Distributed Random Set Theoretic Soft/Hard Data Fusion

by

Bahador Khaleghi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Research on multisensor data fusion aims at providing the enabling technology to combine information from several sources in order to form a unified picture. The literature work on fusion of conventional data provided by non-human (hard) sensors is vast and well-established. In comparison to conventional fusion systems where input data are generated by calibrated electronic sensor systems with well-defined characteristics, research on soft data fusion considers combining human-based data expressed preferably in unconstrained natural language form. Fusion of soft and hard data is even more challenging, yet necessary in some applications, and has received little attention in the past. Due to being a rather new area of research, soft/hard data fusion is still in a fledging stage with even its challenging problems yet to be adequately defined and explored.

This dissertation develops a framework to enable fusion of both soft and hard data with the *Random Set* (RS) theory as the underlying mathematical foundation. Random set theory is an emerging theory within the data fusion community that, due to its powerful representational and computational capabilities, is gaining more and more attention among the data fusion researchers. Motivated by the unique characteristics of the random set theory and the main challenge of soft/hard data fusion systems, i.e. the need for a unifying framework capable of processing both unconventional soft data and conventional hard data, this dissertation argues in favor of a random set theoretic approach as the first step towards realizing a soft/hard data fusion framework.

Several challenging problems related to soft/hard fusion systems are addressed in the proposed framework. First, an extension of the well-known Kalman filter within random set theory, called Kalman evidential filter (KEF), is adopted as a common data processing framework for both soft and hard data. Second, a novel ontology (syntax+semantics) is developed to allow for modeling soft (human-generated) data assuming target tracking as the application. Third, as soft/hard data fusion is mostly aimed at large networks of information processing, a new approach is proposed to enable distributed estimation of soft, as well as hard data, addressing the scalability requirement of such fusion systems. Fourth, a method for modeling trust in the human agents is developed, which enables the fusion system to protect itself from erroneous/misleading soft data through discounting such data on-the-fly. Fifth, leveraging the recent developments in the RS theoretic data fusion literature a novel soft data association algorithm is developed and deployed to extend the proposed target tracking framework into multi-target tracking case. Finally, the multi-target tracking framework is complemented by introducing a distributed classification approach applicable to target classes described with soft human-generated data.

In addition, this dissertation presents a novel data-centric taxonomy of data fusion methodologies. In particular, several categories of fusion algorithms have been identified and discussed based on the data-related challenging aspect(s) addressed. It is intended to provide the reader with a generic and comprehensive view of the contemporary data fusion literature, which could also serve as a reference for data fusion practitioners by providing them with conducive design guidelines, in terms of algorithm choice, regarding the specific data-related challenges expected in a given application.

Acknowledgments

First and foremost, I would like to sincerely express my gratitude to my supervisor Professor Fakhri Karray for his continual support during the course of my PhD studies and his patience, guidance, and encouragement throughout the entire research work conducted for this dissertation. I am also very grateful for the financial support provided to me through funding from the ORF Muses Secret project.

Moreover, I want to thank my committee members Dr Baris Fidan, Dr. Pin-Han Ho, and Dr. Zhou Wang. It has been a genuine pleasure to have had the opportunity to work with such passionate, stimulating, and friendly people. Your collaborations, discussions, and feedback during the course of my studies are immensely appreciated. I should also thank Dr. Alaa Khamis for his truly helpful guidance, brainstorming discussions, and valuable feedback during the first two years of research work for this dissertation. The valuable assistance of my fellow graduate student Saiedeh Razavi and of other graduate and undergraduate students, who collaborated on publications, is also deeply appreciated.

Last but not the least, I would like to thank my parents, my sisters and my wonderful wife for their love, motivation and support throughout my graduate studies.

Dedication

This is dedicated to my father, Dr. Ezzatollah Khaleghi, whose encouragement has always been the main force driving me towards higher scientific achievements. It is also dedicated to my beautiful wife, Laleh, as her continual support played a crucial role paving my way through the challenging route of research work for this dissertation.

Table of Contents

List of Abbreviations

AA:	Actuation Agent
ACL:	Agent Communication Language
ATE:	Average Tracking Error
CFA:	Consensus Filter Agent
CPA:	Consensus Propagation Agent
CI:	Covariance Intersection
CPHD:	Cardinalized Probability Hypothesis Density
CU:	Covariance Union
DFA:	Data Fusion Agent
DPA:	Data Preprocessing Agent
DPF:	Distributed Particle Filtering
DSET:	Dempster-Shafer Evidence Theory
FRST:	Fuzzy Rough Set Theory
GEOINT:	Geospatial Intelligence
GM:	Gaussian Mixture
HCI:	Human Computer Interaction
HDA:	Hard Data Agent
HIA:	Human Interaction Agent
HUMINT:	Human Intelligence
IEA:	Internal Ellipsoid Approximation
IF:	Information Filter
JDL:	Joint Directors of Laboratories
KEF:	Kalman Evidential Filter
KF:	Kalman Filter
LE:	Largest Ellipsoid

List of Abbreviations

LRF:	Laser Range Finder
MAP:	Maximum A-Posterior
MASINT:	Measurement And Signature Intelligence
ML:	Maximum Likelihood
NLP:	Natural Language Processing
OOSM:	Out Of Sequence Measurements
OOST:	Out Of Sequence Tracks
OSINT:	Open Source Intelligence
PDF:	Probability Distribution Function
PDPF:	Parallel Distributed Particle Filter
PF:	Particle Filter
PHD:	Probability Hypothesis Density
P/S:	Player/Stage
RFS:	Random Finite Set
SDA:	Soft Data Agent
SDAA:	Soft Data Association Algorithm
SIGINT:	Signal Intelligence
SMC:	Sequential Monte-Carlo
TBM:	Transferable Belief Model

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Scope

Traditionally data fusion systems have been mostly developed to tackle problems in military domain such as situational awareness, threat assessment, condition monitoring of machinery, and target identification and tracking. As a result, the bulk of research in data fusion literature is based upon a related process model called JDL [9] where input data is primarily provided by physical (electronic) sensors (S-space) such as radar, LIDAR [1], and acoustic sensors with limited input from human observers [185]. However, two recent trends in the data fusion community are changing the conventional role of humans from a passive element to an active analyst and provider of data. The first major trend is data fusion applications that are also interested at targets, which are not primarily physical, such as locations, identity, and interaction of individuals and groups. For instance tackling the military threat of IED [2] involves a hierarchy of physical to nonphysical targets sought, ranging from physical devices and vehicles to human networks, belief systems, cyberconnectivity, and policies. This requires a transition from monitoring and characterizing the physical landscape to the one referred to as human landscape, involving data-rich yet model-poor problems. In other words, the main challenge is to properly model such abundant human-generated data in order to be able to process it in a mathematically sound and coherent manner. The second trend driving the change of human role is the emergence of two new sources of data, namely, human observations (H-space) and the Web-based data (I-space). For instance, about 4 billion cell phones, mostly equipped with

[1]Light Detection and Ranging
[2]Improvised Explosive Devices

GPS sensors, image sensors, etc., are currently used worldwide. One can consider these cell phone users as a formal/informal community of observers supplying information about an evolving situation. Similarly, new websites that enable information sharing by humans such as YouTube, Facebook, Twitter, and Blogs provide a gigantic amount of data.

In contrast to the conventional data provided in the S-space, which could typically be related to the system state in a well-defined and crisp way, hence the name hard data, the data available in the H-space and the I-space are usually more difficult to model and relate to the system state, hence the name soft data. Indeed, processing and utilizing such data involves numerous challenges that are yet to be appropriately identified and understood. The pioneering work by Hall et al. [149] outlines some of these new challenging problems including the soft sensor tasking [3], data and knowledge elicitation, and most importantly representation of data imperfection and second-order uncertainty. The main focus of this dissertation is on the last issue, i.e. we propose the *Random Set* (RS) theory as an appropriate mathematical framework to enable fusion of soft, as well as hard data. The RS theory is an attractive solution to this issue, which has gained researchers' attention recently. Indeed, imperfect data represented in probabilistic, evidential, fuzzy, and possibilistic frameworks are shown to have a corresponding formulation within the random set framework [29, 97]. Due to such powerful representational and computational capabilities, we advocate a RS theoretic approach as an appealing solution to enable fusion of disparate forms of data. Moreover, the main application context assumed in this work is the distributed target tracking in sensor networks. As a result, the novel framework proposed in this dissertation can be considered to lie at the intersection of three active areas of research in the data fusion community, namely, RS theoretic data fusion, soft/hard data fusion, and distributed data fusion (See Figure 1.1).

The essential thesis of this dissertation is a random set theoretic approach to enable fusion of soft/hard data. To substantiate this thesis:

- Section 1.2 of this chapter briefly describes the main contributions of this dissertation.
- Chapter 2 presents an elaborate survey of the data fusion state-of-the-art based on a novel data-centric taxonomy of fusion methodologies. In particular, the unifying power of the random set theory to enable dealing with various data imperfection aspects is discussed in this chapter.
- Chapter 3 proposes the core of our RS theoretic approach to the soft/hard data fusion problem relying on an extension of the Kalman filter within the RS theory called the

[3]Please note throughout this thesis the terms hard sensor and non-human sensor are used interchangeably.

Kalman evidential filter (KEF). A review of the related literature work is presented. This chapter further elaborates on the data modeling schemes developed for both soft and hard data, as well as the multi-agent architecture adopted to implement the proposed framework.

- Chapter 4 considers the problem of single-target tracking using soft/hard data and applies the proposed RS theoretic framework as a solution. The goal is to estimate the state, e.g. position and/or velocity, for the single target of interest over time. Furthermore, two novel approaches to enable distributed data aggregation and modeling of human trustworthiness, thus enhancing the scalability and robustness of the proposed framework, are presented. Lastly, a series of single-target tracking experiments evaluating the proposed framework are detailed.

- Chapter 5 further extends the target tracking problem into multi-target case by proposing a novel soft data association algorithm. As with the single-target tracking problem, the objective of a multi-target tracking system is to estimate the state of multiple targets of interest over time. A review of the background literature work is presented. Similar to chapter 4, a series of multi-target tracking experiments assessing the proposed approach are described. Furthermore, a new distributed target classification method applicable to target classes described using soft human-generated data is discussed.

- Chapter 6 provides a synopsis and critical appraisal of the key results, and finally sketches future research directions.

1.2 Key Contributions

This dissertation adopts a random set theoretic approach to address the problem of soft/hard data fusion for distributed target tracking task.

The following is a summary of the main contributions of this dissertation:

- A comprehensive review of the data fusion state-of-the-art based on a novel data-centric taxonomy is presented. Our review discusses various imperfection aspects of fusion data and the mathematical techniques commonly deployed to deal with them. In particular, the power of RS theory as a unifying approach to unable dealing with the majority of the data imperfection aspects is explored.

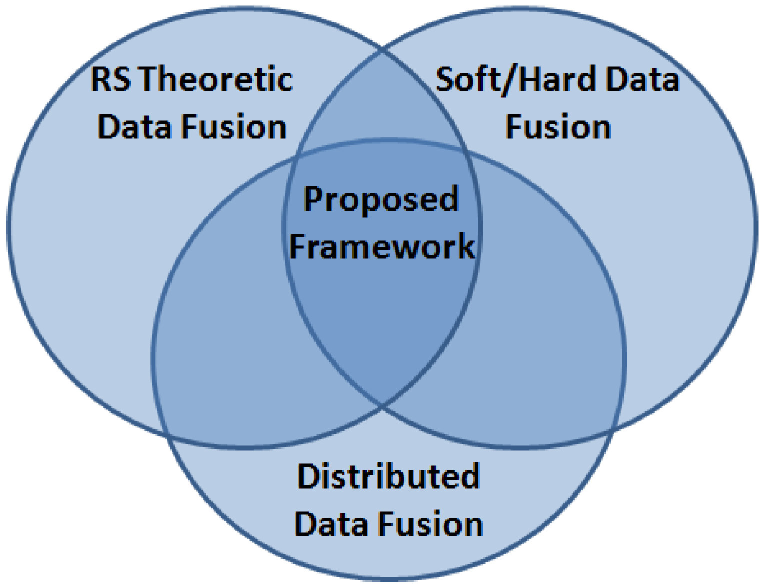

Figure 1.1: The scope of the proposed data fusion framework

- Using the RS theory, a powerful, scalable, and robust single-target tracking framework is developed that is capable of processing both soft human-generated data, as well as hard conventional data. Extensive experimental results demonstrate the tracking performance enhancement achieved using the proposed distributed data aggregation algorithm and the human trust modeling scheme in challenging target tracking scenarios. As soft/hard data fusion applications typically involve a large number of sensor nodes (human and non-human), which could also potentially be unreliable (especially for human type sensor nodes), the proposed distributed data aggregation and human trust modeling approaches tackle two of the most important issues regarding the practicality of the soft/hard fusion systems in the future.

