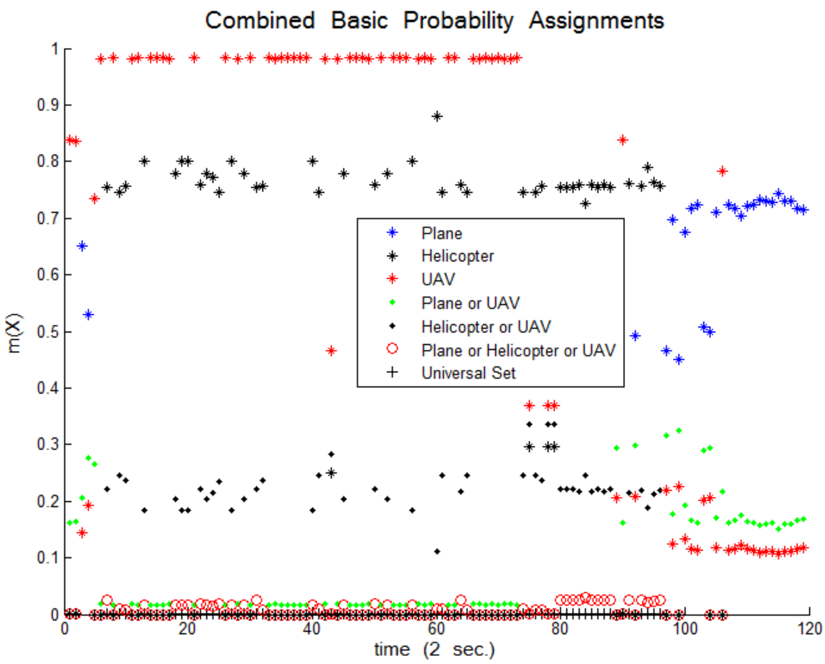


(b) Method that uses the mean

Figure 4.17: Instantaneous mass assignment for altitude

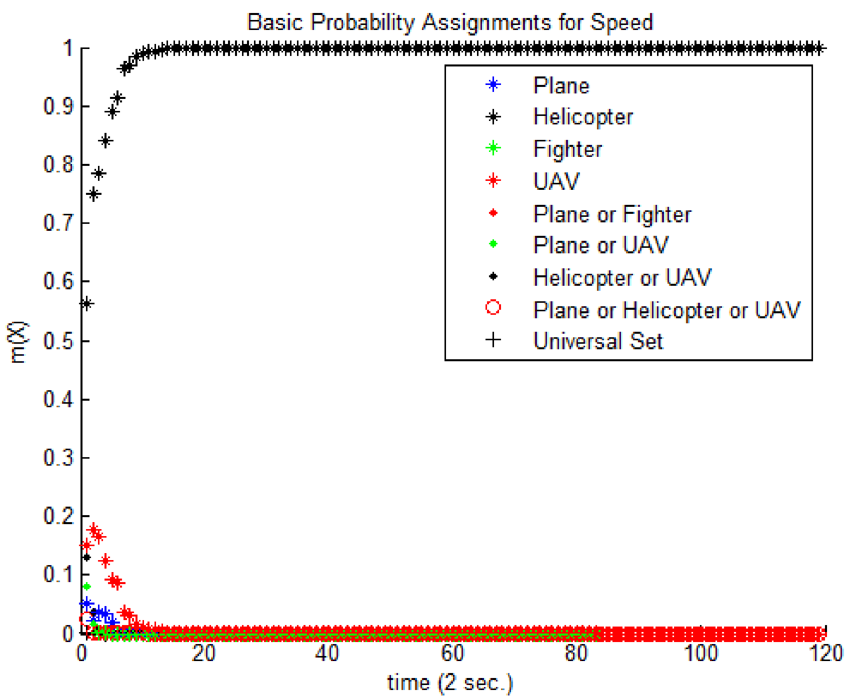


(a) Proposed method

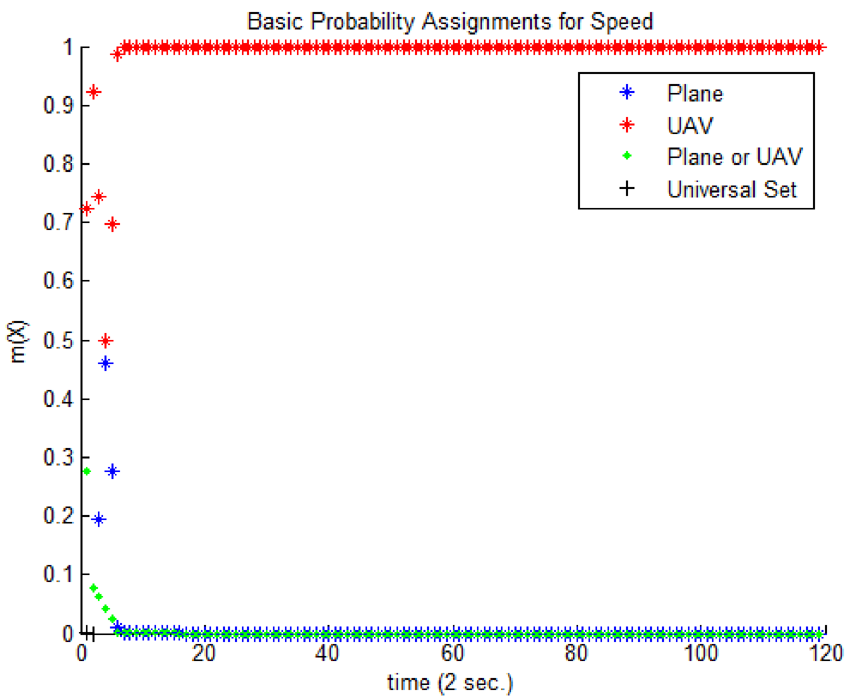


(b) Method that uses the mean

Figure 4.18: Instantaneous combined masses of features (speed and altitude)

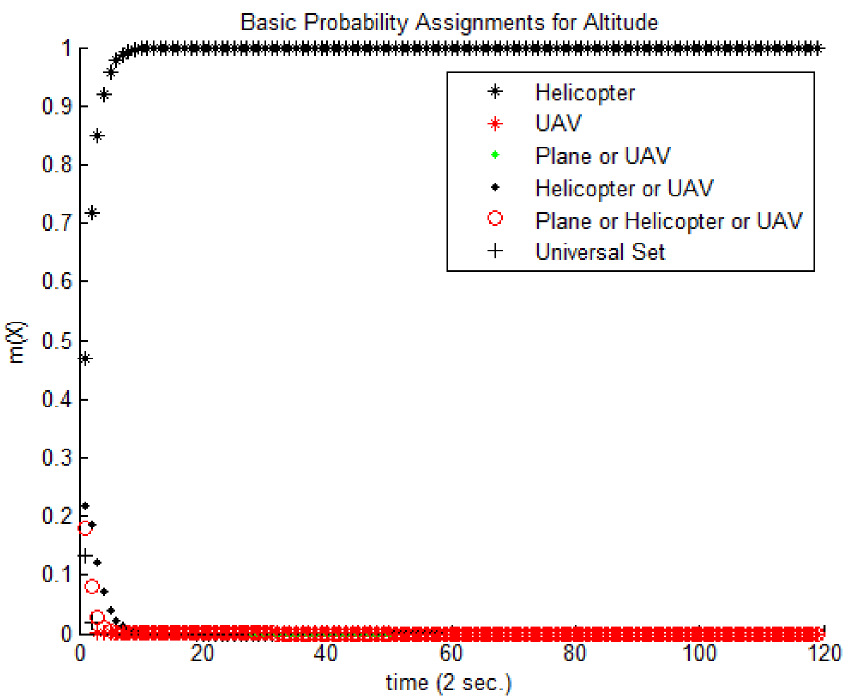


(a) Proposed method

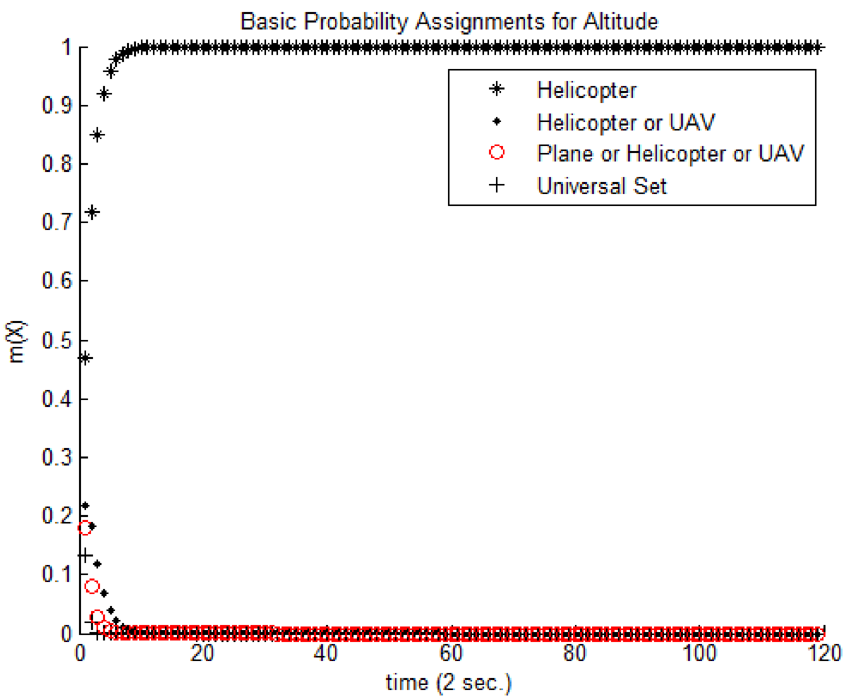


(b) Method that uses the mean

Figure 4.19: Combined masses from initial time to final time for speed

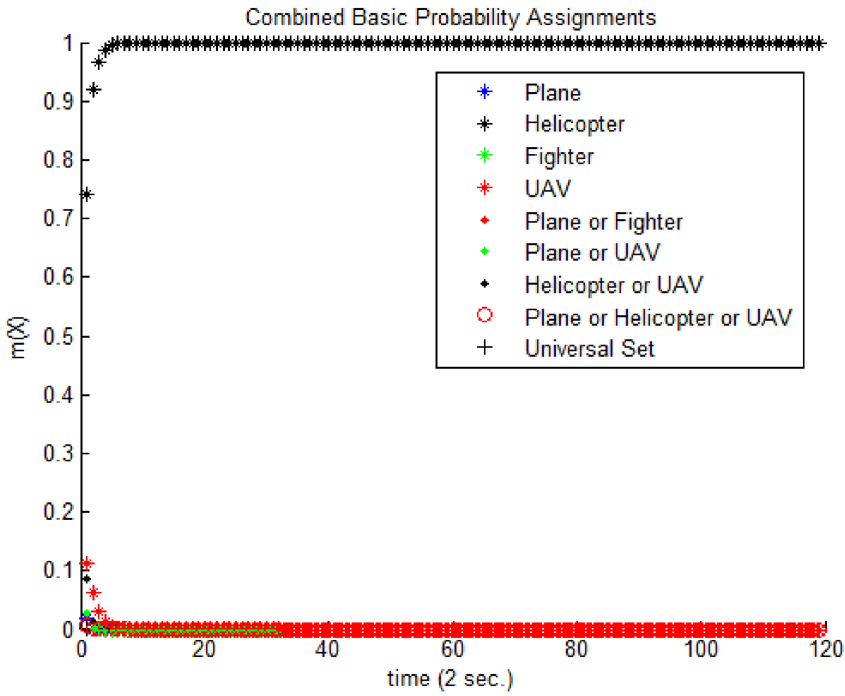


(a) Proposed method

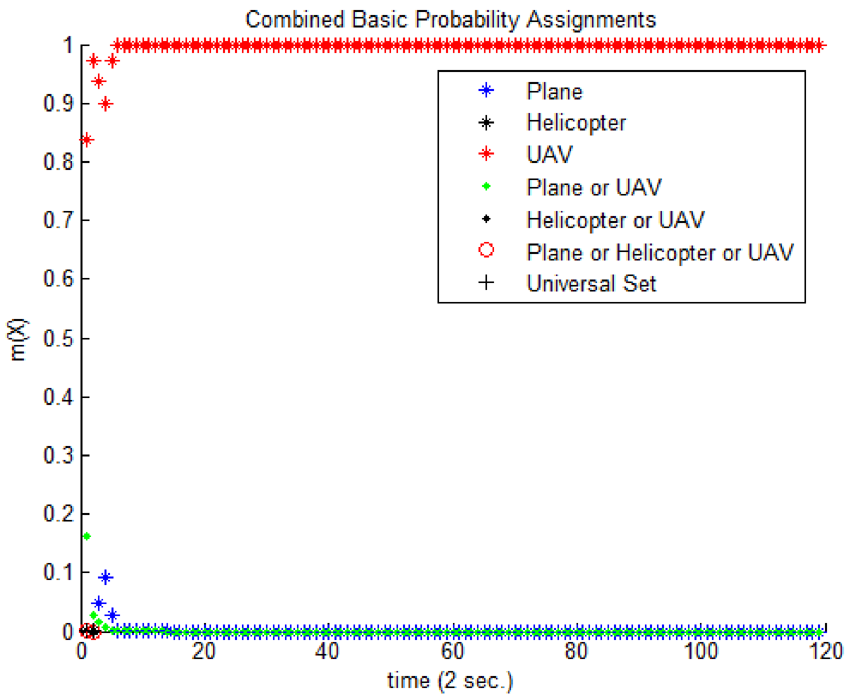


(b) Method that uses the mean

Figure 4.20: Combined masses from initial time to final time for altitude



(a) Proposed method



(b) Method that uses the mean

Figure 4.21: Combined masses from initial time to final time of features (speed and altitude)

4.1.1.4 Evaluation of the Basic Probability Assignment Methodology

The proposed method is tested with real data. Results are compared with another methodology that uses only the mean estimates of the features that can be considered quite similar to the one given in Xu et al. [42]. Test results show that using the whole probability density function of the features provided by the tracker brings significant advantages for classification compared to using only the mean of the features.