- Leveraging recent developments in the RS theoretic data fusion literature, a novel soft data association algorithm (SDAA) is developed to further extend the proposed target tracking framework into multi-target case. The proposed SDAA can be deemed as an augmented nearest neighbor association algorithm that is applicable to soft data modeled using the RS theory and takes into account both the human observer opinions, as well as potential inter target conflict to come up with the final measurement to track associations. A series of multi-target tracking experiments using soft data are conducted to demonstrate the efficiency of the proposed approach. Moreover, the proposed RS theoretic multi-target framework is complemented by introducing a novel distributed classifier applicable to target classes described with vague human-generated data. The preliminary experimental results demonstrate the robustness of the proposed approach to the noisy target data and highly vague target description data.

Chapter 2

A Data-centric Taxonomy of Multisensor Data Fusion Methodologies

2.1 Introduction

Multisensor data fusion is a technology to enable combining information from several sources in order to form a unified picture. Data fusion systems are now widely used in various areas such as sensor networks, robotics, video and image processing, and intelligent system design, to name a few. Data fusion is a wide ranging subject and many terminologies have been used interchangeably. These terminologies and ad hoc methods in a variety of scientific, engineering, management, and many other publications, shows the fact that the same concept has been studied repeatedly. The focus of this chapter is on multisensor data fusion. Thus, throughout this chapter the terms data fusion and multisensor data fusion are used interchangeably.

The data fusion research community have achieved substantial advances, especially in recent years. Nevertheless, realizing a perfect emulation of the data fusion capacity of the human brain is still far from accomplished. This chapter is an endeavor to investigate the data fusion task, including its potential advantages, challenging aspects, existing methodologies, and recent advances. In particular, discussion of the existing data fusion methods relies on a data-centric taxonomy, and explores each method based on the specific data-related challenging aspect(s) addressed. While several general [1, 2, 3] and specific [4, 5, 6, 7, 8] reviews of the data fusion literature exist; this chapter is intended to provide the reader with a generic and comprehensive view of contemporary data fusion methodologies, as well as the most recent developments and emerging trends in the field. The bulk of data fusion research has been dedicated to problems associated with the first level of the Joint Directors of Laboratories (JDL) model [3]. As work on low-level fusion becomes well established and approaches maturity, research on high level fusion tasks is gaining more attention. A discussion of new developments on high level fusion methodologies may be insightful; nonetheless, as the focus of this chapter is on low level fusion, such presentation is left to a future work.

The rest of this chapter is organized as follows: in section 2.2 popular definitions, conceptualizations, and purposes, as well as the major benefits of data fusion, are discussed. The challenging problems pertaining to performing data fusion are described in section 2.3. Section 2.4 provides a discussion of data fusion methodologies based on their data treatment approach. Finally, section 2.5 presents the concluding remarks for this chapter.

2.2 Multisensor Data Fusion

Many definitions for data fusion exist in the literature. Joint Directors of Laboratories (JDL) [9] defines data fusion as a "multi-level, multifaceted process handling the automatic detection, association, correlation, estimation, and combination of data and information from several sources." Klein [10] generalizes this definition, stating that data can be provided either by a single source or by multiple sources. Both definitions are general and can be applied in different fields including remote sensing. In [11], the authors present a review and discussion of many data fusion definitions. Based on the identified strengths and weaknesses of previous work, a principled definition of information fusion is proposed as: "Information fusion is the study of efficient methods for automatically or semi-automatically transforming information from different sources and different points in time into a representation that provides effective support for human or automated decision making". Data fusion is a multi-disciplinary research area borrowing ideas from many diverse fields such as signal processing, information theory, statistical estimation and inference, and artificial intelligence. This is indeed reflected in the variety of the techniques presented in section 2.4.

Generally, performing data fusion has several advantages [12, 2]. These advantages mainly involve enhancements in data authenticity or availability. Examples of the former are improved detection, confidence, and reliability, as well as reduction in data ambiguity, while extending spatial and temporal coverage belong to the latter category of benefits. Data fusion can also provide specific benefits for some application contexts. For example, wireless sensor networks are often composed of a large number of sensor nodes, hence posing a new scalability challenge caused by potential collisions and transmissions of redundant data. Regarding energy restrictions, communication should be reduced to increase the lifetime of the sensor nodes. When data fusion is performed during the routing process, that is, sensor data is fused and only the result is forwarded, the number of messages is reduced, collisions are avoided, and energy is saved.

Various conceptualizations of the fusion process exist in the literature. The most common and popular conceptualization of fusion systems is the JDL model [9]. The JDL classification is based on the input data and produced outputs, and originated from the military domain. The original JDL model considers the fusion process in four increasing levels of abstraction, namely, object, situation, impact, and process refinement. Despite its popularity, the JDL model has many shortcomings, such as being too restrictive and especially tuned to military applications, which have been the subject of several extension proposals [13, 14] attempting to alleviate them. The JDL formalization is focused on data (input/output) rather than processing. An alternative is Dasarathy's framework [15] that views the fusion system, from a software engineering perspective, as a data flow characterized by input/output as well as functionalities (processes). Another general conceptualization of fusion is the work of Goodman et al. [16], which is based on the notion of *random sets*. The distinctive aspects of this framework are its ability to combine decision uncertainties with decisions themselves, as well as presenting a fully generic scheme of uncertainty representation. One of the most recent and abstract fusion frameworks is proposed by Kokar et al. [17]. This formalization is based on category theory and is claimed to be sufficiently general to capture all kinds of fusion, including data fusion, feature fusion, decision fusion, and fusion of relational information. It can be considered as the first step towards development of a formal theory of fusion. The major novelty of this work is the ability to express all aspects of multi-source information processing, i.e., both data and processing. Furthermore, it allows for consistent combination of the processing elements (algorithms) with measurable and provable performance. Such formalization of fusion paves the way for the application of formal methods to standardized and automatic development of fusion systems.

2.3 Challenging Problems of Multisensor Data Fusion

There are a number of issues that make data fusion a challenging task. The majority of these issues arise from the data to be fused, imperfection and diversity of the sensor technologies, and the nature of the application environment as following:

- Data imperfection: data provided by sensors is always affected by some level of impreciseness as well as uncertainty in the measurements. Data fusion algorithms should be able to express such imperfections effectively, and to exploit the data redundancy to reduce their effects.

- Outliers and spurious data: The uncertainties in sensors arise not only from the impreciseness and noise in the measurements, but are also caused by the ambiguities and inconsistencies present in the environment, and from the inability to distinguish between them [18]. Data fusion algorithms should be able to exploit the redundant data to alleviate such effects.

- Conflicting data: fusion of such data can be problematic especially when the fusion system is based on evidential belief reasoning and Dempster's rule of combination [19]. To avoid producing counter-intuitive results, any data fusion algorithm must treat highly conflicting data with special care.

- Data modality: sensor networks may collect the qualitatively similar (homogeneous) or different (heterogeneous) data such as auditory, visual, and tactile measurements of a phenomenon. Both cases must be handled by a data fusion scheme.

- Data correlation: this issue is particularly important and common in distributed fusion settings, e.g. wireless sensor networks, as for example some sensor nodes are likely to be exposed to the same external noise biasing their measurements. If such data dependencies are not accounted for, the fusion algorithm, may suffer from over/under confidence in results.

- Data alignment/registration: sensor data must be transformed from each sensor's local frame into a common frame before fusion occurs. Such an alignment problem is often referred to as sensor registration and deals with the calibration error induced by individual sensor nodes. Data registration is of critical importance to the successful deployment of fusion systems in practice.

- Data association: multi-target tracking problems introduce a major complexity to the fusion system compared to the single-target tracking case [20]. One of these new difficulties is the data association problem, which may come in two forms: measurement-to-track and track-to-track association. The former refers to the problem of identifying from which target, if any, each measurement is originated, while the latter deals with distinguishing and combining tracks, which are estimating the state of the same real-world target [3].

- Processing framework: data fusion processing can be performed in a centralized or decentralized manner. The latter is usually preferable in wireless sensor networks, as it allows each sensor node to process locally collected data. This is much more efficient compared to the communicational burden required by a centralized approach, when all measurements have to be sent to a central processing node for fusion.

- Operational timing: the area covered by sensors may span a vast environment composed of different aspects varying in different rates. Also, in the case of homogeneous sensors, the operation frequency of the sensors may be different. A well-designed data fusion method should incorporate multiple time scales in order to deal with such timing variations in data. In distributed fusion settings, different parts of the data may traverse different routes before reaching the fusion center, which may cause out-of-sequence arrival of data. This issue needs to be handled properly, especially in real-time applications, to avoid potential performance degradation.

- Static vs. dynamic phenomena: the phenomenon under observation may be time-invariant or varying with time. In the latter case, it may be necessary for the data fusion algorithm to incorporate a recent history of measurements into the fusion process [21]. In particular, data freshness, i.e., how quickly data sources capture changes and update accordingly, plays a vital role in the validity of fusion results. For instance in some recent work [22], the authors performed a probabilistic analysis of the recent history of measurement updates to ensure the freshness of input data, and to improve the efficiency of the data fusion process.

- Data dimensionality: the measurement data could be preprocessed, either locally at each of the sensor nodes or globally at the fusion center to be compressed into lower dimensional data, assuming a certain level of compression loss is allowed. This preprocessing stage is beneficial as it enables saving on the communication bandwidth and power required for transmitting data, in the case of local preprocessing [23], or limiting the computational load of the central fusion node, in the case of global preprocessing [24].

While many of these problems have been identified and heavily investigated, no single data fusion algorithm is capable of addressing all the aforementioned challenges. The variety of methods in the literature focus on a subset of these issues to solve, which would be determined based on the application in hand. Our presentation of data fusion literature is organized according to the taxonomy shown in Figure 2.1. The existing fusion algorithms are explored based on how various data-related challenges are treated.

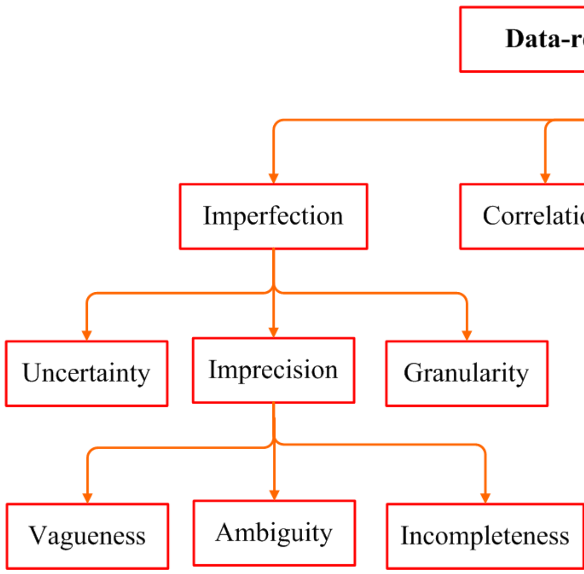

Figure 2.1: Taxonomy of data fusion methodologies: different data fusion algorithms can be roughly categorized based on one of the four challenging problems of input data that are mainly tackled: namely, data imperfection, data correlation, data inconsistency, and disparateness of data form.

2.4 Multisensor Data Fusion Algorithms

Regardless of how different components (modules) of the data fusion system are organized, which is specified by the given fusion architecture, the underlying fusion algorithms must ultimately process (fuse) the input data. As discussed in section 2.3, real-world fusion applications have to deal with several data related challenges. As a result, we decided to explore data fusion algorithms according to our novel taxonomy based on data-related aspects of fusion. Figure 2.1 illustrates an overview of data-related challenges that are typically tackled by data fusion algorithms. The input data to the fusion system may be imperfect, correlated, inconsistent, and/or in disparate forms/modalities. Each of these four main categories of challenging problems can be further subcategorized into more specific problems, as shown in Figure 2.1 and discussed in the following.