When a measurement comes, the algorithm assigns the probability masses instantaneously for every feature and combines them. The computation time spent to reach the overall decision is about 0.2843 seconds on an Intel i7 computer with 8 GB RAM and Matlab 2012a. Considering that the air defense radars produce measurement reports with periods in the order of seconds, the algorithm is certainly fast enough for real time operation.

4.1.2 Combination of Evidence

The combination of evidence as combination of different mass assignments is another important concept in Dempster-Shafer theory. There is a huge literature on this topic and it is given in Section 2.3.

Here, specific to our problem, we propose a new methodology. Instead of combining mapped information we combine the whole information, i.e. the probability density functions. This new methodology postpones the mass assignment until a reasonable amount of information, i.e. probability density functions, is collected. Thus, we claim that this new methodology prevents information loss. If the collected probability density functions are assumed to be independent then they can be multiplied to obtain the overall probability density function at all times. The mass assignment algorithm given in Section 4.1.1.1 can be applied to the new combined probability density function. However, the independence assumption is not correct since the output of the tracker gives the processed data. A de-correlation algorithm is used to overcome this problem.

4.1.2.1 The Combination Methodology

The combination methodology is simple:

- Decorrelate the measurements,
- Combine them by multiplication.

$$p_{final}(x) = \prod_{i=1}^{n} p_i(x) \tag{4.4}$$

4.1.2.2 Analysis of the Combination Methodology

The proposed method is analyzed by comparing the combination performance of some well-known combination rules in the literature. In order to analyze efficiency and reliability of the methodology, two artificial scenarios are constructed. The first scenario is given below and it is illustrated in Figure 4.22.

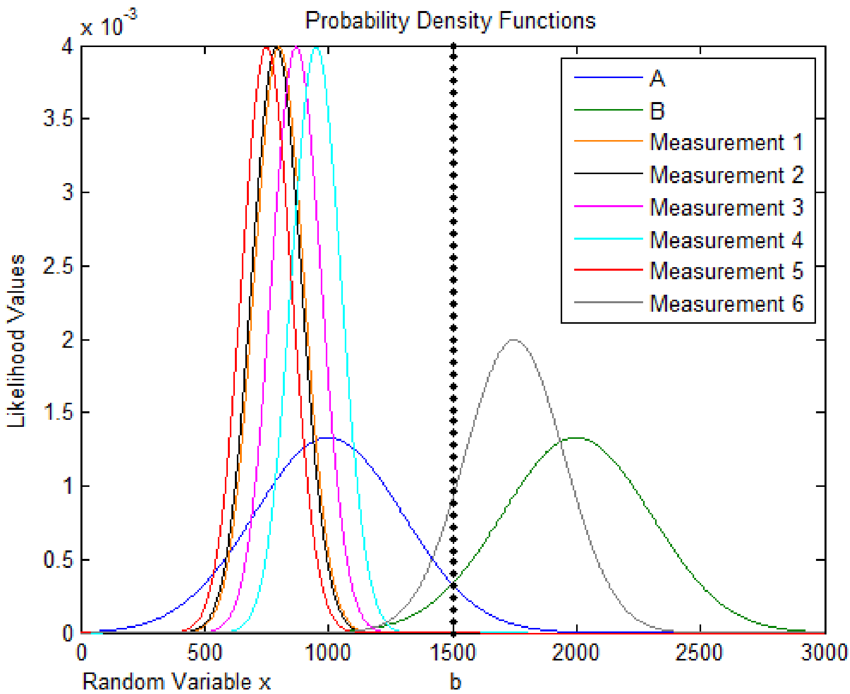


Figure 4.22: Measurement and the prior probability functions

Prior probability density function of class A : $N(x; 1000, 300^2)$

Prior probability density function of class B : $N(x; 2000, 300^2)$

Probability density function of the 1^{st} measurement: $N(x; 800, 100^2)$

Probability density function of the 2^{st} measurement: $N(x; 790, 100^2)$

Probability density function of the 3^{st} measurement: $N(x; 870, 100^2)$

Probability density function of the 4^{st} measurement: $N(x; 950, 100^2)$

Probability density function of the 5^{st} measurement: $N(x; 750, 200^2)$

Probability density function of the 6^{st} measurement: $N(x; 1750, 200^2)$

The basic probability assignment for each measurement is given in Table 4.3, and the results of different combination rules are given in Table 4.4.

The scenario is selected so that the first 5 measurements are consistent, however the last one is different from the first 5. The combination results are given in Table 4.4. First row of the Table 4.4 shows that the unnormalized masses. Note that mass assigned to the empty set (given in the last column) is large and it is an indicator of inconsistency. Different combination algorithms distribute this mass differently. The inconsistent data can be interpreted in several ways:

- The last measurement can be wrong
- Data can be time varying

Table 4.3: Basic Probability Assignments for the measurements

pdfs	Probability Masses		
	$m(\{A\})$	$m(\{B\})$	$m(\{A, B\})$
$N(x; 800, 100^2)$	0.9991	0.0000	0.0009
$N(x; 790, 100^2)$	0.9992	0.0000	0.0008
$N(x; 870, 100^2)$	0.9982	0.0000	0.0018
$N(x; 950, 100^2)$	0.9959	0.0000	0.0041
$N(x; 750, 100^2)$	0.0577	0.8511	0.0912
$N(x; 1750, 200^2)$	0.9994	0.0000	0.0006

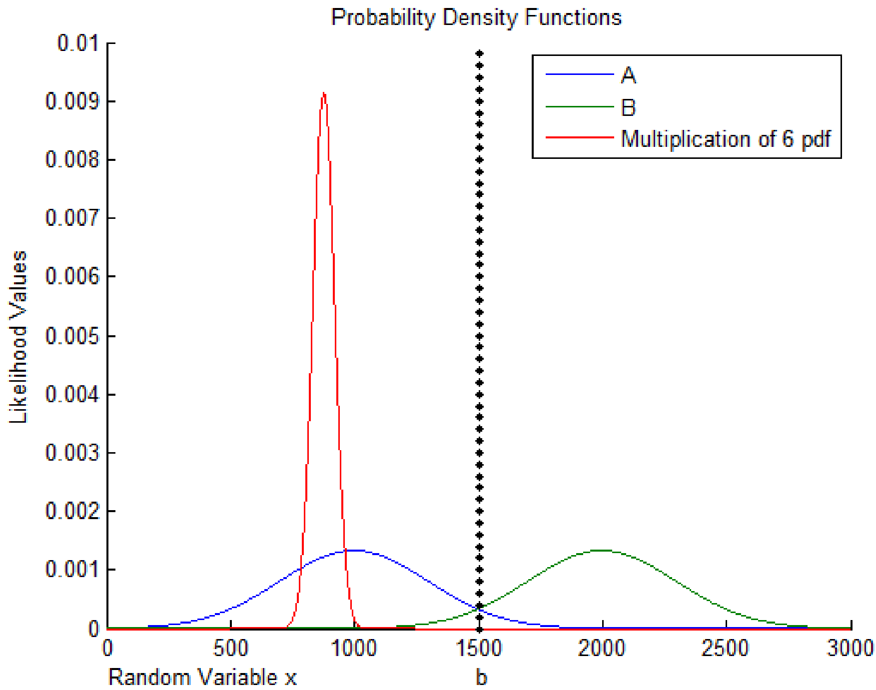


Figure 4.23: Measurement and the prior probability functions

Table 4.4: Combination results

Combination Rule	Probability Masses			
	$m(\{A\})$	$m(\{B\})$	$m(\{A,B\})$	$m(\{\phi\})$
Unnormalized masses	0.1489	0.0000	0.0000	0.8511
Dempster' rule	1.0000	0.0000	0.0000	0.0000
PCR6	0.8758	0.1242	0.0000	0.0000
Yager's rule	0.4552	0.0296	0.5152	0.0000
Proposed method	0.9989	0.0000	0.0011	0.0000

According to the results given in Table 4.4, the proposed method assigns high probability mass to class A like all other combination methods do; however, it assigns some probability mass to uncertainty contrary to Demspter's rule and PCR6 rule. According to the first interpretation given above this result is reasonable. There are five consistent measurements and they provide high probability mass to class A. On the other hand, there is one inconsistent measurement that indicates class B; but considering five consistent measurements this can be a wrong measurement, so the proposed method gives some credit to uncertainty instead of class B. If the interpretation is 'data is time varying' then again it is reasonable not to assign probability mass to class B immediately.

The second scenario is designed to see the effect of uncertain prior probability density functions. The prior probability density functions are selected very close to each other and the performance of the proposed method is analyzed under such uninformative prior cases. The prior and the measurement probability density functions are given below and are illustrated in Figure 4.24.

Prior probability density function of class A : $N(x; 2000, 300^2)$
Prior probability density function of class B : $N(x; 2001, 300^2)$
Probability density function of the 1^{st} measurement: $N(x; 1300, 100^2)$
Probability density function of the 2^{st} measurement: $N(x; 1290, 100^2)$
Probability density function of the 3^{st} measurement: $N(x; 1270, 100^2)$
Probability density function of the 4^{st} measurement: $N(x; 1250, 100^2)$
Probability density function of the 5^{st} measurement: $N(x; 2700, 200^2)$

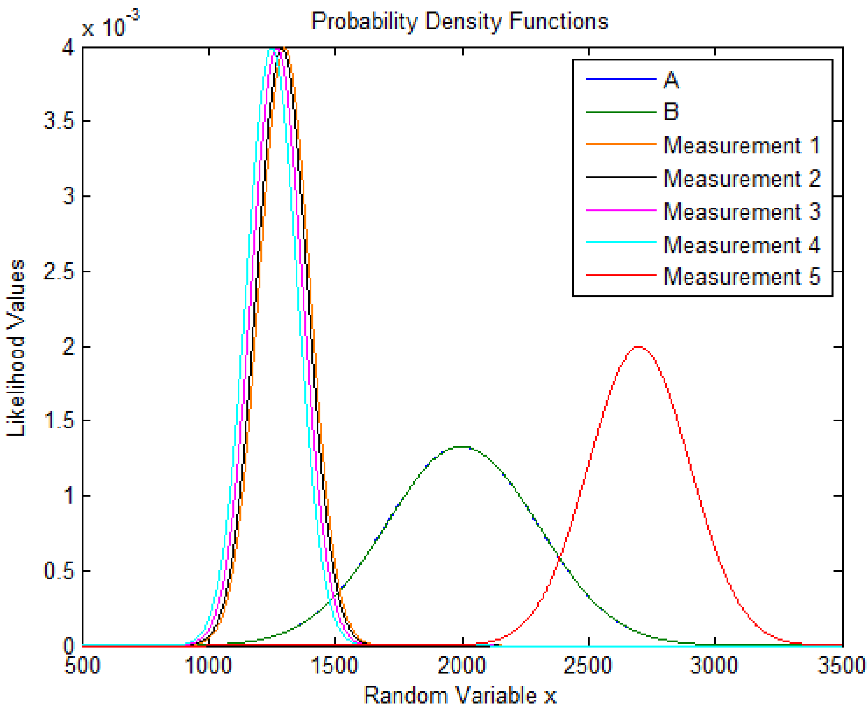


Figure 4.24: Measurement and the prior probability functions

The basic probability assignment for each measurement is given in Table 4.5 and the results of different combination rules are given in Table 4.6.