Various classifications of imperfect data have been proposed in the literature [25, 26, 27]. Our classification of imperfect data is inspired by the pioneering work of Smets' [26] as well as recent elaborations by Dubois and Prade [28]. Three aspects of data imperfection are considered in our classification: uncertainty, imprecision, and granularity.

Data is uncertain when the associated confidence degree, about what is stated by the data, is less than 1. On the other hand, the imprecise data is that data which refers to several, rather than only one, object(s). Finally, data granularity refers to the ability to distinguish among objects, which are described by data, being dependent on the provided set of attributes. Mathematically speaking, assume the given data d (for each described object of interest) to be structured as the following:

$$\textit{object } O \quad \textit{attribute } A \quad \textit{statement } S$$

representing that the data d is stating S regarding the relationship of some attribute(s) A to some object O in the world. Further assume $C(S)$ to represent the degree of confidence we assign to the given statement S. Then, data is regarded to be uncertain if $C(S) < 1$ while being precise, i.e., a singleton. Similarly, data is deemed as imprecise if the implied attribute A or degree of confidence C are more than one, e.g. an interval or set. Please note, the statement part of the data are almost always precise.

The imprecise A or C may be well-defined or ill-defined, and/or, miss some information. Thus, imprecision can manifest itself as ambiguity, vagueness, or incompleteness of data. The ambiguous data refers to those data where the A or C are exact and well-defined yet imprecise. For instance, in the sentence "Target position is between 2 and 5" the assigned attribute is the well-defined imprecise interval [2 5]. The vague data is characterized by having ill-defined attributes, i.e., attribute is more than one and not a well-defined set or interval. For instance, in the sentence "The tower is large" the assigned attribute "large" is not well-defined as it can be interpreted subjectively, i.e., have different meaning from one observer to the other. The imprecise data that has some information missing is called incomplete data. For instance, in the sentence "It is possible to see the chair", only the upper limit on the degree of confidence C is given, i.e., $C < \tau$ for some τ [29].

Consider an information system [30] where a number of (rather than one) objects $O = \{o_1, \ldots, o_k\}$ are described using a set of attributes $A = \{V_1, V_2, \ldots, V_n\}$ with respective domains $D_1, D_2, \ldots, D_n$. Let $F = D_1 \times D_2 \times \ldots \times D_n$ to represent the set of all possible descriptions given the attributes in A, also called the frame. It is possible for several objects to share the same description in terms of these attributes. Let $[o]_F$ to be the set of objects that are equivalently described (thus indistinguishable) within the frame F, also called the equivalence class. Now, let $T \subseteq O$ to represent the target set of objects. In general, it is not possible to exactly describe T using F, because T may include and exclude objects which are indistinguishable within the frame F. However, one can approximate T by the lower and upper limit sets that can be described exactly within F in terms of the induced equivalence classes. Indeed, the Rough set theory, discussed later on in this section, provides a systematic approach to this end. In summary, data granularity refers to the fact that the choice of data frame F (granule) has a significant impact on the resultant data imprecision. In other words, different attribute subset selections $B \subseteq A$ will lead to different frames, and thus different sets of indiscernible (imprecise) objects.

Correlated (dependent) data is also a challenge for data fusion systems and must be treated properly. We consider inconsistency in input data to stem from (highly) conflicting, spurious, or out of sequence data. Finally, fusion data may be provided in different forms, i.e. in one or several modalities, as well as generated by physical sensors (hard data) or human operators (soft data).

We believe such categorization of fusion algorithms is beneficial as it enables explicit exploration of popular fusion techniques according to the specific data-related fusion challenge(s) they target. Furthermore, our taxonomy is intended to facilitate ease of development by supplying fusion algorithm designers with an outlook of the appropriate and established techniques to tackle the data-related challenges their given application may involve. Finally, such exposition would be more intuitive and therefore helpful to non-experts in data fusion by providing them with an easy-to-grasp view of the field.

2.4.1 Fusion of Imperfect Data

The inherent imperfection of data is the most fundamental challenging problem of data fusion systems, and thus the bulk of research work has been focused on tackling this issue. There are a number of mathematical theories available to represent data imperfection [31], such as probability theory [32], fuzzy set theory [33], possibility theory [34], rough set theory [35], and Dempster-Shafer evidence theory (DSET) [36]. Most of these approaches are capable of representing specific aspect(s) of imperfect data. For example, a probabilistic distribution expresses data uncertainty, fuzzy set theory can represent vagueness of data, and evidential belief theory can represent uncertain as well as ambiguous data. Historically, the probability theory was used for a long time to deal with almost all kinds of imperfect information, because it was the only existing theory. Alternative techniques such as fuzzy set theory and evidential reasoning have been proposed to deal with perceived limitations in probabilistic methods, such as complexity, inconsistency, precision of models, and uncertainty about uncertainty [32]. We discuss each of these families of data fusion algorithms, along with their hybridizations that aim for a more comprehensive treatment of data imperfection. Examples of such hybrid frameworks are fuzzy rough set theory (FRST) [37] and fuzzy Dempster-Shafer theory (Fuzzy DSET) [38]. We also describe the new emerging field of fusion using *random sets*, which could be used to develop a unified framework for

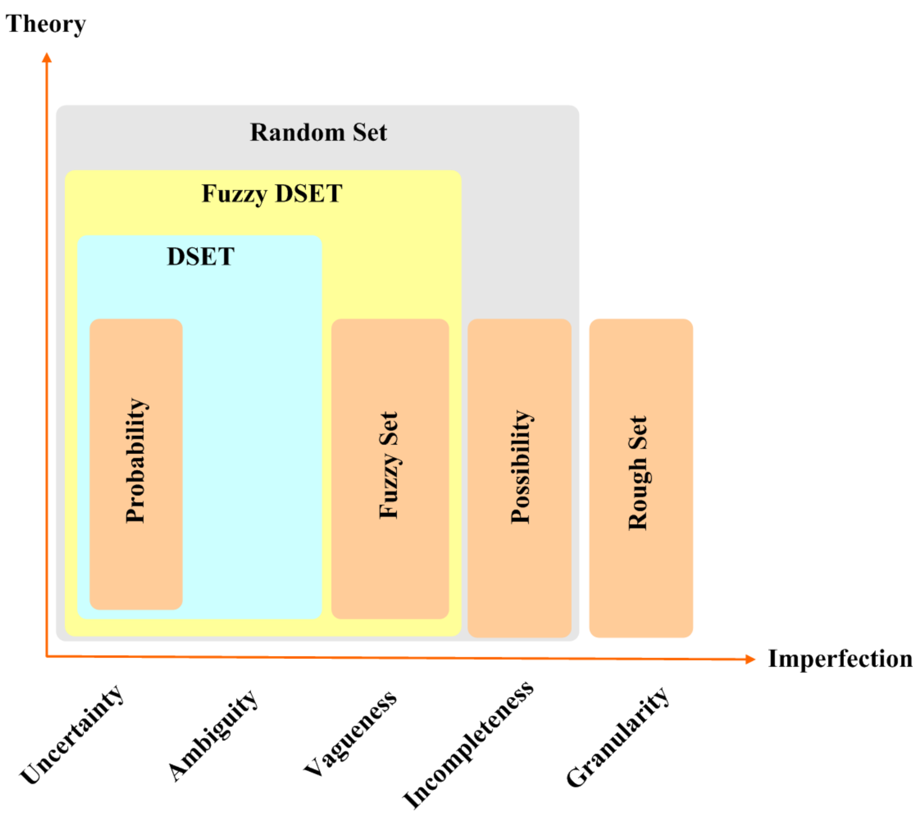

Figure 2.2: Overview of theoretical frameworks of imperfect data treatment (note: the fuzzy rough set theory is omitted from the diagram to avoid confusion)

treatment of data imperfections [39]. Figure 2.2 provides an overview of the aforementioned mathematical theories of dealing with data imperfections. On the x-axis, various aspects of data imperfection, introduced in Figure 2.1, are depicted. The box around each of the mathematical theories designates the range of imperfection aspects targeted mainly by that theory. The interested reader is referred to [29] for a comprehensive review of the classical theories of representing data imperfections, describing each of them along with their inter-relations.

Probabilistic Fusion

Probabilistic methods rely on the probability distribution/density functions to express data uncertainty. At the core of these methods lies the Bayes estimator, which enables fusion of pieces of data, hence the name “Bayesian fusion”. Assuming a state-space representation, the Bayes estimator provides a method for computing the posterior (conditional) probability distribution/density of the hypothetical state x_k at time k given the set of measurements $Z^k = \{z_1, \ldots, z_k\}$ (up to time k) and the prior distribution, as following

$$p(x_k|Z^k) = \frac{p(z_k|x_k)p(x_k|Z^{k-1})}{p(Z^k|Z^{k-1})} \tag{2.1}$$

Where

- $p(z_k|x_k)$ is called the likelihood function and is based on the given sensor measurement model
- $p(x_k|Z^{k-1})$ is called the prior distribution and incorporates the given transition model of the system
- The denominator is a merely a normalizing term to ensure that the probability density function integrates to one

One can apply the Bayes estimator each time and update the probability distribution/density of the system state by fusing the new piece of data, i.e. z_k, recursively. However, both the prior distribution and the normalizing term contain integrals that cannot be evaluated analytically in general. Thus, an analytic solution of the Bayes estimator is occasionally available. Indeed, the well-known Kalman filter (KF) is an exceptional case of the Bayes filter with an exact analytical solution due to enforcing simplifying (and somewhat unrealistic) constraints on the system dynamics to be linear-Gaussian, i.e. the measurement and motion model are assumed to have a linear form and be contaminated with zero-mean Gaussian noise [39]. Nonetheless, the Kalman filter is one of the most popular fusion methods mainly due to its simplicity, ease of implementation, and optimality in a mean-squared error sense. It is a very well established data fusion method whose properties are deeply studied and examined both theoretically and in practical applications. On the other hand, similar to other least-square estimators, the Kalman filter is very sensitive to data corrupted with outliers. Furthermore, the Kalman filter is inappropriate for applications whose error characteristics are not readily parameterized.

When dealing with non-linear system dynamics, one usually has to resort to approximation techniques. For instance, the Extended KF [40] and Unscented KF [41], which are extensions of the Kalman filter applicable to non-linear systems, are based on the first-order and second-order approximations as a Taylor series expansion about the current estimate, respectively. However, both of these methods can only handle non-linearities to a limited extent. Grid-based methods [42] provide an alternative approach for approximating non-linear probability density functions, although they rapidly become computationally intractable in high dimensions.

The Monte Carlo simulation-based techniques such as Sequential Monte Carlo (SMC) [43] and Markov Chain Monte Carlo (MCMC) [44] are among the most powerful and popular methods of approximating probabilities. They are also very flexible as they do not make any assumptions regarding the probability densities to be approximated. Particle filters are a recursive implementation of the SMC algorithm [45]. They provide an alternative for Kalman filtering when dealing with non-Gaussian noise and non-linearity in the system. The idea is to deploy a (weighted) ensemble of randomly drawn samples (particles) as an approximation of the probability density of interest. When applied within the Bayesian framework, particle filters are used to approximate the posterior (conditional) probability of the system state as a weighted sum of random samples. The random samples are usually drawn (predicted) from the prior density (transition model) with their weights updated according to the likelihood of the given measurement (sensing model). This approach to the implementation of particle filters is referred to as sequential importance sampling (SIS). One usually performs a resampling step where the current set of particles is replaced by a new set drawn from it with probabilities proportional to their weights. This step is included in the original proposal of the particle filters [46], which is called sequential importance resampling(SIR).

Similar to the Kalman filter, the particle filters have been shown to be sensitive to outliers in data, and require a set of auxiliary variables to improve their robustness [47]. In addition, when compared to the Kalman filter, particle filters are computationally expensive as they may require a large number of random samples (particles) to estimate the desired posterior probability density. Indeed, they are not suitable for fusion problems involving a high-dimensional state space as the number of particles required to estimate a given density function increases exponentially with dimensionality.