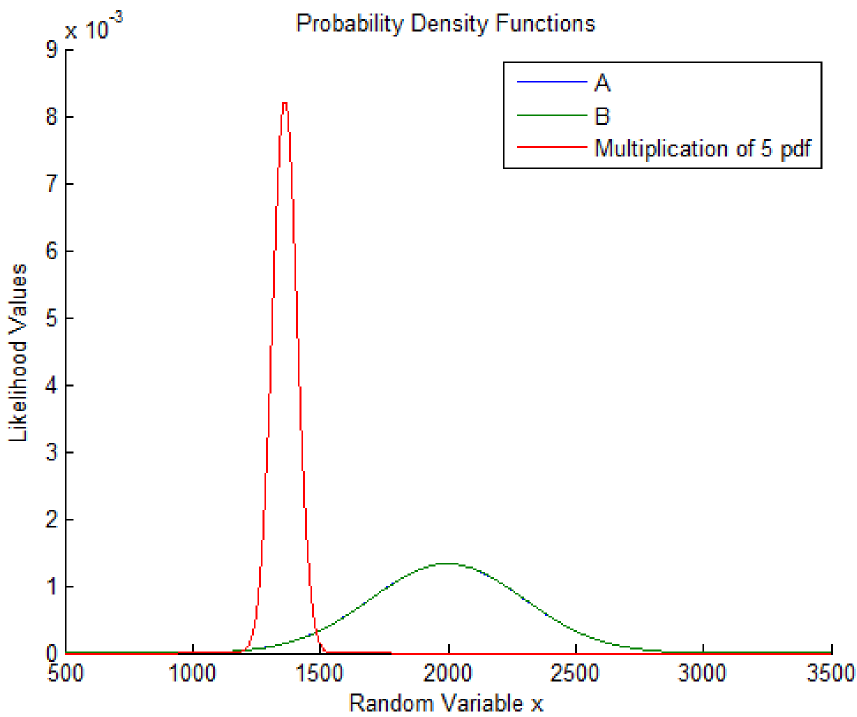


Figure 4.25: Measurement and the prior probability functions

Table 4.5: Basic Probability Assignments for the measurements

pdfs	Probability Masses		
	$m(\{A\})$	$m(\{B\})$	$m(\{A,B\})$
$N(x;1300,100^2)$	0.5017	0.0000	0.4983
$N(x;1290,100^2)$	0.5018	0.0000	0.4982
$N(x;1270,100^2)$	0.5018	0.0000	0.4982
$N(x;1250,100^2)$	0.5019	0.0000	0.4981
$N(x;2700,200^2)$	0.0009	0.5004	0.4987

Table 4.6: Combination results

Combination Rule	Probability Masses			
	$m(\{A\})$	$m(\{B\})$	$m(\{A,B\})$	$m(\{\phi\})$
Unnormalized masses	0.4688	0.0308	0.0307	0.4696
Dempster' rule	0.8840	0.0581	0.0579	0.0000
PCR6	0.6703	0.1248	0.2049	0.0000
Yager's rule	0.4688	0.0308	0.5322	0.0000
Proposed method	0.5017	0.0000	0.4983	0.0000

According to the results in Table 4.6, the proposed method assigns some probability mass to class A and some probability mass to uncertainty as all other combination methods do; however, it does not assign any probability mass to class B, contrary to all other rules. Intution says that this five consistent measurements should provide some credit to class A. But, the last measurement should not provide any credit to class B unless some other measurements strengthen this decision.

4.1.2.3 Experimental Results

The real fighter data that is used in Section 4.1.1.3 is utilized for comparing the methodology with other combination rules. The combination results are obtained only for speed information from initial time to final time. Tables 4.7, 4.8, 4.9, 4.10 and 4.11 show overall combination results for Unnormalized Dempster's rule, Dempster's rule, PCR6 rule, Yager's rule and the proposed method respectively.

Table 4.7: Unnormalized probability masses

$m(\{\phi\})$	$m(\{P\})$	$m(\{F\})$	$m(\{P,F\})$	$m(\{U\})$	$m(\{F,U\})$
0.8507	0.0205	0.1196	0.0091	0.0000	0.0000

Table 4.8: Dempster's rule

$m(\{\phi\})$	$m(\{P\})$	$m(\{F\})$	$m(\{P,F\})$	$m(\{U\})$	$m(\{F,U\})$
0.0000	0.1377	0.8016	0.0607	0.0000	0.0000

Table 4.9: PCR6 rule

$m(\{\phi\})$	$m(\{P,U\})$	$m(\{H,F\})$	$m(\{P,H,U\})$	$m(\{P,H,F\})$	$m(\{H,F,U\})$	$m(\{\Theta\})$
0.0000	0.0098	0.0093	0.3372	0.4062	0.2067	0.0308

Table 4.10: Yager's rule

$m(\{\phi\})$	$m(\{P\})$	$m(\{F\})$	$m(\{P,F\})$	$m(\{U\})$	$m(\{F,U\})$	$m(\{\Theta\})$
0.0000	0.0205	0.1196	0.0091	0.0000	0.0000	0.8507

Table 4.11: Proposed method

$m(\{F\})$	$m(\{P, F\})$
0.7795	0.2205

The results show that PCR6 rule and Yager's rule could not converge to a singleton set whereas Dempster's rule and the proposed method do. Dempster's rule gives some credit to set $\{P\}$, whereas the proposed method does not. The air vehicle slows down towards the end of the scenario. This causes increase in probability masses of the set $\{P\}$ for Dempster's rule. On the other hand, the proposed method increases the mass of the set $\{P, F\}$ instead of the set $\{P\}$, because there are enough consistent measurements that support fighter decision. It seems reasonable to increase the mass corresponding to the set $\{P, F\}$ until enough information comes that supports the set $\{P\}$.

4.1.2.4 Evaluation of the Combination Methodology

The proposed method can be evaluated as follows according to the general terms that are explained in section 2.3:

$$p_{final}(x) = \prod_{i=1}^{n} p_i(x)$$

✓ It satisfies the Commutativity property.

✓ It satisfies the Associativity property.

✓ For the total uninformative case it satisfies the Neutral Impact of VBA. However, it is hard to obtain such a condition for this method.

✓ It is coherent to possible cases and this can be seen in section 4.1.2.3.

× Reliability of the sources is not taken into consideration.

CHAPTER 5

A NOVEL METHODOLOGY FOR TARGET IDENTIFICATION

5.1 Target Identification Problem

Target Identification is another important concept for air defense systems like target classification. Air defense systems supply plenty of information to identify a target as friend, hostile or neutral. The information sources that are used in this study are as follows:

- Interrogation of friend or foe (IFF) - Mod 4
- Interrogation of friend or foe (IFF) - Mod 3
- Restricted area breach (RAB)
- Air corridor usage (ACU)
- Human-eye identification information (HII)

The information that is supplied by these sources is verbal. For example the information comes in the form 'The answer given to an IFF Mod-4 interrogation is not valid' or 'Target X is made a RAB'. Because of non-quantitative nature of the information as valid/invalid or all/nothing it cannot be used directly in a fusion system that uses numerical values. To fuse them, all information that comes from different types of sources should be converted by some mapping to a form that a decision making system can use. This can be possible by a rule-based inference system.

Rule-based systems are used in both target classification and target identification problems [13, 31]. A typical example for a rule-based system is given below [31]:

Example

Assume that $\Theta = \{f(friend), h(hostile), n(neutral)\}$

Rule: If a target responds to IFF correctly, then it is identified as friend. Basic probability assignments are made as follows:

$m(\{f\}) = 0.9$
$m(\{\Theta\}) = 0.1$

Note that in this example, even if the target responds to IFF correctly some mass value is assigned to the set Θ which would be useful when combining this information with another one, which may be conflicting. This type of basic probability assignment is quite coarse and is far from being objective.

Target identification problem can be considered as a multi-criteria decision making problem. We also want to assign probabilities to alternatives beyond making a decision. Besides, the information sources supply the information asynchronously. Hence, in this thesis we propose to use analytic hierarchy process for assigning probability masses in a distinct manner and to use Dempster-Shafer Theory for asynchronous information fusion.

5.2 Analytic Hierarchical Process for Target Identification based on Dempster-Shafer Theory

5.2.1 Hierarchy

Analytic Hierarchy Process (AHP) gives us the weights of events that can be used in mass assignment. To do that we have to first define a hierarchy. The hierarchy for target identification is a two level hierarchy that is illustrated in Figure 5.1.

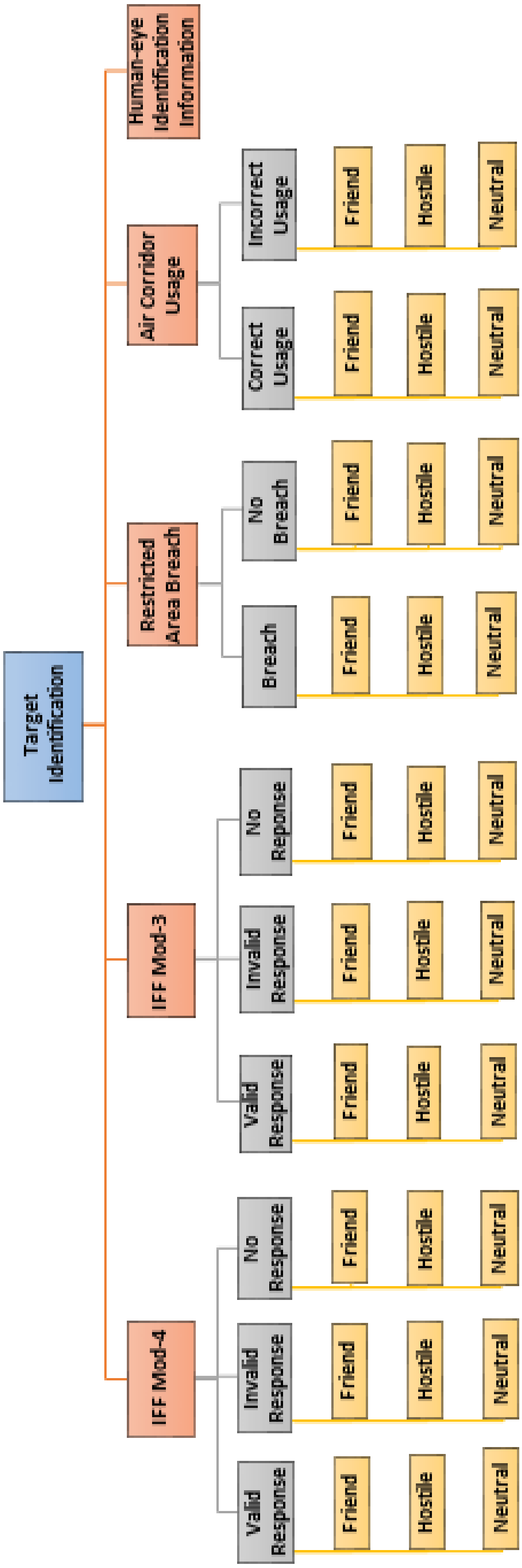


Figure 5.1: AHP Model for Target Identification

The first level of hierarchy defines all possible information sources and computes their relevance to the decision. The second level of hierarchy concentrates on final decision.

5.2.2 Pair-wise Comparison

Second step is the pair-wise comparison. Pair-wise comparison is applied to see which alternative is more effective on decisions. This aim is achieved by asking some questions to an expert group. For this purpose a questionnaire is built for evaluating every level of the hierarchy separately. The way of preparing and applying the questionnaire is new and is somehow different from the AHP studies given in the literature.

The questionnaire consists of two main parts:

1. Defining the order of importance of the information sources

2. Defining friend, foe (hostile) and neutral probability masses for every response of the information sources.

5.2.2.1 First Level: Order of Importance of the Information Sources

For pair-wise comparison it is asked which information source is more important, and how much it is important for identifying the target. The question of how much important is answered according to Table 5.1.

Every importance level shows how much more important the chosen alternative is compared with the other alternative. 2, 4, 6, 8 are the intermediate values. The question of which information source is more important is answered according to Table: 5.2

Table 5.1: Evaluation table for the Order of Importance of Information Sources

Importance Level	Explanation
1	Equally important
2	
3	More important
4	
5	Much more important
6	
7	Very much more important
8	
9	Absolutely more important

Table 5.2: Evaluation table for the Order of Information Sources

Order	Explanation
0	Both alternatives are equally important
1	First alternative is more important
2	Second alternative is more important

In order to explain how the questions are answered, three generic examples are given below.

Question:

	Which one?	How much?
Considering IFF Mod-4 and IFF Mod-3, which one of these two sources is more important and how much?		