An attractive alternative for particle filters when dealing with high dimensions, are the MCMC algorithms. The underlying idea is to ease the burden of high-dimensional density approximation by using a Markov chain to evolve the samples, instead of simply drawing them randomly (and independently) at each step. Here, a Markov chain is a sequence of random samples generated according to a transition probability (kernel) function with Markovian property, i.e. the transition probabilities between different sample values in the state space depend only on the random samples' current state. It has been shown that one can always use a well-designed Markov chain that converges to a unique stationary density of interest (in terms of drawn samples) [44]. The convergence occurs after a sufficiently large number of iterations, called the burn-in period. Metropolis et al. [48] were the first to deploy this technique for solving problems involving high-dimensional density approximation. Their method was later extended by Hastings [49] and is referred to as the Metropolis-Hastings algorithm. The algorithm works by successively sampling a candidate point from some jumping (proposal) distribution, which is the conditional probability of a potential sample given a current sample. The obtained candidate point is accepted with a probability that is determined based on the ratio of the density at the candidate and current points. The Metropolis-Hastings algorithm is sensitive to the sample initialization and the choice of jumping distribution. Indeed, the burn-in period may be significantly longer for an inappropriate choice of initial samples and/or jumping distribution. Research on the so-called optimal starting point and jumping distribution is the subject of active work. The starting point is typically set as close as possible to the center of distribution, e.g. the distribution's mode. Also, random walks and independent chain sampling are two of the commonly adopted approaches for jumping distribution.

The popular Gibbs sampler is a special case of the Metropolis-Hastings algorithm where the candidate point is always accepted. The keys advantage of this method is that it considers only univariate conditional distributions, which usually have simpler form and are thus much easier to simulate than complex full joint distributions [50]. Accordingly, the Gibbs sampler simulates n random variables sequentially from the n univariate conditionals rather than generating a single n-dimensional vector in a single pass using the full joint distribution. One of the difficulties of applying MCMC methods in practice is to estimate the burn-in time, although it is often suggested that provided a large enough sample size, the burn-in time is not that important. Nonetheless, the effect of burn-in time may not be neglected when parallel processing schemes are deployed to implement MCMC methods [51]. With parallel MCMC the computational load is divided into several pieces, and thus the individual sample sizes may not be as large. To alleviate this problem, the convergence diagnostics methods [52] are commonly used to determine the burn-in time. This has to be done with caution as these methods can potentially introduce some biases of their own into the computations.

Evidential Belief Reasoning

The theory of belief functions was founded by Dempster's work [53], and was later mathematically formalized by Shafer [36] toward a general theory of reasoning based on evidence. It is a popular method to deal with uncertainty and imprecision and is based on a theoretically attractive evidential reasoning framework. Dempster-Shafer theory introduces the notion of assigning beliefs and plausibilities to possible measurement hypotheses along with a collection of combination rules to fuse them.

Mathematically speaking, consider X to represent all possible states of a system (also called the frame of discernment) and the power set 2^X to represent the set of all possible subsets of X. In contrast to probability theory that assigns a probability mass to each element of X, Dempster-Shafer theory assigns belief mass m to each element E of 2^X, which represent possible propositions regarding the system state x. Function m has two properties as follows

1. $m(\phi) = 0$
2. $\sum_{E \in 2^X} m(E) = 1$

Intuitively for any proposition E, $m(E)$ represents the proportion of available evidence that supports the claim that the actual system state x belongs to E. Usually, m is non-zero for only a limited number of sets called the focal elements. Using m, a probability interval can be obtained for E as below:

$$bel(E) \leq P(E) \leq pl(E) \tag{2.2}$$

Where:

- $bel(E)$ is called belief of E and is defined as $bel(E) = \sum_{B \subseteq E} m(B)$
- $pl(E)$ is called plausibility of E and is defined as $pl(E) = \sum_{B \cap E \neq \phi} m(B)$

Evidence from sensors is usually fused using the Dempster's rule of combination. Consider two sources of information with belief mass functions m_1 and m_2, respectively. The joint belief mass function $m_{1,2}$ is computed as follows:

$$m_{1,2}(E) = (m_1 \oplus m_2)(E) = \frac{1}{1-K} \sum_{B \cap C = E \neq \phi} m_1(B) m_2(C) \tag{2.3}$$

$$m_{1,2}(\phi) = 0 \tag{2.4}$$

Where K represents the amount of conflict between the sources and is given by:

$$K = \sum_{B \cap C = \phi} m_1(B) m_2(C) \tag{2.5}$$

D-S Theory has established itself as a promising and popular approach to data fusion especially in the last few years. Nonetheless there are issues such as the exponential complexity of computations (in general worst case scenario) as well as the possibility of producing counterintuitive results when fusing conflicting data using Dempster's rule of combination. Both of these issues have been heavily studied in the literature and numerous strategies have been proposed to resolve or alleviate them. Several family of complexity reduction approaches based on graphical techniques [58], parallel processing schemes [59], reducing the number of focal elements [60], and coarsening the frame of discernment to approximate the original belief potentials [61] have been studied. Some works have also deployed the finite set representation of focal elements to facilitate fusion computations [62].

As mentioned, the latter issue of fusing conflicting data using Dempster's rule of combination has been an active area of fusion research and has been studied extensively, especially in recent years. Many solutions to this issue have been proposed, which are discussed in detail in section 2.4.3.

Fusion and fuzzy reasoning

Fuzzy set theory is another theoretical reasoning scheme for dealing with imperfect data. It introduces the novel notion of partial set membership, which enables imprecise (rather than crisp) reasoning [33]. A fuzzy set $F \subseteq X$ is defined by the gradual membership function $\mu_F(x)$ in the interval $[0, 1]$ as below:

$$\mu_F(x) \in [0, 1] \quad \forall x \in X \tag{2.6}$$

Where the higher the membership degree, the more x belongs to F. This makes fuzzy data fusion an efficient solution where vague or partial sensory data is fuzzified using a gradual membership function. Fuzzy data can then be combined using fuzzy rules to produce fuzzy fusion output(s). Fuzzy fusion rules can be divided into conjunctive and disjunctive categories. Examples of the former are the following:

$$\mu_1^{\cap}(x) = min[\mu_{F_1}(x), \mu_{F_2}(x)] \quad \forall x \in X \tag{2.7}$$

$$\mu_2^{\cap}(x) = \mu_{F_1}(x).\mu_{F_2}(x) \quad \forall x \in X \tag{2.8}$$

which represent the standard intersection and product of two fuzzy sets, respectively. Some examples of the latter fuzzy fusion category are

$$\mu_1^{\cup}(x) = max[\mu_{F_1}(x), \mu_{F_2}(x)] \quad \forall x \in X \tag{2.9}$$

$$\mu_2^{\cup}(x) = \mu_{F_1}(x) + \mu_{F_2}(x) - \mu_{F_1}(x).\mu_{F_2}(x) \quad \forall x \in X \tag{2.10}$$

which represent the standard union and algebraic sum of two fuzzy sets, respectively. Conjunctive fuzzy fusion rules are considered appropriate when fusing data provided by equally reliable and homogeneous sources. On the other hand, disjunctive rules are deployed when (at least) one of the sources is deemed reliable, though which one is not known, or when fusing highly conflictual data. Accordingly, some adaptive fuzzy fusion rules have been developed, as a compromise between the two categories, that can be applied in both cases. The following fusion rule proposed in [64] is an example for adaptive fuzzy fusion:

$$\mu_{Adaptive}(x) = max\Big\{\frac{mu_i^{\cap}(x)}{h(\mu_{F_1}(x), \mu_{F_2}(x))}, min\{1 - h(\mu_{F_1}(x), \mu_{F_2}(x)), \mu_j^{\cup}(x)\}\Big\} \forall x \in X \tag{2.11}$$

where $h(\mu_{F_1}, \mu_{F_2})$ measures the degree of conflict between the gradual membership functions μ_{F_1} and μ_{F_2} defined as

$$h(\mu_{F_1}(x), \mu_{F_2}(x)) = max(min\{\mu_{F_1}(x), \mu_{F_2}(x)\}) \quad \forall x \in X \tag{2.12}$$

and $\mu_i^{\cap}$ and $\mu_j^{\cup}$ are the desired conjunctive and disjunctive fuzzy fusion rules, respectively.

In contrast to the probability and evidence theories, which are well suited to modeling the uncertainty of membership of a target in a well-defined class of objects, fuzzy sets theory is well suited to modeling the fuzzy membership of a target in an ill-defined class. Yet, similar to probability theory that requires prior knowledge of probability distributions, fuzzy sets theory requires prior membership functions for different fuzzy sets. Due to being a powerful theory to represent vague data, fuzzy set theory is particularly useful to represent and fuse vague data produced by human experts in a linguistic fashion. Furthermore, it has been often integrated with probabilistic [155, 66] and D-S evidential [38, 67] fusion algorithms in a complementary manner.

Possibilistic fusion

Possibility theory was founded by Zadeh [34] and later extended by Dubois and Prade [68, 69]. It is based on fuzzy set theory, but was mainly designed to represent incomplete rather than vague data. Indeed possibility theory's treatment of imperfect data is similar in spirit to probability and D-S evidence theory with a different quantification approach [29]. The model of imperfect data in possibility theory is the possibility distribution $\pi_B(x) \in [0,1]\ \forall x \in X$, which characterizes the uncertain membership of an element x in a (well-defined) known class B. This is distinguished from the gradual membership function $\mu_F(x)$ of fuzzy set theory, which characterizes the membership of x in an ill-defined fuzzy set F. Another important distinction is the normalization constraint that requires that at least one value is totally possible, i.e. $\exists x^* \in X\ s.t.\ \pi_B(x^*) = 1$. Given the possibility distribution $\pi_B(x)$, the possibility measure $\Pi(U)$ and necessity measure $N(U)$ of an event U are defined as below

$$\Pi(U) = max_{x \in U}\{\pi_B(x)\} \quad \forall U \subseteq X \tag{2.13}$$

$$N(U) = min_{x \notin U}\{1 - \pi_B(x)\} \quad \forall U \subseteq X \tag{2.14}$$

A possibility degree $\Pi(U)$ quantifies to what extent the event U is plausible, while the necessity degree $N(U)$ quantifies the certainty of U, in the face of incomplete information expressed by $\pi(x)$ [70]. The possibility and necessity measures can also be interpreted as a special case of upper and lower probabilities, in connection with the probability theory [71].

The data combination rules used for possibilistic fusion are similar to those deployed for fuzzy fusion. The main difference is that possibilistic rules are always normalized. The choice of appropriate fusion rules is dependent on the how agreeable the data sources are, and also what is known about their reliability [69]. However, the basic symmetric conjunctive and disjunctive fusion rules of fuzzy set theory are sufficient only for restricted cases. There are a number of enhancements of possibilistic fusion methods that allow for handling more difficult fusion scenarios. For instance, assuming $0 \leq \lambda_i \leq 1$ to represent the perceived reliability of the i^{th} source for a set of unequally reliable sources, one can modify the associated possibility distribution π_i of the source using the discounting approach as $\pi_i^{'} = max(\pi_i, 1 - \lambda)$ to incorporate its reliability into the fusion process [68].

Although possibility theory has not been commonly used in the data fusion community, some researchers have studied its performance in comparison to probabilistic and evidential fusion approaches [72], where it was shown to be capable of producing competitive results. Also, possibilistic fusion is argued to be most appropriate in poorly informed environments (no statistical data available) as well as in fusion of heterogeneous data sources [64]. For example, a recent work by Benferhat and Sossai [73] has demonstrated the effectiveness of possibilistic fusion for robot localization in partially known indoor environments.