Answer: If one thinks that IFF Mod-4 is absolutely more important than the IFF Mod-3, one should answer the above question as follows:

	Which one?	How much?
Considering IFF Mod-4 and IFF Mod-3, which one of these two sources is more important and how much?	1	9

"1" in "which one" column shows that the first alternative IFF Mod-4 is more important than the second alternative IFF Mod-3. "9" in "how much" column shows that the chosen alternative is absolutely more important than the other alternative.

Question:

	Which one?	How much?
Considering air corridor usage information and human-eye identification information, which one of these two sources is more important and how much?		

Answer: If one thinks that human-eye identification information is much more important than the air corridor usage information, one should answer the above question as follows:

	Which one?	How much?
Considering air corridor usage information and human-eye identification information, which one of these two sources is more important and how much?	2	5

"2" in "which one" column shows that the second alternative human-eye identification information is more important than the first alternative air corridor usage information. "5" in "how much" column shows that the chosen alternative is much more important than the other alternative.

Question:

	Which one?	How much?
Considering air corridor usage information and human-eye identification information, which one of these two sources is more important and how much?		

Answer: If one thinks that both of the alternatives are equally important, one should answer the above question as follows:

	Which one?	How much?
Considering air corridor usage information and human-eye identification information, which one of these two sources is more important and how much?	0	1

"0" in which one column and "5" in how much column show that the both of the alternatives are equally important.

The Questionnaire

The questionnaire built for the first level of the hierarchy is given in Table 5.3 below.

Table 5.3: The questionnaire for built for the first level of the hierarchy

	Which one?	How much?
Considering IFF Mod-4 and IFF Mod-3, which one of these two sources is more important and how much?		
Considering IFF Mod-4 and restricted area breach information, which one of these two sources is more important and how much?		
Considering IFF Mod-4 and air corridor usage information, which one of these two sources is more important and how much?		
Considering IFF Mod-4 and human-eye identification information, which one of these two sources is more important and how much?		
Considering IFF Mod-3 and restricted area breach information, which one of these two sources is more important and how much?		
Considering IFF Mod-3 and air corridor usage information, which one of these two sources is more important and how much?		
Considering IFF Mod-3 and human-eye identification information, which one of these two sources is more important and how much?		
Considering restricted area breach information and air corridor usage information, which one of these two sources is more important and how much?		
Considering restricted area breach information and human-eye identification information, which one of these two sources is more important and how much?		
Considering air corridor usage information and human-eye identification information, which one of these two sources is more important and how much?		

5.2.2.2 Second Level: Friend, foe (hostile) and neutral probability mass assignments

To assign masses to classes, pairwise comparison is applied again with a new set of questions. The structure of the questions is also changed. Instead of asking which alternative is more important now we ask which event is more probable. The quantization levels of the answers of the "how much probable" part is done according to Table 5.4:

Table 5.4: Evaluation table for defining friend, foe and neutral probability masses

Probability Level	Explanation
1	Equally probable
2	
3	More probable
4	
5	Much more probable
6	
7	Very much more probable
8	
9	Absolutely more probable

Every probability level shows that how much the chosen event is more probable than the other event. 2, 4, 6, 8 are the intermediate values. The question of which event is more probable is answered according to Table 5.5

Table 5.5: Evaluation table for the Order of Information Sources

Order	Explanation
0	Both events are equally probable
1	First event is more probable
2	Second event is more probable

In order to explain how the questions are answered, three generic examples are given below.

Question:

	Which one?	How much?
If a target responds IFF Mod-4 correctly, which identity is more probable among being friend or hostile? How much?		

Answer: If one thinks that target is more probable to be a friend, because of the target's correct response to IFF Mod-4, one should answer the above question as follows:

	Which one?	How much?
If a target responds IFF Mod-4 correctly, which identity is more probable among being friend or hostile? How much?	1	7

"1" in which one column shows that the first identity friend is more probable than the second identity hostile. "7" in how much column shows that the chosen event is very much more probable than the other event.

Question:

	Which one?	How much?
If a target responds IFF Mod-4 correctly, which identity is more probable among being hostile or neutral? How much?		

Answer: If one thinks that target is more probable to be a neutral, because of the target's correct response to IFF Mod-4, one should answer the above question as follows:

	Which one?	How much?
If a target responds IFF Mod-4 correctly, which identity is more probable among being hostile or neutral? How much?	2	5

"2" in which one column shows that the second identity neutral is more probable than the first identity. "5" in how much column shows that the chosen event is much more probable than the other event.

Question:

	Which one?	How much?
If a target responds IFF Mod-4 correctly, which identity is more probable among being friend or neutral? How much?		

Answer: If one thinks that both of the alternatives are equally important, one should answer the above question as follows:

	Which one?	How much?
If a target responds IFF Mod-4 correctly, which identity is more probable among being friend or neutral? How much?	0	1

"0" in which one column and "1" in how much column show that the both of the events are equally probable.

A questionnaire is built for the second level of the hierarchy and it is given in Appendix A.

5.2.2.3 Weight Calculation

Weight calculation is explained by using a generic example:

Example: Suppose that pair-wise comparison matrix A for the first hierarchy level is obtained as follows:

Table 5.6: Pair-wise comparison matrix for information sources

	IFF Mod-4	IFF Mod-3	RAB	ACU	HII
IFF Mod-4	1	5	9	9	6
IFF Mod-3	$1/5$	1	5	7	6
RAB	$1/9$	$1/5$	1	9	3
ACU	$1/9$	$1/7$	$1/9$	1	$1/3$
HII	$1/9$	$1/6$	$1/3$	3	1

After building the pair-wise comparison matrix, the first step is summing all column elements as shown in last row of Table 5.7. Smaller sum indicates that this source is more important compared to the other sources in determining the target identity.

Table 5.7: Pair-wise comparison matrix for information sources after the first step

	IFF Mod-4	IFF Mod-3	RAB	ACU	HII
IFF Mod-4	1	5	9	9	6
IFF Mod-3	$1/5$	1	5	7	6
RAB	$1/9$	$1/5$	1	9	3
ACU	$1/9$	$1/7$	$1/9$	1	$1/3$
HII	$1/9$	$1/6$	$1/3$	3	1
	1.5333	6.5095	15.4444	29.0000	16.3333

Later, all column elements are normalized according to the sums in Table 5.7 and the normalized values are obtained as in Table 5.8. Normalization gives large values to important sources.

Table 5.8: Normalized Pair-wise comparison matrix for information sources

	IFF Mod-4	IFF Mod-3	RAB	ACU	HII
IFF Mod-4	0.6522	0.7681	0.5827	0.3103	0.3673
IFF Mod-3	0.1304	0.1536	0.3237	0.2414	0.3673
RAB	0.0725	0.0307	0.0647	0.3103	0.1837
ACU	0.0725	0.0219	0.0072	0.0345	0.0204
HII	0.0725	0.0256	0.0216	0.1034	0.0612
	1	1	1	1	1

Finally, the weights are found by calculating the arithmetic mean of each row as shown in last column of Table 5.9 that gives an average importance value.

Table 5.9: Normalized Pair-wise comparison matrix for information sources with weights

	IFF Mod-4	IFF Mod-3	RAB	ACU	HII	
IFF Mod-4	0.6522	0.7681	0.5827	0.3103	0.3673	0.5361
IFF Mod-3	0.1304	0.1536	0.3237	0.2414	0.3673	0.2433
RAB	0.0725	0.0307	0.0647	0.3103	0.1837	0.1324
ACU	0.0725	0.0219	0.0072	0.0345	0.0204	0.0313
HII	0.0725	0.0256	0.0216	0.1034	0.0612	0.0569
	1	1	1	1	1	1

5.2.3 Basic Probability Assignment for Target Identification

In order to calculate probability masses, firstly, the prepared questionnaires are applied to a small group that consists of 4 people. Pair-wise comparison matrices give the order of importance and the raw probability masses for each information source response. Final order of importance and raw probability masses are calculated using arithmetic mean method. Order of importance values are given in Table 5.10

Table 5.10: Final order of importance values of the information sources

IFF Mod-4	0.6297
IFF Mod-3	0.1966
RAB	0.0610
ACU	0.0636
HII	0.0490

The first level of AHP gives the weights of all sources. The probability mass of each class is obtained by considering the two levels together. The pair-wise comparison matrices of the second level are obtained again by the answers of 4 people. The interpretation of the matrices is similar to the first level. Applying same operations raw probability masses of each information source response are obtained. Results are given in Table 5.11. Finally the raw probability masses are multiplied by the corresponding weights given in Table 5.10 to obtain final probability masses.

As an example final probability mass calculation is given for one information source response. Rest of the probability mass calculation is trivial.

To find the probability masses for the target that gives valid response to IFF Mod-4, firstly the raw probability masses for IFF Mod-4 valid response are multiplied with the importance level of IFF Mod-4.

$$0.6297. \begin{bmatrix} 0.7668 \\ 0.0985 \\ 0.1347 \end{bmatrix} = \begin{bmatrix} 0.4829 \\ 0.0620 \\ 0.0848 \end{bmatrix}$$

and probability masses are obtained as follows:

Friend	0.4829
Hostile	0.0620
Neutral	0.0848

After calculating probability masses for the pre-defined identities, mass of set Θ is computed as follows:

$$\begin{aligned} m(\{\Theta\}) &= 1 - (0.4829 + 0.0620 + 0.0848) \\ &= 0.3703 \end{aligned}$$

These steps are repeated for all responses to calculate the final probability masses. Final probability masses are given in Table 5.12.

Table 5.11: Raw probability masses

	IFF Mod-4			IFF Mod-3			Restrcited Area Breach		Air Corridor Usage	
	Valid Response	Invalid Response	No Response	Valid Response	Invalid Response	No Response	Breach	No breach	Correct Usage	Incorrect Usage
Friend	0.7668	0.2115	0.1875	0.6806	0.1410	0.2330	0.0806	0.6108	0.6774	0.2647
Hostile	0.0985	0.5091	0.4953	0.0933	0.5426	0.4781	0.6476	0.1481	0.0994	0.4902
Neutral	0.1347	0.2794	0.3172	0.2261	0.3165	0.2889	0.2718	0.2412	0.2233	0.2451

Table 5.12: Final probability masses

	IFF Mod-4			IFF Mod-3			Restrcited Area Breach		Air Corridor Usage		Human-eye Identification Information
	Valid Response	Invalid Response	No Response	Valid Response	Invalid Response	No Response	Breach	No breach	Correct Usage	Incorrect Usage	
Friend	0.4829	0.1332	0.1181	0.1338	0.0277	0.0458	0.0049	0.0373	0.0431	0.0168	0.0490
Hostile	0.0620	0.3206	0.3119	0.0183	0.1067	0.0940	0.0395	0.0090	0.0063	0.0312	
Neutral	0.0848	0.1759	0.1997	0.0445	0.0622	0.0568	0.0166	0.0147	0.0142	0.0156	
Unknown	0.3703	0.3703	0.3703	0.8034	0.8034	0.8034	0.9390	0.9390	0.9364	0.9364	0.9510

5.2.4 Analysis of the Novel Basic Probability Assignment Methodology for Target Identification

The method is firstly analyzed to observe the reactions of the system to different outcomes of interrogations or other clues of being friend, hostile or neutral. 5 different experiments were conducted for this purpose. The number of interrogations for IFF is chosen as 10 since air defense radars make 10 interrogations for each target. Restricted area breach, air corridor usage and human-eye identification information are used at different rates. Dempster's rule is used for information fusion. Note that it satisfies the commutativity property so the order of the responses or the order of the information sources is not important.

The second part of the analysis section is about the convergence rates and decision rates of each source. 12 experiments were conducted for this purpose.

5.2.4.1 The First Group of Experiments

For the first scenario, 10 IFF interrogations were made. 5 interrogations ended up with valid response and 5 interrogations ended up with no response. Basic probability assignments were mainly shaped according to the these responses. This result seems reasonable considering IFF Mod-4 is the most reliable information source according to the answers of the Analytic Hierarchy Process questionnaire that is given in Table 5.3.

The first scenario gives probability of being friend prominence, whereas the second scenario gives probability of being hostile prominence. At the beginning of the third scenario, getting no response from IFF interrogations causes an increase in the mass of being hostile; however, incoming valid answer to the following interrogations changes the decision from hostile to friend. Scenario 4 shows how the probability masses are assigned, when two of the three situations (valid response/invalid response/no response) of IFF reveal.

Scenario 1:

Probability masses are computed as shown in Figure 5.2, with the following responses from the information sources.

- 5 times **IFF Mod-4** interrogations ended up with valid response.
- 1 time **Restricted Areas** are breached.
- 4 times **Air Corridors** are used correctly.
- 5 **IFF Mod-4** interrogations ended up with no response.
- 3 times **Restricted Areas** are not breached.