Rough set based fusion

Rough set is a theory of imperfect data developed by Pawlak [35] to represent imprecise data, ignoring uncertainty at different granularity levels. Indeed, the Rough set theory enables dealing with data granularity. It provides means of approximating a crisp target set T within a given frame F_B designated by the set $B \subseteq A$, which is the specific set of attributes chosen to describe objects. The approximation is represented as a tuple $\langle B_*(T), B^*(T)\rangle$, where $B_*(T)$ and $B^*(T)$ represent the lower and upper approximations of set T within frame F_B, respectively, and are defined as below [74]

$$B_*(T) = \{o|[o]_{F_B} \subseteq T\} \tag{2.15}$$

$$B^*(T) = \{o|[o]_{F_B} \cap T \neq \phi\} \tag{2.16}$$

and $B_*(T) \subseteq \ T \ \subseteq B^*(T)$. The lower approximation $B_*(T)$ can be interpreted as a conservative approximation that includes only objects that are definitely a member of T, whereas the upper approximation $B^*(T)$ is more liberal in including all objects that can possibly belong to T. Based on this approximation, the boundary region of T is defined as $BN_B(T) = B^*(T) - B_*(T)$ which is the set of objects that can neither be classified as belonging nor not-belonging to T. Accordingly, a set T is considered rough if $BN_B(T) \neq \phi$.

Within the data fusion framework, T can be considered as representing the imprecise set of (target) states of a system (instead of abstract objects). Then, Rough set theory would allow the approximation of possible states of the system based on the granularity of input data, i.e. F_B. Once approximated as rough sets, data pieces can be fused using classic set theory conjunctive or disjunctive fusion operators, i.e. intersection or union, respectively.

In order to perform fusion successfully, data granules must be neither too fine nor too rough. In the case of data granules being too fine, i.e. $[o]_{F_B}$ being singletons, the Rough set theory reduces to classical set theory. On the other hand, for very rough data granules, i.e. $[o]_{F_B}$ being very large subsets, the lower approximation of data is likely to be empty, resulting in total ignorance. The major advantage of Rough set compared to other alternatives is that it does not require any preliminary or additional information such as data distribution or membership function [75]. Rough set theory allows for fusion of imprecise data approximated based merely on its internal structure (granularity).

Due to being a relatively new theory and not well understood within fusion community, Rough set theory has been rarely applied to data fusion problems. Some work has been reported on data fusion systems using Rough set theory [76, 77], where it provides a means to select the most informative set of attributes (sensors) regarding the goal of the fusion system, e.g. classification of objects. The idea is to use a rough integral as the measure of relevance for each sensor, and filter out sensors below the given threshold.

Hybrid fusion approaches

The main idea behind development of hybrid fusion algorithms is that different fusion methods such as fuzzy reasoning, D-S evidence theory, and probabilistic fusion should be not be competing, as they approach data fusion from different (possibly complementary) perspectives. At the theoretical level, hybridization of fuzzy set theory with D-S evidence theory has been studied frequently [78, 38] aiming at providing a framework for more comprehensive treatment of data imperfection. Among many such proposals, the work by Yen [38] is perhaps the most popular approach that extends D-S evidence theory into the fuzzy realm while maintaining its major theoretical principles. Yen's theory of fuzzy D-S evidence theory has been frequently used in the literature. For instance, Zue and Basir [67, 79] developed a hybrid fusion system applied to an image segmentation problem, which is based on a fuzzy Dempster-Shafer evidential reasoning scheme.

Combination of fuzzy set theory with Rough set theory (FRST), proposed by Dubois and Prade, is another important theoretical hybridization existing in the literature [37]. In spite of being a powerful representation tool for vague as well as ambiguous data, the original FRST has some limitations such as relying on special fuzzy relations. This issue has been recently addressed by Yeung et al. [80] in an attempt to generalize FRST to arbitrary fuzzy relations. Application of FRST to data fusion has not often been investigated in the fusion literature as Rough set theory itself is still not an established data fusion approach. Nonetheless, some preliminary work has been reported [81].

Random set theoretic fusion

The principles of random sets theory were first proposed to study integral geometry in 1970s [82]. The unifying capability of random set theory has been shown by several researchers [83, 84, 16], among them, the work of Goodman et al. [16] has been most successful in gaining attention. The most notable work on promoting random set theory as a unified fusion framework has been done by Mahler in his papers [20, 16, 85] and recent book [39]. In particular, in his book he attempts to present a detailed exposition of random set theory and its application to general single-target as well as multi-target data fusion problems.

Random set theory is usually deemed as an ideal framework for extending the popular Bayes filter from single-target (modeled by a random variable) into multi-target (modeled by a random set). Accordingly, the majority of research work has been focused on applying random set theory to tracking of multiple targets. This generalization is not a straightforward procedure and is only possible provided that an appropriate calculus of random finite sets is formulated [20]. Indeed, within random set theory data, i.e. target states and measurements, are modeled as random sets of finite size instead of conventional vectors. Having done this, priors and likelihood functions are constructed that are capable of modeling a wide range of different phenomena. For instance, phenomena related to the system dynamics such as target disappearance/appearance, extended/unresolved targets, and target spawning, as well as measurement-related phenomena such as missed detection and false alarms can be explicitly represented.

Obviously, one can not expect to solve for this multi-target tracking analytically (as was not the case for single-target Bayes filter). Therefore, different approximation techniques are devised to compute the Bayes update equation. The moment matching techniques have been very successful in approximating the single-target Bayes filter. For instance, Kalman filter relies on propagating the first two moments (i.e. mean and covariance) while alpha-beta filters match only the first moment. In case of multi-target tracking, the first moment is the Probability Hypothesis Density (PHD), which is used to develop a filter with the same title, i.e. PHD filter [86]. There is also a higher order extension of this filter called Cardinalized Probability Hypothesis Density (CPHD) filter [87, 88], which propagates the PHD as well as the full probability distribution of the random variable representing the number of targets. Both PHD and CPHD filters involve integrals that prevent direct implementation of a closed form solution. As a result two approximation methods, namely, Gaussian Mixture (GM) and Sequential Monte Carlo (SMC), have been used in the literature to further ease the implementation stage for these filters [89, 90]. Both of these methods have been evaluated and shown to compare favorably with alternative approaches such as JPDA [88] and MHT [91], while being less computationally demanding than either. One important advantage of the (C)PHD family of filters is to avoid the data association problem, but this also means that maintaining track continuity can become a challenging task. For a review of recent work on the (C)PHD filter, the interested reader is referred to [92].

Random set theory has also been recently shown to be able to efficiently solve fusion related tasks such as target detection [93], tracking [94], identification [29], sensor management [95], and soft/hard data fusion [96]. Nonetheless, further research through more complex test scenarios in diverse applications should be performed to prove its performance as a unifying framework for fusion of imperfect data.

2.4.2 Fusion of Correlated Data

Many data fusion algorithms, including the popular KF approach, require either independence or prior knowledge of the cross covariance of data to produce consistent results. Unfortunately, in many applications fusion data is correlated with potentially unknown cross covariance. This can occur due to common noise acting on the observed phenomena [100] in centralized fusion settings, or the rumor propagation issue, also known as data incest or double counting problem [101], where measurements are inadvertently used several times in distributed fusion settings [3]. If not addressed properly, data correlation can lead to biased estimation, e.g. artificially high confidence value, or even divergence of fusion algorithm [102]. For KF-based systems, the optimal KF approach exists that allows for maintaining cross covariance information between updates [3]. However, it is not typically desirable, as it is shown to scale quadratically with the number of updates [103]. Also, in case of data incest, an exact solution is to keep track of pedigree information which includes all sensor measurements that have contributed to a certain estimate [104]. This solution is not appealing as it does not scale well with the number of fusion nodes [105]. Most of the proposed solutions to correlated data fusion attempt to solve it by either eliminating the cause of correlation or tackling the impact of correlation in fusion process.

Table 2.1: Comparison of Imperfect Data Fusion Frameworks

Framework	**Characteristics**	**Capabilities**	**Limitations**
Probabilistic [32, 40, 45]	Represents sensory data using probability distributions fused together within Bayesian framework	Well-established and understood approach to treat data uncertainty	Considered incapable of addressing other data imperfection aspects

Table 2.1: Comparison of Imperfect Data Fusion Frameworks

Framework	**Characteristics**	**Capabilities**	**Limitations**
Evidential [36, 54, 55, 56, 58]	Relies on probability mass to further characterize data using belief and plausibilities and fuses using Dempsters' combination rule	Enables fusion of uncertain and ambiguous data	Does not deal with other aspects of data imprecision, inefficient for fusion of highly conflicting data
Fuzzy reasoning [155, 66, 67]	Allows vague data representation, using fuzzy memberships, and fusion based on fuzzy rules	Intuitive approach to deal with vague data esp. human generated	Limited merely to fusion of vague data
Possibilistic [29, 72, 64]	Similar in data representation to probabilistic and evidential frameworks and fusion to fuzzy framework	Allows for handling incomplete data common in poorly informed environment	Not commonly used and well understood in fusion community
Rough set theoretic [35, 99, 75, 77]	Deals with ambiguous data using precise approximate lower and upper bounds manipulated using classical set theory operators	Does not require any preliminary or additional information	Requires appropriate level of data granularity

Table 2.1: Comparison of Imperfect Data Fusion Frameworks

Framework	**Characteristics**	**Capabilities**	**Limitations**
Hybridization [78, 38, 67, 79]	Aims at providing a more comprehensive treatment of imperfect data	Deploys fusion framework in a complementary rather than competitive fashion	Rather ad-hoc generalization of one fusion framework to subsume other(s), extra computational burden
Random set theoretic [20, 16, 85, 39]	Relies on random subsets of measurement/state space to represent many aspects of imperfect data	Can potentially provide a unifying framework for fusion of imperfect data	Relatively new and not very well appreciated in fusion community

Eliminating data correlation

Data correlation is especially problematic in distributed fusion systems and is commonly caused by data incest. The data incest situation itself happens when the same information takes several different paths from the source sensor to the fusion node or due to cyclic paths through which the information recirculates from output of a fusion node back to the input [106, 3]. This issue can be eliminated (before fusion) either explicitly by removal of data incest [107] or implicitly through reconstruction of measurements [108]. The former family of approaches usually assume a specific network topology as well as fixed communication delays, although recent extensions consider the more general problem of arbitrary topologies with variable delays using graph theoretic algorithms [109, 110]. The latter approaches attempt to form a decorrelated sequence of measurements by reconstructing them such that the correlation with previous intermediate updates from current intermediate state updates is removed. The decorrelated sequence is then fed to the global fusion processor as input to a filtering algorithm. Extensions in this family consider more complex fusion scenarios with existence of clutter, data association, and interacting targets [111].

Data fusion in presence of unknown correlations

Instead of removing data correlation, one can design a fusion algorithm that accounts for correlated data. Covariance Intersection (CI) [100] is the most common fusion method to deal with correlated data. CI was originally developed to avoid the problem of covariance matrix underestimation due to data incest. It solves this problem in general form for two data sources (i.e. random variables) by formulating an estimate of the covariance matrix as a convex combination of the means and covariances of the input data. CI has been shown to be optimal, in terms of finding the upper bound for the combined covariances [112], as well as theoretically sound and applicable to any probability distribution function, from information theory perspective [113].

On the other hand, CI requires a non-linear optimization process and is therefore computationally demanding. Furthermore, it tends to overestimate the intersection region, which results in pessimistic results and consequent degradation of fusion performance. Some faster variants of CI have been proposed attempting to alleviate the former issue [114, 115]. The Largest Ellipsoid (LE) algorithm was developed, as an alternative to CI, to address the latter issue [116]. LE provides a tighter estimate of covariance matrix by finding the largest ellipse that fits within the intersection region of the input covariances. It has been recently argued that LE's formula derivation for the center of largest ellipsoid is not appropriate and a new algorithm, called Internal Ellipsoid Approximation (IEA), is proposed to accomplish this task [117]. One major limitation with all these methods is their inability to facilitate fusion of correlated data within a more powerful fusion framework than KF-based techniques, such as particle filters [3]. Very recently, a fusion framework based on an approximation to the generalized CI algorithm, called Chernoff fusion method, is proposed, which tackles the generic problem of fusing any number of correlated PDFs [118].

2.4.3 Fusion of Inconsistent Data

The notion of data inconsistency, as applied in this chapter, is in a generic sense and encompasses spurious, as well as disordered and conflicting data. We explore various techniques in the data fusion literature which are developed to tackle each of the three aspects of data inconsistency.

Spurious data

Data provided by sensors to the fusion system may be spurious due to unexpected situations such as permanent failures, short duration spike faults, or slowly developing failure [18]. If fused with correct data, such spurious data can lead to dangerously inaccurate estimates. For instance, KF would easily break down if exposed to outliers [119]. The majority of work on treating spurious data has been focused on identification/prediction and subsequent elimination of outliers from the fusion process. Indeed, the literature work on sensor validation is partially aiming at the same target [120, 121, 122]. The problem with most of these techniques is the requirement for prior information, often in the form of specific failure model(s) As a result, they would perform poorly in a general case where prior information is not available or unmodeled failures occur [123]. Recently, a general framework for detection of spurious data has been proposed that relies on stochastic adaptive modeling of sensors and is thus not specific to any prior sensor failure model [18, 124]. It is developed within the Bayesian fusion framework by adding a term to the common formulation that

Table 2.2: Summary of Correlated Data Fusion Methods

Framework	**Algorithms**	**Characteristics**
Correlation elimination	Explicit removal [107, 109, 110]	Usually assumes a specific network topology and fixed communication delays
	Measurement reconstruction [108, 111]	Applicable to more complex fusion scenarios
Correlation presence	Covariance Intersection [100, 112]	Avoids the covariance underestimation problem, yet computationally demanding and rather pessimistic
	Fast CI [114, 115]	Enhanced efficiency through alternative non-linear optimization processes
	Largest Ellipsoid [116]	Provides a tighter (less pessimistic) covariance estimate, yet limited to KF-based fusion like the others

represents the probabilistic estimate that the data is not spurious conditioned upon the data and the true state. The intended effect for this term is increasing the variance of the posterior distribution when data from one of the sensors is inconsistent with respect to the other. Extensive experimental simulations have shown the promising performance of this technique in dealing with spurious data [123].

Out of sequence data

The input data to the fusion system is usually organized as discrete pieces each labeled with a timestamp designating its time of origin. Several factors such as variable propagation times for different data sources as well as having heterogeneous sensors operating at multiple rates can lead to data arriving out of sequence at the fusion system. Such out of sequence measurements (OOSM) can appear as inconsistent data to the fusion algorithm. The main issue is how to use this, usually old, data to update the current estimate while taking care of the correlated process noise between the current time and the time of the delayed measurement [125]. A trivial solution to OOSM is to simply discard it. Such solution would cause information loss and severe fusion performance degradation if OOSM is prevalent in the input data. Another intuitive solution is to store all input data in order and reprocess it once OOSM is received. This approach yields optimal performance yet is impractical due to having intense computational and storage requirements. There has been considerable amount of research done in this area in the last decade due to the increasing popularity of distributed sensing and tracking systems [125]. We explore these methods according to their assumed number of step lags as well as number of tracking targets.

Most of the early work on OOSM assumed only single-lag data. For example, an approximate sub-optimal solution to OOSM called "Algorithm B" [126] as well as its famous optimal counterpart "Algorithm A" [127], both assume single-lag data. Some researchers have proposed algorithms to enable handling of OOSM with arbitrary lags [128, 129, 130]. Among these methods the work in [130] is particularly interesting as it provides a unifying framework for treating OOSM with "Algorithm A" as special case. Nonetheless, it was shown in [131] that this approach along with many other multi-lag OOSM methods are usually very expensive in terms of computational complexity and storage. The same authors proposed an extension to the "Algorithm A" and "Algorithm B" called "algorithm Al1"' and "Algorithm Bl1", respectively. They further showed that these new algorithms have requirements similar to their single-lag counterparts and are therefore recommended for practical applications, especially "Algorithm Bl1" is preferred due to being almost optimal and very efficient. Some recent work also investigates the OOSM problem in case of having both single-lag and multiple-lag data, termed the mixed-lag OOSM problem.

The proposed algorithm is claimed to handle all three types of OOSM data and is shown to be suboptimal in the linear MMSE sense under one approximation [132].

The bulk of research on the OOSM problem has been traditionally concentrated on an OOSM filtering algorithm that considers only a single-target, and does not address issues pertinent to data association and the presence of clutter that arise in multi-target fusion scenarios [133]. This problem has received attention in recent years and several methods tackling various aspects of OOSM in multi-target tracking have been proposed. In [133], a multitarget OOSM dwell-based tracking algorithms is proposed which includes gating, likelihood computation, and hypothesis management; and the single-lag and two-lag OOSM problems are discussed. In [134], the authors present a generic framework that enables straightforward extension of many single-target OOSM solutions to efficient algorithms in the multi-target data association case. The problem of out of sequence data for disordered tracks, instead of measurements, termed OOST, is explored in [135]. The OOST problem is solved using equivalent measurements obtained from individual sensor tracks, which are then used in an augmented state framework to compute the joint density of the current target state and the target state corresponding to the delayed data. Generally, in comparison with the OOSM problem, the OOST problem is much less studied in the literature. More recently, the three popular algorithms for the OOSM problem proposed by Bar-Shalom [136] are adapted to handle the OOST problem. This work is expected to improve the research community's understanding of the OOST problem.

Conflicting data

Fusion of conflicting data, when for instance several experts have very different ideas about the same phenomenon, has long been identified as a challenging task in the data fusion community. In particular, this issue has been heavily studied for fusion within the Dempster-Shafer evidence theory framework. As shown in a famous counterexample by Zadeh [137], naive application of Dempster's rule of combination to fusion of highly conflicting data results in unintuitive results. Since then Dempster's rule of combination has been subject to much criticism for rather counter-intuitive behavior [138]. Most of the solutions proposed alternatives to Dempster's rule of combinations [139, 140, 141, 142]. On the other hand, some authors have defended this rule, arguing that the counter-intuitive results are due to improper application of this rule [143, 144, 39]. For example, in [39] Mahler shows that the supposed unintuitive result of Dempster's combination rule can be resolved using a simple corrective strategy, i.e. to assign arbitrary small but non-zero belief masses to hypotheses deemed extremely unlikely. Indeed, proper application of Dempsters rule of combination requires satisfaction of the following three constraints: (1) independent sources providing independent evidences, (2) homogeneous sources defined on a unique frame of discernment and (3) a frame of discernment containing an *exclusive* and *exhaustive* list of hypotheses.

These constraints are too restrictive and difficult to satisfy in many practical applications. As a result DSET has been extended to more flexible theories such as Transferable Belief Model (TBM) [140] and DezertSmarandache theory (DSmT) [141]. The former theory extends DSET by refuting the *exhaustivity* constraint, i.e open-world assumption, and allowing elements outside the frame of discernment to be represented by the empty set. The latter refutes the *exclusivity* constraint allowing compound elements, i.e. elements of the hyper power set, to be represented. The theoretical justification for TBM was recently presented by Smets [19]. In this work, he provides an exhaustive review of the existing combination rules in an attempt to shed light on their applicability as well as theoretical soundness. He argues that the majority of proposed combination rules are ad-hoc in nature and lack appropriate theoretical justification. It is also demonstrated that most of the alternative combination rules are indeed conjunctive fusion operators that redistribute the global (or partial) conflicting belief mass among some elements of the power set. This relies on the notion that if experts agree on some evidence, they are considered reliable, and otherwise at least one of them is unreliable and the disjunctive fusion rules are deployed. But disjunctive rules usually result in degradation in data specificity. Therefore, the reliability of the expert sources must be either known *a priori* or estimated [145].

Fusion of conflicting data within the Bayesian probabilistic framework has also been explored by some authors. For example, Covariance Union (CU) algorithm is developed to complement the CI method, and enable data fusion where input data is not just correlated but may also be conflicting [146]. Furthermore, a new Bayesian framework for fusion of uncertain, imprecise, as well as conflicting data was proposed recently [147]. Authors exploit advances in the Bayesian research arena to develop Bayesian models with similar theoretical properties as TBM and DSmT theories allowing for consistent probabilistic fusion of conflicting data.

2.4.4 Fusion of Disparate Data

The input data to a fusion system may be generated by a wide variety of sensors, humans, or even archived sensory data. Fusion of such disparate data in order to build a coherent and accurate global view or the observed phenomena is a very difficult task. Nonetheless, in some fusion applications such as human computer interaction (HCI), such diversity of sensors is necessary to enable natural interaction with humans. Our focus of discussion is on fusion of human generated data (soft data) as well as fusion of soft and hard data, as research in this direction has attracted attention in recent years. This is motivated by the inherent limitations of electronic (hard) sensors and recent availability of communication infrastructure that allow humans to act as soft sensors [149]. Furthermore, while a tremendous amount of research has been done on data fusion using conventional sensors, very limited work has studied fusion of data produced by human and non-human sensors. An example of preliminary research in this area includes the work on generating a dataset for hard/soft data fusion intended to serve as a foundation and a verification/validation resource for future research [150, 151]. Also in [149], the authors provide a brief review on ongoing work on dynamic fusion of soft/hard data, identifying its motivation and advantages, challenges, and requirements. Very recently, a Dempster-Shafer theoretic framework for soft/hard data fusion is presented that relies on a novel conditional approach to updating as well as a new model to convert propositional logic statements from text into forms usable by Dempster-Shafer theory [152]. Furthermore, some new work investigates the problem of uncertainty representation for linguistic data [153]. The authors describe various types of uncertainty inherent in the nature of human languages as well as some tools to perform linguistic disambiguation such as lexicons, grammars, and dictionaries.

Table 2.3: Overview of Inconsistent Data Fusion Methodologies

Inconsistency Aspect	**Problem**	**Resolution Strategy**	**Characteristics**
Outlier	If fused with correct data, can lead to dangerously inaccurate estimates	Sensor validation techniques [120, 121, 122]	Identification/predication and subsequent removal of outliers, typically restricted to specific prior-known failure models
		Stochastic adaptive sensor modeling [123]	General framework for detection of spurious data without prior knowledge
Disorder	Update current estimate using old measurements (OOSM)	Ignore, reprocess, or use backward/forward prediction [126, 127, 125, 130, 148]	Mostly assume single-lag delays and linear target dynamics
	Update current estimate using old track estimates (OOST)	Use augmented state framework to incorporate delayed estimates [135, 136]	Much less understood and studied in the literature
Conflict	Non-intuitive results while fusing highly conflicting data using Dempsters' combination rule	Numerous alternative combination rules [139, 140, 141, 142]	Mostly ad-hoc in nature without proper theoretical justification
		Apply corrective strategies while using Dempsters' rule [143, 144, 39]	Defend validity of Dempsters' rule provided that certain constraints are satisfied

Another new direction of work is focused on the so called *human centered data fusion* paradigm that puts emphasis on the human role in the fusion process [154]. This new paradigm allows human to participate in the data fusion process not merely as soft sensors, but also as hybrid computers and ad-hoc teams (hive mind). It relies on emerging technologies such as virtual worlds and social network software to support humans in their new fusion roles. In spite of all these developments, research on hard/soft data fusion as well as human centered fusion is still in the fledging stage, and believed to provide rich opportunities for further theoretical advancement and practical demonstrations in the future.

2.5 Chapter Summary

This chapter presented a critical review of data fusion state-of-the-art methodologies based on a novel *data centric* taxonomy. Data fusion is a multi-disciplinary research field with a wide range of potential applications in areas such as defense, robotics, automation and intelligent system design, pattern recognition, etc. This has been and will continue to act as the driving force behind the ever-increasing interest in research community in developing more advanced data fusion methodologies and architectures.

Based on this exposition, it is clear that research on data fusion systems is becoming more and more common-place. There are a number of areas in the data fusion community that will most likely be highly active in the near future. For instance, the ever-increasing demand for data fusion on extremely large scales, such as sensor networks and the Web, will drive intense research on highly scalable data fusion algorithms based on distributed architectures. In addition, the availability and abundance of non-conventional data in the form of human-generated reports or Web documents will lead to the development of new and powerful fusion frameworks capable of processing a wide variety of data forms. Such fusion frameworks could potentially be realized by exploiting strong mathematical tools for modeling imperfect data, such as random set theory. This is the main motivation behind our RS theoretic soft/hard data fusion framework introduced in the next chapter.