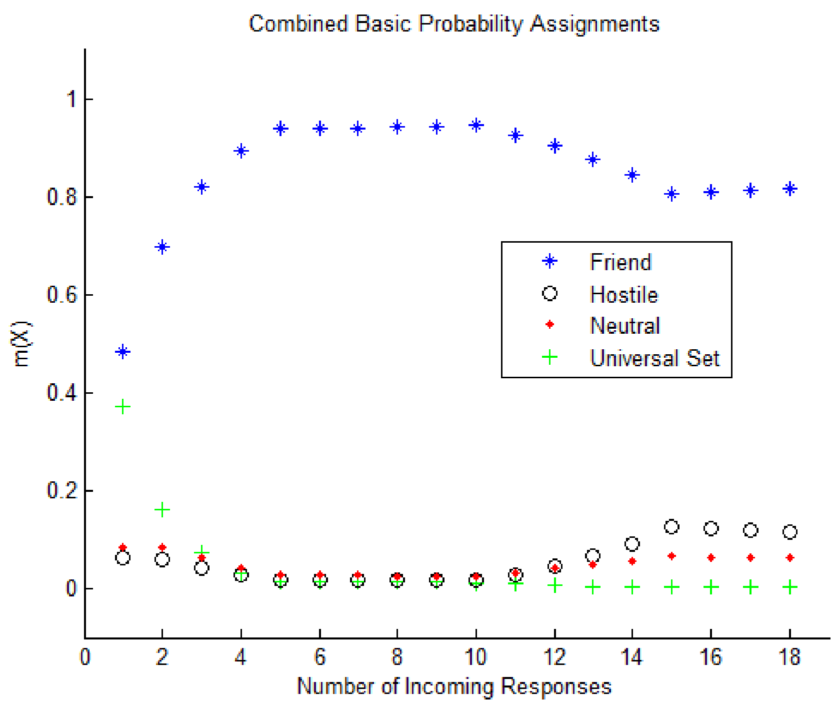


Figure 5.2: Combined basic probability assignments for the above responses

Scenario 2:

Probability masses are computed as shown in Figure 5.3, with the following responses from the information sources.

- 5 times **IFF Mod-4** interrogations ended up with invalid response.
- 2 times **Restricted Areas** are not breached.
- 4 times **Air Corridors** are used correctly.
- 5 **IFF Mod-4** interrogations ended up with no response.
- 2 times **Restricted Areas** are breached.

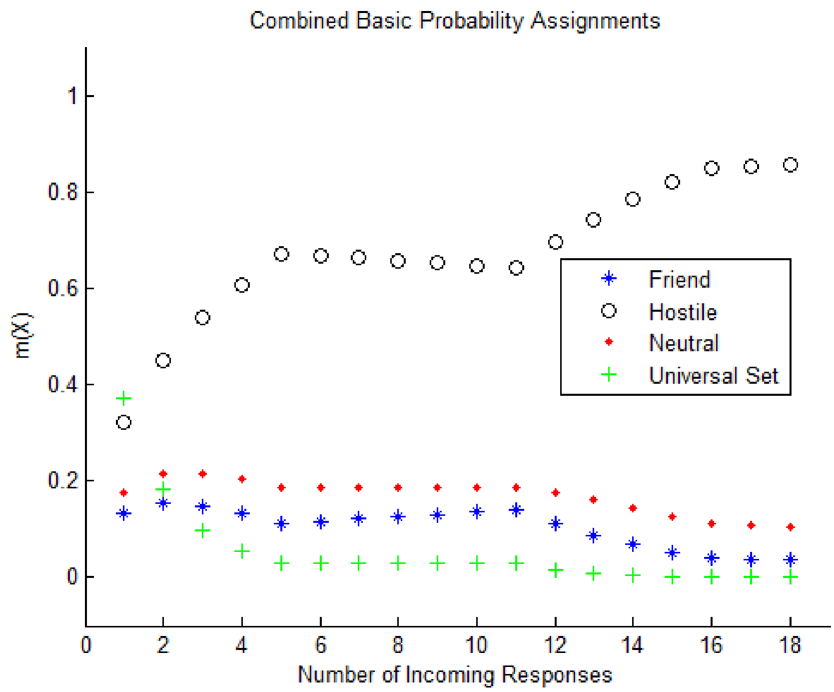


Figure 5.3: Combined basic probability assignments for the above responses

Scenario 3:

Probability masses are computed as shown in Figure 5.4, with the following responses from the information sources.

- 5 times **IFF Mod-4** interrogations ended up with no response.
- 2 times **Restricted Areas** are not breached.
- 4 times **Air Corridors** are used correctly.
- 5 **IFF Mod-4** interrogations ended up with valid response.

- 2 times **Restricted Areas** are breached.

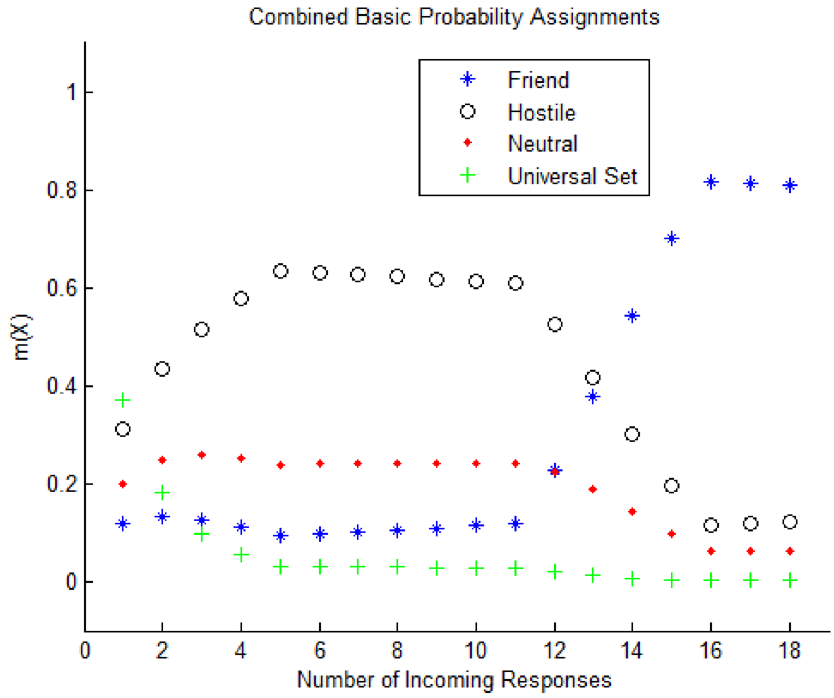


Figure 5.4: Combined basic probability assignments for the above responses

Scenario 4:

Probability masses are computed as shown in Figure 5.5, with the following responses from the information sources.

- 5 times **IFF Mod-4** interrogations ended up with invalid response.
- 2 times **Restricted Areas** are not breached.
- 4 times **Air Corridors** are used correctly.
- 3 **IFF Mod-4** interrogations ended up with valid response.
- 2 **IFF Mod-4** interrogations ended up with no response.
- 2 times **Restricted Areas** are breached.

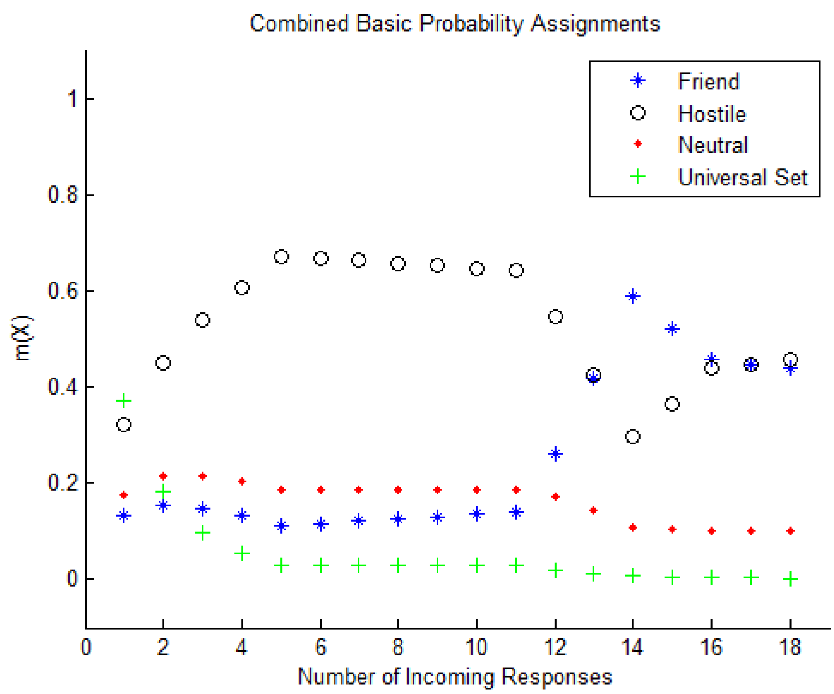


Figure 5.5: Combined basic probability assignments for the above responses

Scenario 5:

Probability masses are computed as shown in Figure 5.6, with the following responses from the information sources.

- 3 times **IFF Mod-4** interrogations ended up with invalid response.
- 2 times identified as neutral according to the **Human-eye Identification Information**.
- 2 times **Restricted Areas** are not breached.
- 4 times **Air Corridors** are used correctly.
- 3 **IFF Mod-4** interrogations ended up with valid response.
- 4 **IFF Mod-4** interrogations ended up with no response.
- 2 times **Restricted Areas** are breached.

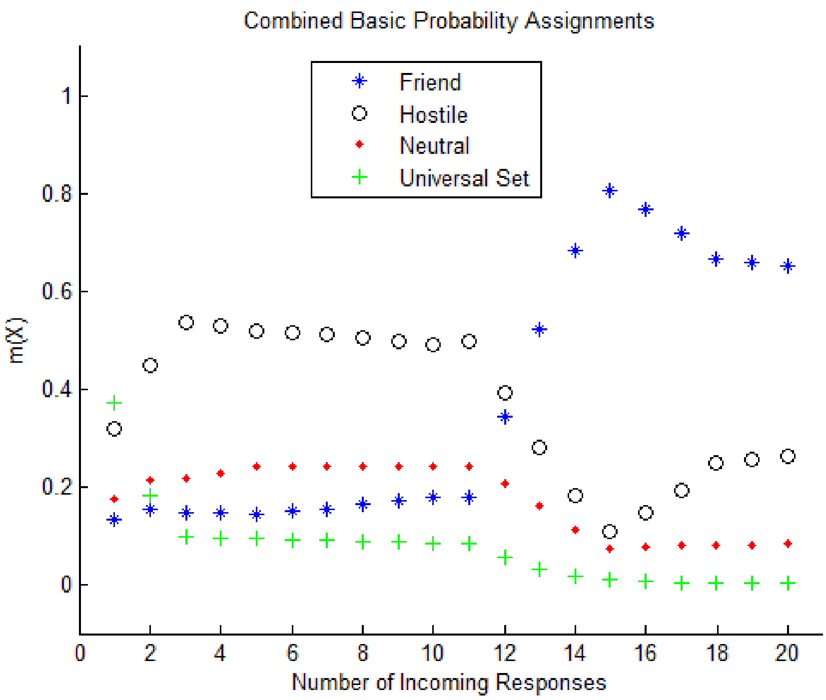


Figure 5.6: Combined basic probability assignments for the above responses

5.2.4.2 The Second Group of Experiments

The aim of the second group of experiments is to evaluate the influence of each response of each information source. For this purpose, it is assumed that there is only one information source and it gives the same response at each interrogation. The time of converging to the indicated identity with a high probability is analyzed.

Human-eye identification information is in the form of all/nothing, so order of importance of human-eye identification information is directly used for assigning the probability mass of the incoming identity.

The convergence speed to friend and hostile decisions are presented in Tables 5.13 and 5.14 respectively.

Table 5.13: Convergences speed for friend decision

IFF Mod-4 - valid response	8
IFF Mod-3 - valid response	48
RAB - no breach	207
ACU - correct usage	159
HII - friend	92

Table 5.14: Convergences speed for hostile decision

IFF Mod-4 - invalid response	21
IFF Mod-4 - no response	26
IFF Mod-3 - invalid response	93
IFF Mod-3 - no response	114
RAB - breach	197
ACU - incorrect usage	341
HII - hostile	92

The second columns of the tables show the number of responses in order to converge to 0.9999 for the corresponding information source responses.

The results show that

- The decision of friend is given quicker than the decision of hostile, as expected for all information sources.

- The convergence speed is proportional to order of importance of the information sources, as expected for all responses.

Figures from 5.7 to 5.18 give more information about the rate of convergence.

IFF Mod-4

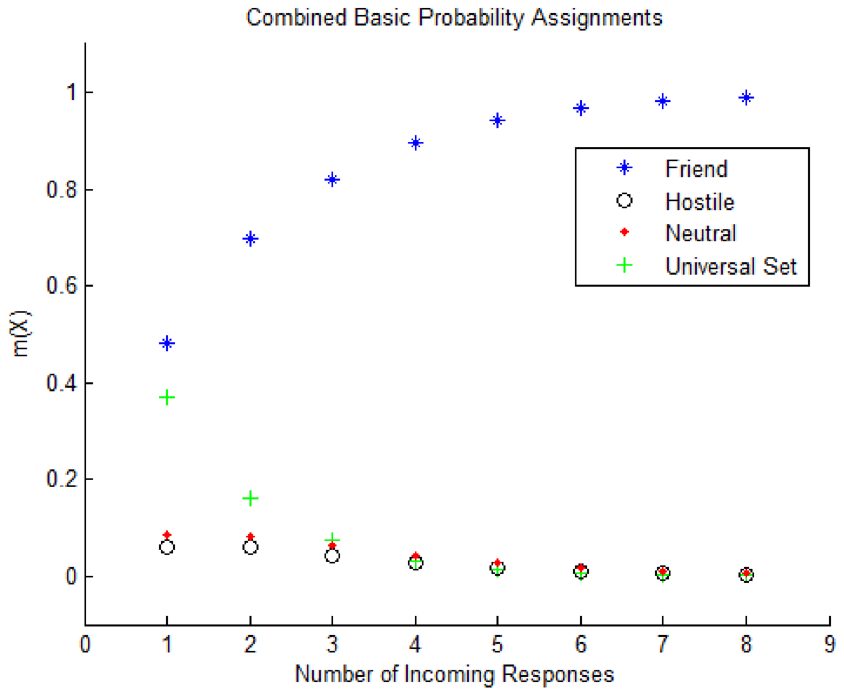


Figure 5.7: IFF Mod-4 valid response

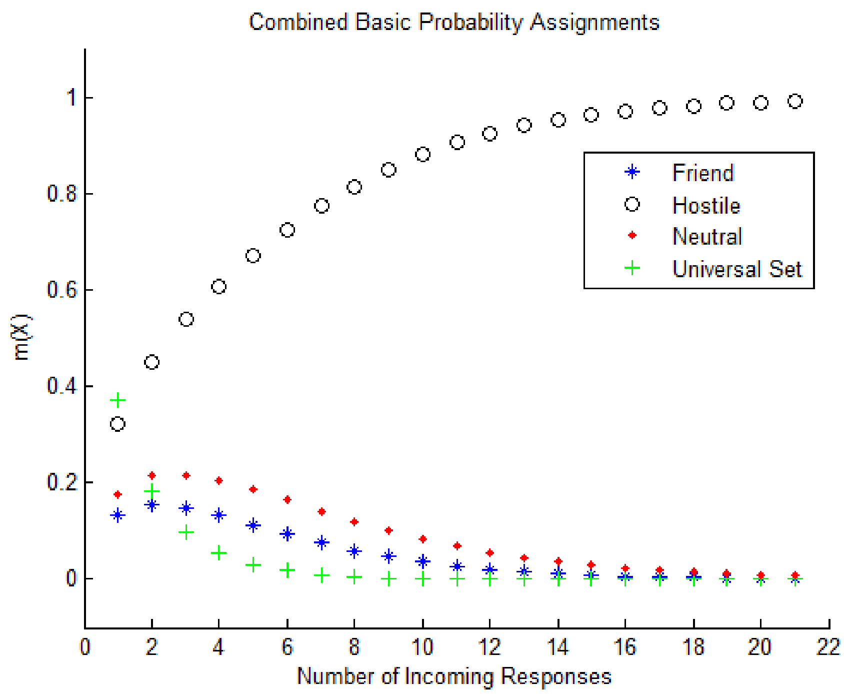


Figure 5.8: IFF Mod-4 invalid response

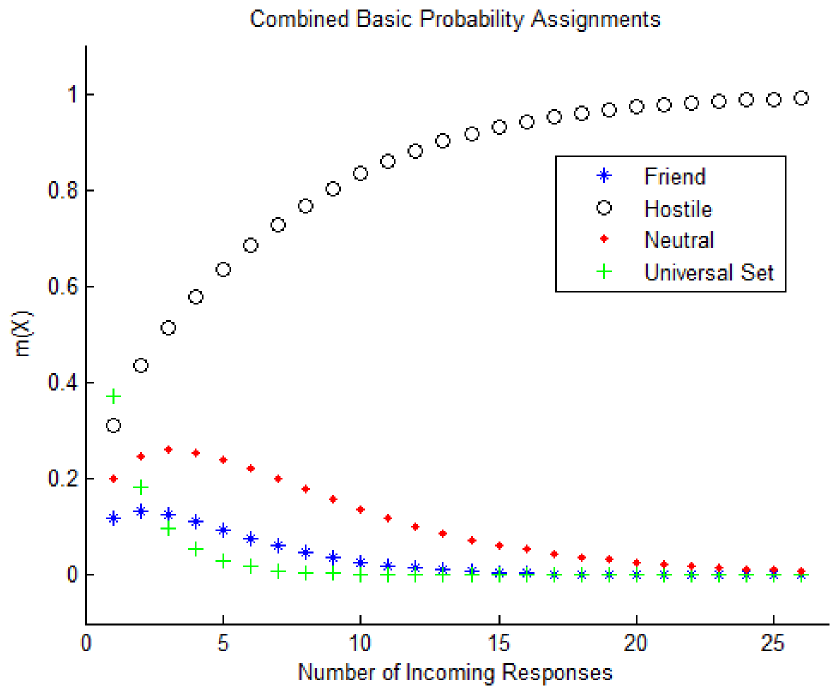


Figure 5.9: IFF Mod-4 invalid response

IFF Mod-3

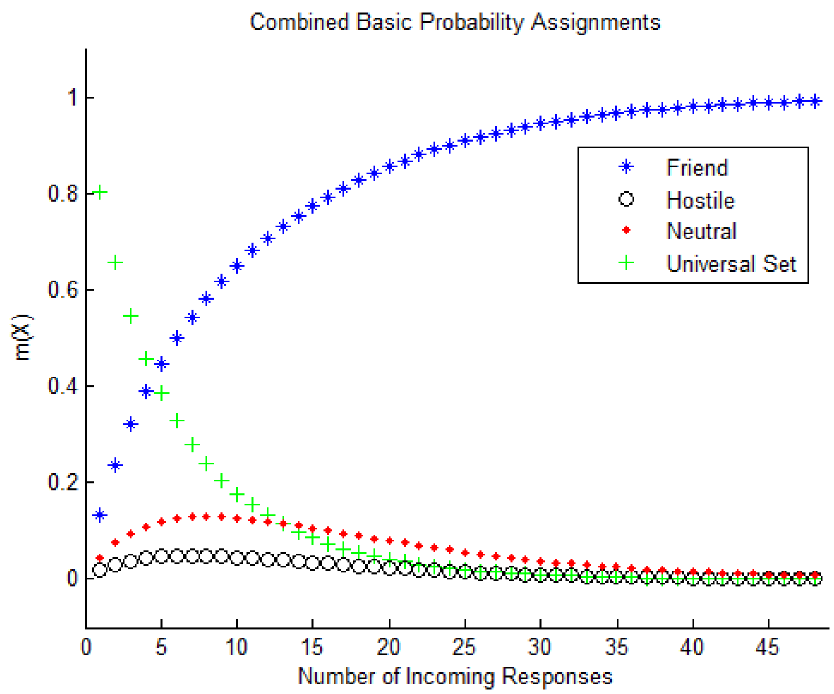


Figure 5.10: IFF Mod-3 valid response

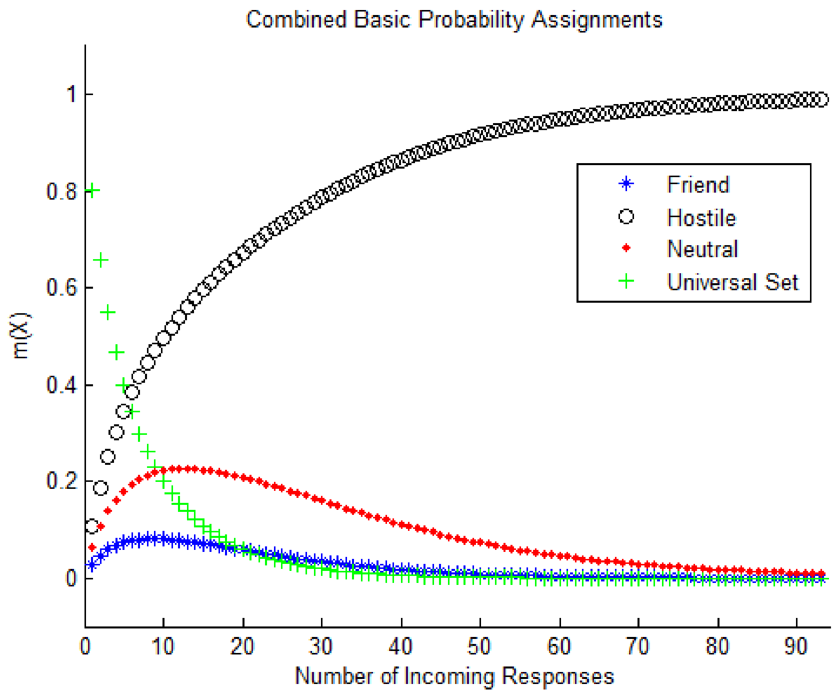


Figure 5.11: IFF Mod-3 invalid response

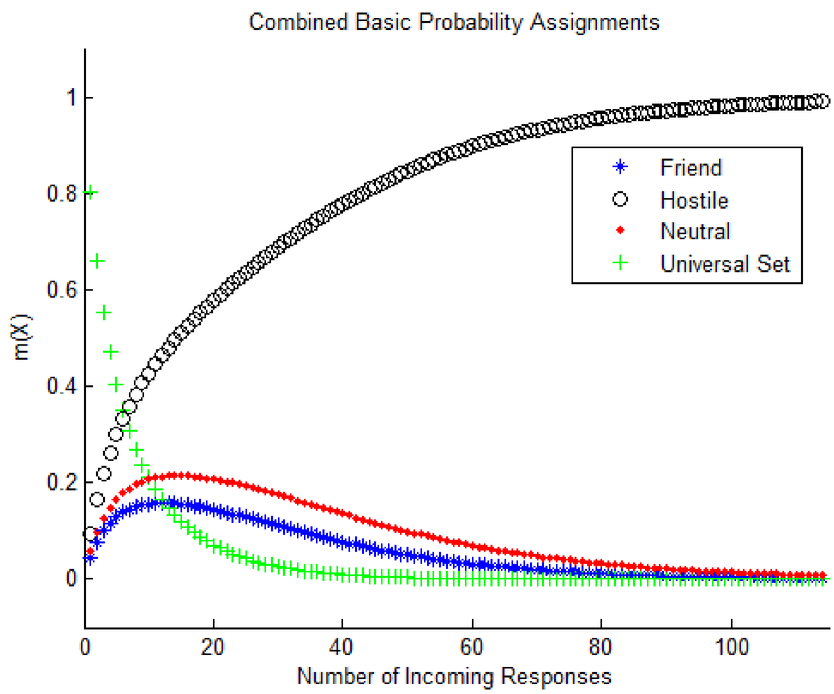


Figure 5.12: IFF Mod-3 invalid response

Restricted Area Breach

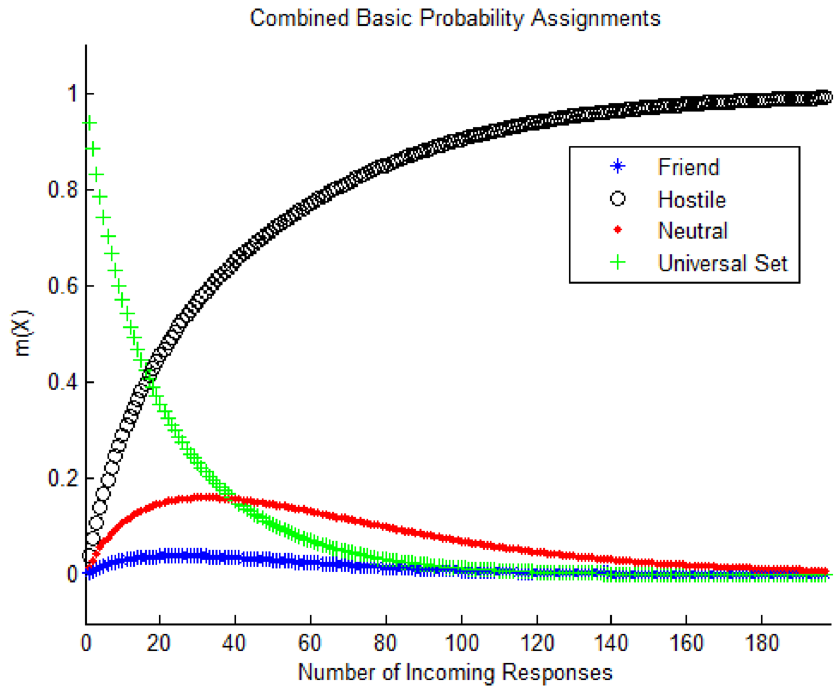


Figure 5.13: Restricted Area Breach - Breach

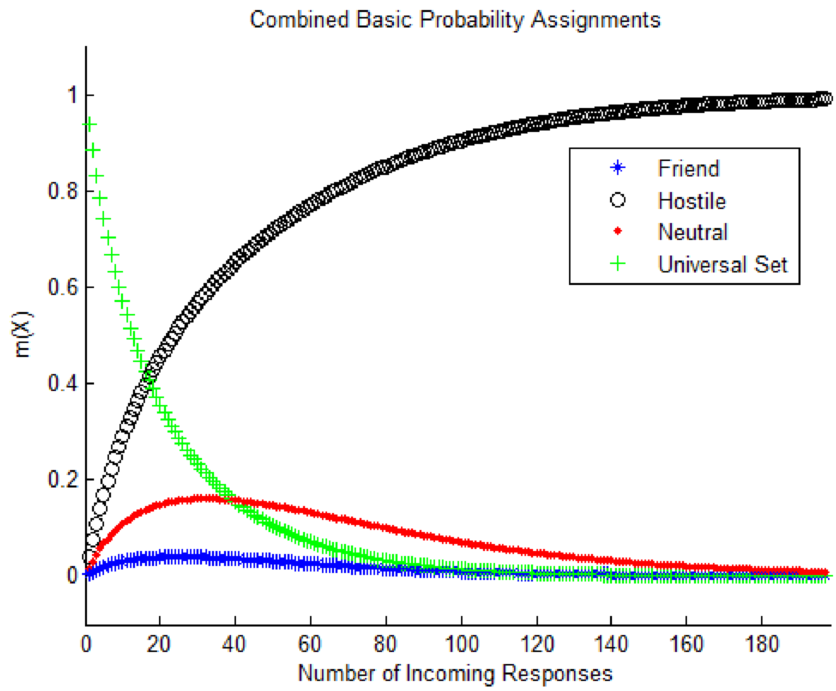


Figure 5.14: Restricted Area Breach - No Breach

Air Corridor Usage

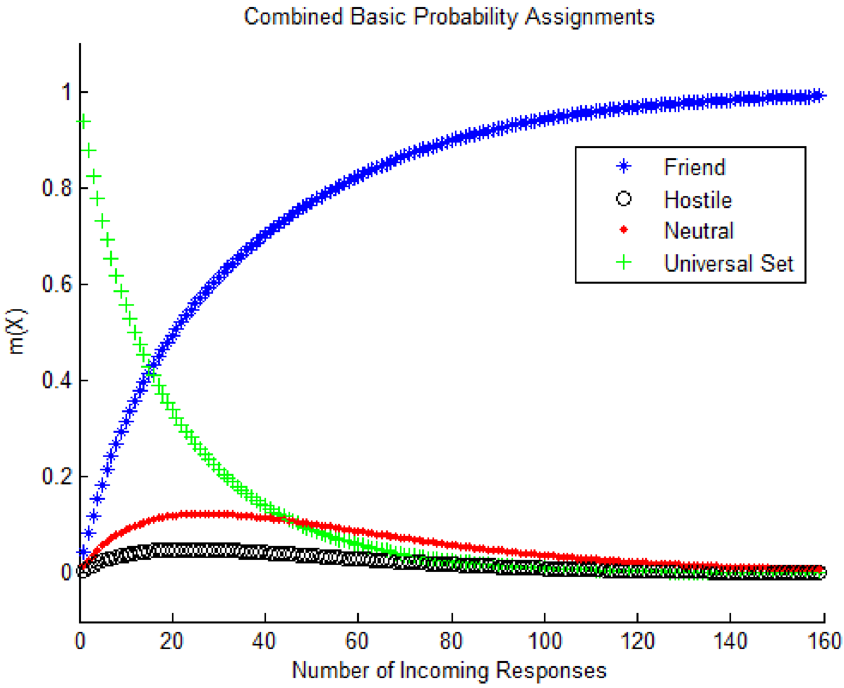


Figure 5.15: Air Corridor Usage - Correct

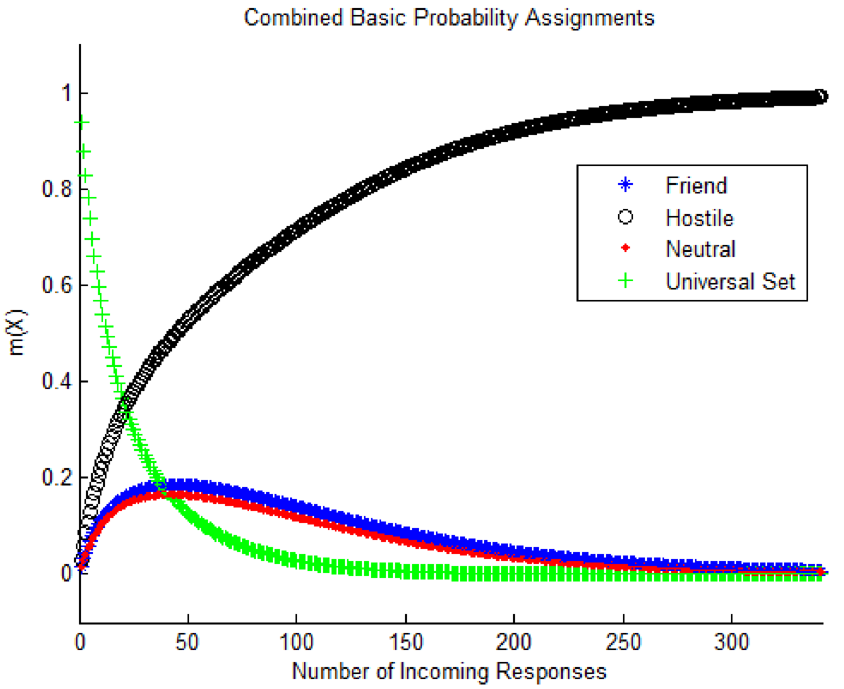


Figure 5.16: Air Corridor Usage - Incorrect

Human-eye Identification Information

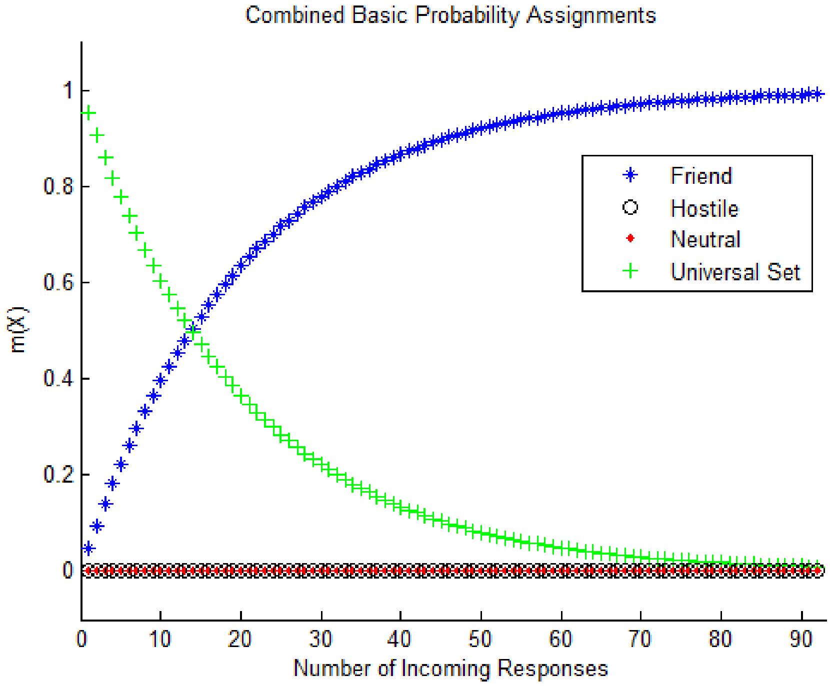


Figure 5.17: Human-eye Identification - Friend

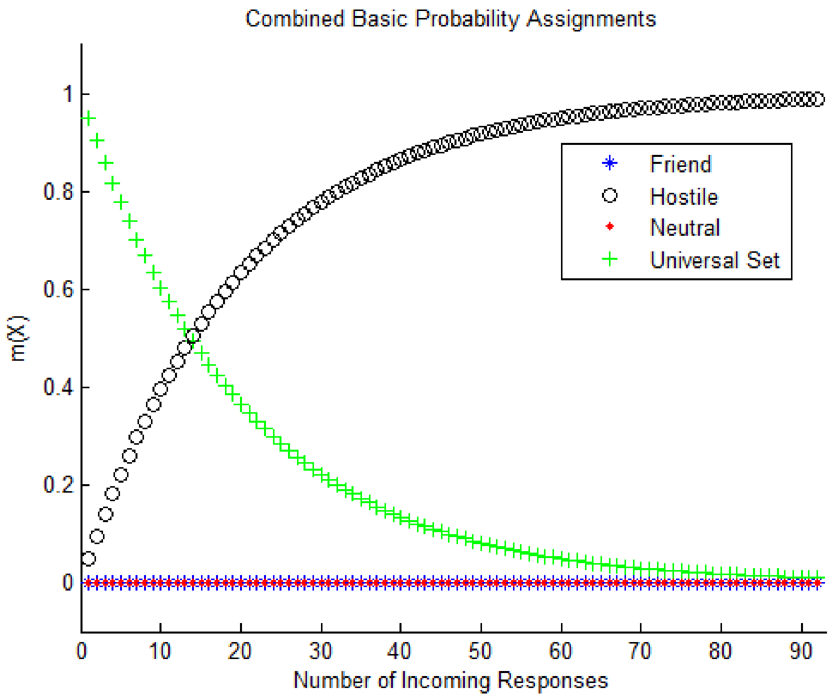


Figure 5.18: Human-eye Identification - Hostile

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

In this thesis target classification and target identification problems, which have close relationship with target tracking, are concerned. In this respect we have concentrated on air vehicles.

Classification means the determination of target type as bomber and surveillance plane, helicopter, fighter or UAV and it is made by using the probability density function of the state provided by the tracker. On the other hand, identification is the determination of a target as friend, hostile or neutral and it is made by using IFF interrogations, restricted area breach, air corridor usage and human-eye identification information.

Both of the problems require the combination of evidences as they are collected in time and from different sources. Dempster-Shafer Theory is utilized for this purpose. Dempster-Shafer Theory requires mass assignment to the events for combination of evidence. For this purpose we propose different methodologies in sense of target classification and identification.