Chapter 3

RS Theoretic Soft/Hard Data Fusion

3.1 Introduction

The literature work on fusion of conventional data provided by non-human (hard) sensors is vast and well-established as presented in reviews given in [2, 3]. In contrast to the conventional fusion systems where input data are assumed to be provided by calibrated electronic sensor systems (with well-defined characteristics), research on soft data fusion systems aims at developing approaches to enable combining human-generated data expressed preferably in unconstrained natural language. It is motivated mainly by asymmetric urban military operations, where human-generated data are shown to be of crucial importance [152]. Soft data fusion by itself is a challenging problem and has received little attention in the past [165]. Fusion of soft and hard data is even more challenging yet necessary in some applications. Recent developments in the literature such as human-centered data fusion paradigm [185] as well as several pioneering work on soft/hard fusion [188, 149, 152, 150, 151, 166] are an indicative of the new trend towards more general data fusion frameworks to achieve efficient processing of both soft and hard data. In this chapter we describe the main building blocks underpinning the novel RS theoretic soft/hard data fusion system developed in this work, namely, the KEF, the soft and hard data modeling schemes, and the multiagent architecture deployed for the system implementation.

In order to develop a soft/hard fusion system one has to deploy an appropriate mathematical framework to represent inherent data imperfections and allow for performing inference using available data. Probabilistic, Dempster-Shafer, fuzzy, possibilistic and rough set theoretic are among some of the most commonly used frameworks in the data fusion community. Each of these frameworks has its own advantages and drawbacks, nonetheless, none of them is able to fully address all aspects of imperfect data [97]. Random set (RS) theory is a candidate solution to this issue, which has gained researchers' attention recently. Indeed, imperfect data represented in probabilistic, evidential, fuzzy, and possibilistic frameworks are shown to have a corresponding formulation within the random set framework [29]. Due to such powerful representational and computational capabilities, in this chapter RS theory is proposed as an appealing approach to enable fusion of disparate forms of data. The recent research work on RS theoretic fusion has been mostly restricted to multitarget tracking problems [167, 168, 39] inspired by the pioneering works of Mahler [169, 170]. Very recently RS theory was deployed to develop a soft data fusion system capable of state estimation using natural language propositions [171]. Nonetheless, to the best of our knowledge, a RS theoretic approach to the soft/hard data fusion problem does not exist in the literature. The main contribution of this chapter is to develop a RS theoretic approach to represent both soft and hard data in a unified manner. In particular, assuming a target tracking application, a novel ontology is proposed to enable the formatting and interpretation of soft data provided as language statement(s). The natural (raw) soft data is assumed to be preprocessed by an appropriate natural language processing (NLP) algorithm.

The rest of this chapter is organized as follows. In section 3.2 a brief review of the related literature work is presented. The RS theoretic fusion approach for soft/hard data processing is discussed in section 3.3. Lastly, section 3.4 concludes this chapter.

3.2 Background and Related Work

In this section we briefly discuss the background literature pertinent to the main context of the proposed fusion system, i.e. soft/hard data fusion, as well as, the RS theoretic data fusion. It is worth to mention that the literature work on the RS theoretic multi-target tracking is rather established and is not detailed in this section. Interested reader is referred to [39] for a more elaborate discussion.

3.2.1 Soft/Hard Data Fusion

The main focus of research studies in this area is to achieve fusion of human-generated data (soft data), as well as fusion of soft and hard data. This is motivated by the inherent limitations of electronic (hard) sensors, and recent availability of communication infrastructure that allows humans to act as soft sensors [149]. Furthermore, while a tremendous amount of research has been done on data fusion using conventional sensors, very limited work has studied fusion of data produced by human and non-human sensors. An example of the preliminary research in this area is the work on generating a dataset for hard/soft data fusion intended to serve as a foundation and a verification/validation resource for future research [150, 151]. Also in [149], authors provide a brief review of ongoing work on dynamic fusion of soft/hard data, identifying its motivation and advantages, challenges, and requirements. Very recently, a Dempster-Shafer theoretic framework for soft/hard data fusion was presented that relies on a novel conditional approach to updating, as well as a new model to convert propositional logic statements from text into forms usable by Dempster-Shafer theory [152]. Furthermore, some new work examines the uncertainty representation problem for the case of linguistic data [153]. Authors examine various types of uncertainty inherent in the nature of human languages, as well as some of the existing tools to enable linguistic disambiguation such as lexicons, grammars, and dictionaries.

Another new direction of work in this area is focused on the so called *human centered data fusion* paradigm that puts emphasis on human role in the fusion process [185, 154]. This new paradigm considers humans as active participants in the data fusion process and not merely as soft sensors but also as hybrid computers and *ad hoc* teams (hive mind). It relies on emerging technologies such as virtual worlds and social network software to support humans in their new fusion roles. In spite of these accomplishments, research on hard/soft data fusion, as well as human-centered fusion is still in its fledging stage and is believed to provide rich opportunities for further theoretical advancement and practical demonstrations in the future [97].

3.2.2 Random Set Theoretic Fusion

The principles of random sets theory were first proposed to study integral geometry in 1970s [82]. The unifying capability of random set theory has been studied early on by several researchers [83, 84, 16], among them the work of Goodman et al. [16] has gained the most attention. The most notable recent work on promoting random set theory as a unified fusion framework has been done by Mahler in his papers [177, 85] and recent book [92]. In particular, in his book he makes an attempt to provide a detailed exposition of random set theory and its application to general single-target, as well as multi-target data fusion problems.

The RS theory is typically deployed within a Beyesian framework to provide a generalization of the (single-target) Bayes filtering to multi-target applications. As a result, the majority of research work has been focused on applying RS theory to tracking of multiple targets. This generalization is not a straightforward procedure and requires the development of an appropriate calculus of random finite sets where both target state and measurement(s) are modeled as random sets of finite size instead of the conventional vector representation, hence the name *random finite set* (RFS). Having done this, the prior and likelihood functions are also reconstructed to enable explicit modeling of a wide range of different phenomena observed in multi-target fusion systems such target disappearance/appearance, extended/unresolved targets, and target spawning, which are otherwise rather challenging to deal with using the more traditional multi-target tracking approaches such as JPDA [178] and MHT [179]. The RS theory can also be deployed to perform single-target Bayes filtering using the generalized (non-conventional) measurement(s) that cannot be readily expressed as vectors. For instance, human-generated data in the form of natural language statements or expert rules can be fused using RS theoretic representation of data within the Bayesian framework. This potential capability of RS theory has been rarely exploited in the data fusion literature. This observation has been the main motive for developing the soft/hard data fusion framework described in this chapter.

3.3 Random Set Theoretic Soft/Hard Data Fusion

Our pioneering work [96] was probably the first paper to argue in favor of a RS theoretic approach to deal with various imperfection aspects of soft data. This section presents the basic building blocks of our RS theoretic soft/hard fusion system, i.e. the RS theoretic representation of unconventional data imperfections, a generalization of the popular Kalman filter (KF), derived using the aforementioned RS theoretical formalism, called Kalman evidential filter (KEF), the soft and hard data modeling approaches, and the multi-agent organization of the proposed fusion system.

3.3.1 RS Theoretic Representation of Ambiguous and Vague Data

Although soft human-generated data can express several types of imperfections (See section 2.4), in this work we focus on two of the most common data imperfection aspects observed in soft data, namely, vagueness (fuzziness) and ambiguity. The exposition in this section is meant to outline the basic ideas underpinning the RS theoretic representation of such data. As discussed earlier in chapter 2, the vague data is a particular type of imprecise data characterized by having ill-defined attributes, i.e. attribute is not a singleton and not a well-defined set or interval mostly due to the being open to (subjective) interpretation.

Such data is usually modeled using a fuzzy membership function $f(u)$ with values in $[0,1]$ representing the membership degree of u in the fuzzy set specified by f.

Assume A to be a uniformly distributed random number on $[0,1]$, the random set representation of a fuzzy membership function $f(u)$ is [39]

$$\Sigma_A(f) = \{u | A \leq f(u)\} \tag{3.1}$$

For each random $a \in [0,1]$, the subset $\Sigma_a(f) \subseteq \Omega$ (Ω is the universe of discourse) is called the level set of the function f cut by the (hyper)plane given by a. Due to the uniformity of A,

$$P(u \in \Sigma_A(f)) = P(A \leq f(u)) = f(u) \tag{3.2}$$

Where $P()$ represents the probability distribution. In other words, the random set $\Sigma_A(f)$ faithfully represents the fuzzy information contained in $f(u)$.

The data ambiguity can also be defined as another form of imprecision where both the attribute and the statement confidence are well-defined yet imprecise. As discussed in section 2.4, the Dempster-Shafer theory is commonly deployed to represent data ambiguity using a basic mass assignment function (b.m.a.) m. For the case of ambiguous and vague data, the m can be extended into a fuzzy b.m.a. function $o(f)$ defined on fuzzy membership function(s) $f(u)$. The RS theoretic representation of o can be constructed by first dividing $A = [0,1]$ up into intervals $A_1, \ldots, A_e$ with respective lengths $o_1, \ldots, o_e$. Considering a generalized RS representation of a fuzzy membership function f_i as [39]

$$W_i = \{(u,a) | o^+_{i-1} < a < o^+_{i-1} + o_i f_i(u)\} \tag{3.3}$$

Where $o^+_0 = 0$ and $o^+_i = \sum_{k=1}^{i} o_k$ for $i = 1, \ldots, e$. Please note the W_i will be mutually disjoint.

Let

$$W_0 = W_1 \cup \ldots \cup W_e \tag{3.4}$$

The RS theoretic representation of $o(f)$ would be $\Sigma_A(W_0)$

3.3.2 Kalman Evidential Filter

The KEF is derived by applying RS theory within the Bayesian estimation framework and is capable of processing imprecise as well as vague data. The details of the KEF derivation are not presented in this section and can be found in [92]. Similar to the KF, the KEF is formulated within a state-space modeling scheme. However instead of working in the original space, KEF relies on the so-called fuzzy Dempster-Shafer states which are indeed a special case of fuzzy basic mass assignment function (fuzzy b.m.a.) on the original state space X. Mathematically speaking, a fuzzy D-S state is a non-negative function $\mu(f)$ defined on fuzzy membership functions $f(x)$ where $x \in X$. Furthermore, function $\mu(f)$ complies with the following three constraints:

1. $\mu(f) = 0$ for all but a finite number of fuzzy membership functions f that constitute its focal set,

2. The fuzzy membership functions of the focal set are normalized Gaussian multivariate in form, i.e.

$$f(x) = N_C(x - x_0) = exp(-\frac{1}{2}(x - x_0)^T C^{-1}(x - x_0)) \tag{3.5}$$

   where x_0 and C correspond to the mean and covariance of the multivariate distribution, respectively,

3. $\sum_f \mu(f) = 1$.

The KEF performs prediction and correction steps (similar to KF) that update the fuzzy D-S state according to the target motion model and input measurement(s), respectively (see Figure 3.1). At the prediction step, KEF estimates the next fuzzy D-S state of the system $\mu_{t+1|t}$ assuming underlying target dynamics in the original state-space $x \in X$ to be linear-Gaussian as shown below:

$$x_{t+1} = Fx_t + V \tag{3.6}$$

where vector V is a Gaussian random vector with zero mean and covariance Q. The fusion problem considered in our experiments is a single-target tracking system in 2D. Therefore, the goal of KEF is to estimate the target position, i.e. $x = [X\ Y]^T$, where X and Y are the estimated target x and y coordinates in the original space. Please note the

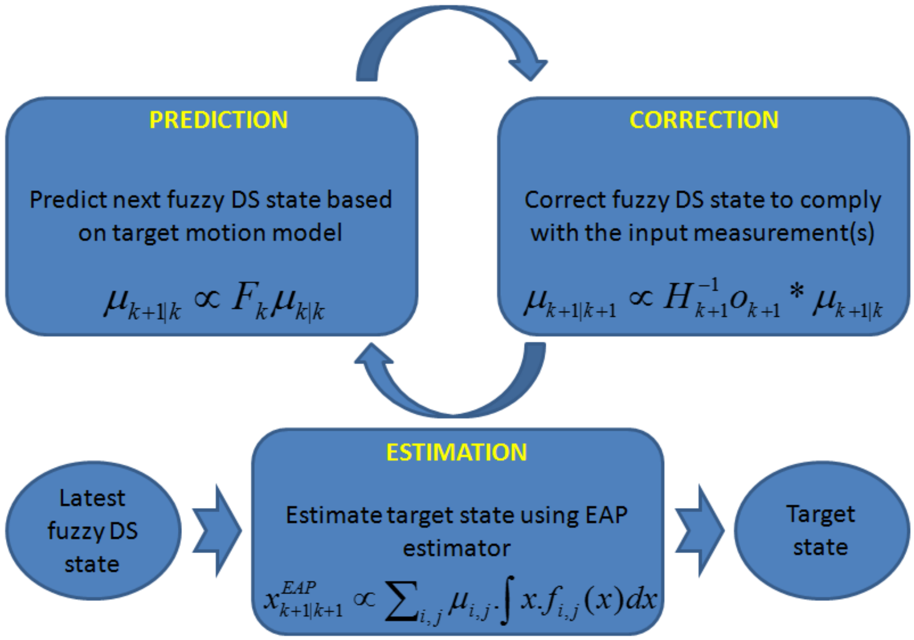

Figure 3.1: The KEF state evolution and target state estimation steps.

linearity constraint is imposed not by the RS theory but indeed by the KEF formulation to ensure all derived equations remain in closed-form. The matrix F in 3.6 represents the target motion model and is defined as

$$F = \begin{bmatrix} 1 + \frac{\tau_x}{X} & 0 \\ 0 & 1 + \frac{\tau_y}{Y} \end{bmatrix} \tag{3.7}$$

where τ_x and τ_y represent the target velocity along the x and y axes, respectively.