The methodology proposed for target classification utilizes all information provided by the tracker which is given in the form of a Gaussian probability density function, unlike the alternative methods in the literature which use only the mean. The proposed method for target classification is tested with real data. Results are compared with another method that uses only the mean estimates of the features that can be considered quite similar to the one given in Xu et. al [42]. Test results show that using the whole probability density function of the features provided by the tracker brings significant advantages for classification compared to using only the mean of the features.

The proposed fusion method is also inspired from the idea of using all available information and preventing information loss. For this purpose probability density functions of the measurements are combined before converting them into probability masses. That gives us a new combination method. The methodology is tested with real data and compared with the major combination rules existing in the literature. Test results show that the proposed method gives reasonable results for the cases that probability density function of the state is available.

The proposed method for target identification utilizes all available information as all other algorithms in the literature do. However, differing from the existing algorithms, it uses Analytic Hierarchy Process for basic mass assignment. In the proposed methodology the evidences are evaluated according to their importance. In order to obtain importance of the information sources and for mass assignment according to this importances level, Analytic Hierarchy Process is adapted to the target identification problem. The adaptation and the way of mass assignment are novel in the literature and constitute another main contribution of this thesis.

6.2 Future Work

In this thesis, for classification, features are selected according to their reliability and direct availability from the tracker. Other than the speed and the altitude, acceleration and micro Doppler can be added into the feature vector to improve the performance of the target classifier. In order to do that related prior information should be collected for the predefined target types.

The probability density function of the speed is obtained by using Monte Carlo methods from the velocity vector as a non-parametric density. Unscented transform is a way for estimating a nonlinear function of a Gaussian random variable as a Gaussian. This approach may be used to obtain the probability density function of the speed.

The questionnaires that are constituted for target identification were applied to a small group. Applying these questionnaires to large expert groups and including electronic support measures to the information sources will improve identification performance.

REFERENCES

[1] Janos Aczél and Thomas L Saaty. Procedures for synthesizing ratio judgements. *Journal of mathematical Psychology*, 27(1):93–102, 1983.

[2] Edit Adamcsek. The analytic hierarchy process and its generalizations. *Unpublished doctoral dissertation, Thesis, Eotvos Lor and Un iversit y*, 2008.

[3] Donka Angelova and Lyudmila Mihaylova. Joint target tracking and classification with particle filtering and mixture kalman filtering using kinematic radar information. *Digital Signal Processing*, 16(2):180–204, 2006.

[4] Mahmood R Azimi-Sadjadi, De Yao, Qiang Huang, and Gerald J Dobeck. Underwater target classification using wavelet packets and neural networks. *Neural Networks, IEEE Transactions on*, 11(3):784–794, 2000.

[5] Yaakov Bar-Shalom. Multitarget-multisensor tracking: Applications and advances. volume iii. *Norwood, MA, Artech House, Inc., 2000*, 2000.

[6] Indrani Basak and Thomas Saaty. Group decision making using the analytic hierarchy process. *Mathematical and computer modelling*, 17(4):101–109, 1993.

[7] Valerie Belton and Tony Gear. On a short-coming of Saaty's method of analytic hierarchies. *Omega*, 11(3):228–230, 1983.

[8] Azzedine Bendjebbour, Yves Delignon, Laurent Fouque, Vincent Samson, and Wojciech Pieczynski. Multisensor image segmentation using dempster-shafer fusion in markov fields context. *Geoscience and Remote Sensing, IEEE Transactions on*, 39(8):1789–1798, 2001.

[9] Samuel Blackman and Robert Popoli. Design and analysis of modern tracking systems. *Boston, MA: Artech House*, 1999.

[10] Isabelle Bloch. Some aspects of dempster-shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. *Pattern Recognition Letters*, 17(8):905–919, 1996.

[11] Philip L Bogler. Shafer-dempster reasoning with applications to multisensor target identification systems. *Systems, Man and Cybernetics, IEEE Transactions on*, 17(6):968–977, 1987.

[12] Richard R Brooks, Parameswaran Ramanathan, and Akbar M Sayeed. Distributed target classification and tracking in sensor networks. *Proceedings of the IEEE*, 91(8):1163–1171, 2003.

[13] Adam Caromicoli and Thomas Kurien. Multitarget identification in airborne surveillance. In *1989 Orlando Symposium*, pages 161–176. International Society for Optics and Photonics, 1989.

[14] Gordon Crawford and Cindy Williams. A note on the analysis of subjective judgment matrices. *Journal of mathematical psychology*, 29(4):387–405, 1985.

[15] Milan Daniel. Associativity in combination of belief functions; a derivation of minc combination. *Soft Computing*, 7(5):288–296, 2003.

[16] Arthur P Dempster. Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pages 325–339, 1967.

[17] Florentin Smarandache Jean Dezert. Proportional conflict redistribution rules for information fusion. *Advances and Applications of DSmT for Information Fusion (Collected works), second volume: Collected Works*, 2:3, 2006.

[18] Didier Dubois and Henri Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4(3):244–264, 1988.

[19] James S Dyer. Remarks on the analytic hierarchy process. *Management science*, 36(3):249–258, 1990.

[20] Mihai C Florea, Anne-Laure Jousselme, Dominic Grenier, and Eloi Bosse. Combining belief functions and fuzzy membership functions. In *AeroSense 2003*, pages 113–122. International Society for Optics and Photonics, 2003.

[21] Mihai Cristian Florea, Anne-Laure Jousselme, Éloi Bossé, and Dominic Grenier. Robust combination rules for evidence theory. *Information Fusion*, 10(2):183–197, 2009.

[22] Ernest Forman and Kirti Peniwati. Aggregating individual judgments and priorities with the analytic hierarchy process. *European journal of operational research*, 108(1):165–169, 1998.

[23] Le Hegarat-Mascle, Isabelle Bloch, Daniel Vidal-Madjar, et al. Application of dempster-shafer evidence theory to unsupervised classification in multisource remote sensing. *Geoscience and Remote Sensing, IEEE Transactions on*, 35(4):1018–1031, 1997.

[24] Frederick Thomas Jane et al. *Jane's all the world's aircraft*. McGraw-Hill, 1909.

[25] Wen Jiang, Dejie Duanmu, Xin Fan, and Yong Deng. A new method to determine basic probability assignment under fuzzy environment. In *Systems and Informatics (ICSAI), 2012 International Conference on*, pages 758–762. IEEE, 2012.

[26] Audun Jøsang, Milan Daniel, and Patrick Vannoorenberghe. Strategies for combining conflicting dogmatic beliefs. In *Proceedings of the 6th International Conference on Information Fusion*, pages 1133–1140, 2003.

[27] Marie-Hélène Masson and Thierry Denoeux. Ecm: An evidential version of the fuzzy< i> c</i>-means algorithm. *Pattern Recognition*, 41(4):1384–1397, 2008.

[28] Wei Mei, Gan-lin Shan, and Yue-feng Wang. A second-order uncertainty model for target classification using kinematic data. *Information Fusion*, 12(2):105–110, 2011.

[29] R Ramanathan and LS Ganesh. Group preference aggregation methods employed in ahp: An evaluation and an intrinsic process for deriving members' weightages. *European Journal of Operational Research*, 79(2):249–265, 1994.

[30] Branko Ristic, Neil Gordon, and Amanda Bessell. On target classification using kinematic data. *Information Fusion*, 5(1):15–21, 2004.

[31] Branko Ristic and Ph Smets. Target identification using belief functions and implication rules. *Aerospace and Electronic Systems, IEEE Transactions on*, 41(3):1097–1103, 2005.

[32] Christoph Romer and Abraham Kandel. Constraints on belief functions imposed by fuzzy random variables. *Systems, Man and Cybernetics, IEEE Transactions on*, 25(1):86–99, 1995.

[33] Thomas L Saaty. Group decision making and the ahp. In *The Analytic Hierarchy Process*, pages 59–67. Springer, 1989.

[34] Thomas L Saaty. An exposition of the ahp in reply to the paper "remarks on the analytic hierarchy process". *Management science*, 36(3):259–268, 1990.

[35] TL Saaty, LG Vargas, and RE Wendell. Assessing attribute weights by ratios, 1983.

[36] Fabien Salzenstein and Abdel-Ouahab Boudraa. Unsupervised multi sensor data fusion approach. *Sixth International Symposium on Signal Processing and its Applications*, 1:152–155, 2001.

[37] Thomas L Satty. The analytic hierarchy process, 1980.

[38] Glenn Shafer. *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976.

[39] Florentin Smarandache and Jean Dezert. A simple proportional conflict redistribution rule. *JAIF Journal*, pages 1–36, 2005.

[40] Philippe Smets and Robert Kennes. The transferable belief model. *Artificial intelligence*, 66(2):191–234, 1994.

[41] Philippe Smets and Branko Ristic. Kalman filter and joint tracking and classification based on belief functions in the tbm framework. *Information Fusion*, 8(1):16–27, 2007.

[42] Peida Xu, Yong Deng, Xiaoyan Su, and Sankaran Mahadevan. A new method to determine basic probability assignment from training data. *Knowledge-Based Systems*, 46:69–80, 2013.

[43] Ronald R Yager. On the dempster-shafer framework and new combination rules. *Information sciences*, 41(2):93–137, 1987.

[44] Ronald R Yager. A class of fuzzy measures generated from a dempster–shafer belief structure. *International Journal of Intelligent Systems*, 14(12):1239–1247, 1999.

[45] Lotfi A Zadeh. *On the validity of Dempster's rule of combination of evidence*. Electronics Research Laboratory, University of California, 1979.

[46] Sajjad Zahir. Clusters in a group: Decision making in the vector space formulation of the analytic hierarchy process. *European Journal of Operational Research*, 112(3):620–634, 1999.

[47] Yue Min Zhu, Olivier Dupuis, Michele Rombaut, Layachi Bentabet, Daniel Babot, et al. Automatic determination of mass functions in dempster-shafer theory using fuzzy c-means and spatial neighborhood information for image segmentation. *Optical Engineering*, 41(4):760–770, 2002.

APPENDIX A

THE QUESTIONNAIRE FOR THE SECOND LEVEL OF THE HIERARCHY OF AHP MODEL FOR TARGET IDENTIFICATION

This questionnaire is built for the second level of the hierarchy. According to the target's response to any of the questions, target is wanted to be identified as friend or hostile or neutral.

Tables from A.1 to A.6 are for IFF interrogations. There are 3 tables for IFF Mod-4 and 3 tables for IFF Mod-3. These 3 tables correspond to 3 different responses as valid response, invalid response and no response respectively.

Table A.1: Questions for IFF Mod-4 valid response situation

	Which one?	How much?
If a target gives valid response to IFF Mod-4, which identity is more probable among being friend or hostile? How much?		
If a target gives valid response to IFF Mod-4, which identity is more probable among being friend or neutral? How much?		
If a target gives valid response to IFF Mod-4, which identity is more probable among being hostile or neutral? How much?		

Table A.2: Questions for IFF Mod-4 invalid response situation

	Which one?	How much?
If a target gives invalid response to IFF Mod-4, which identity is more probable among being friend or hostile? How much?		
If a target gives invalid response to IFF Mod-4, which identity is more probable among being friend or neutral? How much?		
If a target gives invalid response to IFF Mod-4, which identity is more probable among being hostile or neutral? How much?		

Table A.3: Questions for IFF Mod-4 no response situation

	Which one?	How much?
If a target gives no response to IFF Mod-4, which identity is more probable among being friend or hostile? How much?		
If a target gives no response to IFF Mod-4, which identity is more probable among being friend or neutral? How much?		
If a target gives no response to IFF Mod-4, which identity is more probable among being hostile or neutral? How much?		

Table A.4: Questions for IFF Mod-3 valid response situation

	Which one?	How much?
If a target gives valid response to IFF Mod-3, which identity is more probable among being friend or hostile? How much?		
If a target gives valid response to IFF Mod-3, which identity is more probable among being friend or neutral? How much?		
If a target gives valid response to IFF Mod-3, which identity is more probable among being hostile or neutral? How much?		

Table A.5: Questions for IFF Mod-3 invalid response situation

	Which one?	How much?
If a target gives invalid response to IFF Mod-3, which identity is more probable among being friend or hostile? How much?		
If a target gives invalid response to IFF Mod-3, which identity is more probable among being friend or neutral? How much?		
If a target gives invalid response to IFF Mod-3, which identity is more probable among being hostile or neutral? How much?		

Table A.6: Questions for IFF Mod-3 no response situation

	Which one?	How much?
If a target gives no response to IFF Mod-3, which identity is more probable among being friend or hostile? How much?		
If a target gives no response to IFF Mod-3, which identity is more probable among being friend or neutral? How much?		
If a target gives no response to IFF Mod-3, which identity is more probable among being hostile or neutral? How much?		

Tables A.7 and A.8 are for the restricted area breach information. First table is for breach case and second table is for no breach case.

Table A.7: Questions for RAB breach situation

	Which one?	How much?
If a target breaches the restricted area, which identity is more probable among being friend or hostile? How much?		
If a target breaches the restricted area, which identity is more probable among being friend or neutral? How much?		
If a target breaches the restricted area, which identity is more probable among being hostile or neutral? How much?		

Table A.8: Questions for RAB no breach situation

	Which one?	How much?
If a target does not breaches the restricted area, which identity is more probable among being friend or hostile? How much?		
If a target does not breaches the restricted area, which identity is more probable among being friend or neutral? How much?		
If a target does not breaches the restricted area, which identity is more probable among being hostile or neutral? How much?		

Tables A.9 and A.10 are for the air corridor usage information. First table is for correct usage case and second table for incorrect usage case.

Table A.9: Questions for ACU correct usage situation

	Which one?	How much?
If a target uses air corridors correctly, which identity is more probable among being friend or hostile? How much?		
If a target uses air corridors correctly, which identity is more probable among being friend or neutral? How much?		
If a target uses air corridors correctly, which identity is more probable among being hostile or neutral? How much?		

Table A.10: Questions for ACU incorrect usage situation

	Which one?	How much?
If a target uses air corridors incorrectly, which identity is more probable among being friend or hostile? How much?		
If a target uses air corridors incorrectly, which identity is more probable among being friend or neutral? How much?		
If a target uses air corridors incorrectly, which identity is more probable among being hostile or neutral? How much?		