Given the current fuzzy D-S state of KEF at time t as $\mu_{t|t}$ with focal sets of the form $f_i(x) = N_{C_i}(x - x_i)$ for $i = 1, \ldots, e$ and corresponding mass of μ_i, the focal sets of the predicted fuzzy D-S state $\mu_{t+1|t}$ are computed as the following

$$f_i^{'}(x) = N_{D_i}(x - x_i^{'}) \tag{3.8}$$

where $D_i = Q + FC_iF^T$ and $x_i^{'} = Fx_i$.

The corresponding mass assigned to the focal sets $\mu_i^{'}$ would also be computed according to

$$\mu_i{}' = \frac{\omega_i.\mu_i}{\sum_{i=1}^{e} \omega_i^{'}.\mu_i^{'}} \tag{3.9}$$

where $\omega_i = \sqrt{\frac{detC_i}{detD_i}}$.

At the correction step, KEF assumes an underlying sensing model for the original measurement z of the form

$$z_t = Hx_t + W \tag{3.10}$$

where matrix H characterizes the sensing model and W is a Gaussian random vector with zero mean and covariance R. Given the predicted fuzzy D-S state as $\mu_{t+1|t}$ and the new fuzzy D-S measurement o_{t+1} with fuzzy focal sets $g_j(z)$ in normalized Gaussian form, i.e. $g_j(z) = N_{C_j}(z - z_j)$, and corresponding mass of o_j for $j = 1, \ldots, d$, the corrected fuzzy D-S state after fusing new measurement $\mu_{t+1|t+1}$ would have the focal sets $f_{i,j}(x)$ with the corresponding mass $\mu_{i,j}$ computed as [92]

$$f_{i,j}(x) = N_{E_{i,j}}(x - e_{i,j}) \tag{3.11}$$

$$\mu_{i,j} = \frac{\mu_i.o_j.\omega_{i,j}.N_{C_j+H^TD_iH}(z_j - Hx_i)}{\sum_{i'=1}^{e}\sum_{j'=1}^{d} \mu_{i'}.o_{j'}.\omega_{i',j'}.N_{C_{j'}+H^TD_{i'}H}(z_{j'} - Hx_{i'})} \tag{3.12}$$

where

$$e_{i,j} = x_i + K_{i,j}(z_j - Hx_i) \tag{3.13}$$

$$E_{i,j} = (I - K_{i,j}H)D_i \tag{3.14}$$

$$\omega_{i,j} = \sqrt{\frac{detC_j.detD_i}{detE_{i,j}.det(C_j + H^TD_iH)}} \tag{3.15}$$

and $K_{i,j} = D_iH^T(HD_iH^T + C_j)^{-1}$ may be considered as the KEF gain.

Note that F, H, V and W used in the sensing and target motion modeling are not required to be constant and are indeed in the form of F_t, H_t, V_t and W_t. To simplify the

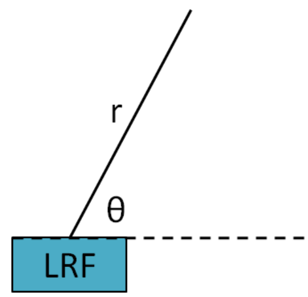

Figure 3.2: Laser range finder range measurement scheme.

notation, the subscript t has been omitted. Furthermore, one can notice that the number of Gaussian components produced by the KEF correction step grows exponentially. However, this issue can be substantially alleviated by merging similar components and removing the ones with negligible weights [180].

3.3.3 Hard Data Modeling

The target is assumed to be observed by laser range finder (LRF)[1] sensors that provide the hard data . The LRF has a $180\,^{\circ}$ field of view and supplies data as a set of range measurements L in the form shown below (see also Figure 3.2).

$$L = \left\{ r_\theta | \theta = 1, \ldots, 180 \right\} \tag{3.16}$$

The LRF deploys a simple tracking scheme based on the comparison between two consecutive measurement sets $L_{t+\Delta t_{hard}}$ and L_t to estimate the approximate relative position of the target with respect to a LRF sensor. The final hard data sent to the fusion system z_{hard} is in the form of $z_{hard} = [z_x \; z_y]^T$, where z_x and z_y are estimated target x and y coordinates with respect to the global coordinate system. Each LRF is modeled using a linear sensing model (see equation 3.10) with $H = I$ as $z_{hard} = x$ and sensor noise covariance R determined experimentally. Indeed, hard data provided by LRF are simply considered as a fuzzy D-S measurement with $d = 1$, i.e. having only a single focal set in the form of Gaussian 2D fuzzy membership function $g(z) = N_R(z - z_{hard})$.

[1]Laser Range Finder

Report = Report1 AND Report2 AND Report3

Reportx = Expression OR Expression OR Expression

Expression=< e_qualifier >< target_ID >< a_qualifier >< action >< speed >< d_qualifier >< direction >

< *_qualifier >∈ {*slightly, perhaps, almost, certainly*}

< target_ID >∈ {*robotx*}

< *action* >∈ {*moves, speeding, stopping*}

< *speed* >∈ {*slow, regular, fast*}

< *direction* >∈ {*N, NE, E, SE, S, SW, W, NW*}

Figure 3.3: Syntax considered for the soft data report.

3.3.4 Soft Data Modeling

The soft data supplied by a human observer can generally be a report provided in unconstrained natural language statement form. Nonetheless, to simplify the modeling process it is assumed that reports comply to a specific syntax and semantics determined by a predefined ontology. This assumption is valid as ideally any unconstrained report related to the target tracking application can be transformed into the desired predefined form using an appropriate natural language processing scheme. Figure 3.3 shows the syntax considered for the human-generated reports. As illustrated each report is comprised of several (up to three in the current implementation) expressions. This allows for representing any ambiguities that human agents may be willing to express regarding the true state of the target. Each expression is indeed a language sentence describing target dynamics in terms of action, speed, and direction. In addition, three qualifiers, namely expression qualifier ($e\_qualifier$), action qualifier ($a\_qualifier$), and direction qualifier ($d\_qualifier$) have been incorporated into the sentence. These qualifiers allow the human observer to express his/her confidence in the corresponding term from very high, i.e. *certainly*, to very low, i.e. *slightly*. The reported target action could be *speeding*, *stopping*, as well as *moving*. The first two possibilities are included in the new soft data ontology in order to enable our system to model richer soft data reporting on target acceleration/deceleration. This is inspired by the ability of human sensors to perform pattern recognition and report on complex high-level target dynamics, in this case variation of target speed over time, which are typically difficult to obtain using conventional hard sensors.

Once produced according to the specified syntax, each qualitative human-generated report must be interpreted to be transformed into an equivalent quantitative representation. Figure 3.4 presents the schematic of the semantics used to interpret soft data expressions.

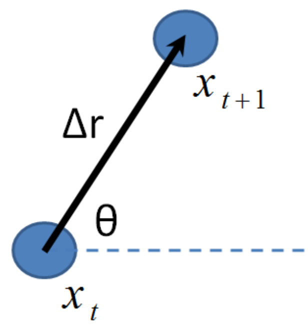

Figure 3.4: Schematic of the semantics used to interpret soft data.

As shown, target displacement from time t to $t + \Delta t_{soft}$ is determined using a vector designating the amount Δr as well as direction θ of the movement. Note that for representational convenience the Cartesian coordinate system is replaced by the polar coordinate system. Each j^{th} expression is modeled as a 2D Gaussian fuzzy membership function, i.e.:

$$g_j(z) = N_{C_j}(z - z_j) \tag{3.17}$$

which is defined on the original measurement space and where z_j and C_j represent the distribution mean and covariance, respectively. The Δr_j is determined according to target action and speed reported in the given expression, i.e. the faster the speed the larger the Δr_j. Similarly, θ_j is determined using the reported target direction. The formulation used for soft data allows for the representation of fuzziness and subjectivity inherent in the expression terms reported by the human observers. For example, let target speed be reported as "fast"; the term "fast" may have slightly different meaning to different people and is thus inherently vague. Given Δr_j and θ_j, the new target measurement z_j^t according to expression j at time t is computed as shown below

$$z_j{}^t = z_j{}^{t-1} + [\Delta r_j \times \cos(\theta_j) \times \Delta t_{soft} \; \Delta r_j \times \sin(\theta_j) \times \Delta t_{soft}]^T \tag{3.18}$$

where $z_j^0 = [X_{init} \; Y_{init}]^T$ is the initial target position estimate.

In addition, the action and direction qualifier terms of the expression j are deployed to determine the covariance C_j associated with the distribution mean (assuming Δr_j and θ_j to be independent) as shown below.

$$C_j = \begin{bmatrix} var(\Delta r_j) & 0 \\ 0 & var(\theta_j) \end{bmatrix} \tag{3.19}$$

The $var(\Delta r_j)$ and $var(\theta_j)$ are chosen directly proportional to the inverse of the confidence level expressed by the expressions' qualifier, i.e., the higher the confidence, the lower the variance. The mass assigned to each expression o_j is computed according to the normalized relative weight of that expression given by its qualifier ($e\_qualifier$). The more certain the $e\_qualifier$, the higher the assigned mass. Finally, for the soft data that report on target acceleration/deceleration, i.e. $action = (speeding\ or\ stopping)$, the corresponding target speed components are updated using the reported target positive or negative acceleration acc and direction θ as (see Table 4.1)

$$\tau_x = \tau_x + acc \times cos(\theta) \times \Delta t_{soft}$$
$$\tau_y = \tau_y + acc \times sin(\theta) \times \Delta t_{soft}. \tag{3.20}$$

3.3.5 System Organization

Our prototype soft/hard data fusion system is implemented within the multi-agent paradigm using a Java-based framework called JADE[2]. An overview of the system architecture is illustrated in Figure 3.5. There are several type of agents developed to handle different tasks involved in the system. The core of the system is the data fusion agent (DFA), where the KEF is deployed to process data and compute the target position estimate. Two separate agents are in charge of collecting soft and hard data, namely, the soft data agent (SDA) and the hard data agent (HDA). Furthermore, to interact with the human observer and obtain reports, a GUI [3] agent named the human interaction agent (HIA) is developed. Once obtained by the HIA, the raw human reports are then sent to the SDA to be validated using the predefined domain ontology discussed in the previous section. The target is assumed to be a mobile robot and is simulated using a powerful robotic simulation platform called Player/Stage [4] (P/S). Player is a server that acts as a hardware abstraction layer and allows different client programs (usually written in C/C++) to communicate with the sensors/actuators of the robot through simple standard interfaces. On the other hand, Stage is the underlying software that provides the simulated robotic platform as well as experimental environment to the Player. Another agent is called the actuation agent (AA), which controls target (robot) movement precisely and also reports on the exact robot position, which is used as the ground truth in the target tracking experiments presented in chapters 4 and 5. The data preprocessing agent (DPA) is in charge of preprocessing both

[2]Java Agent Development Framework
[3]Graphical User Interface
[4]http://playerstage.sourceforge.net

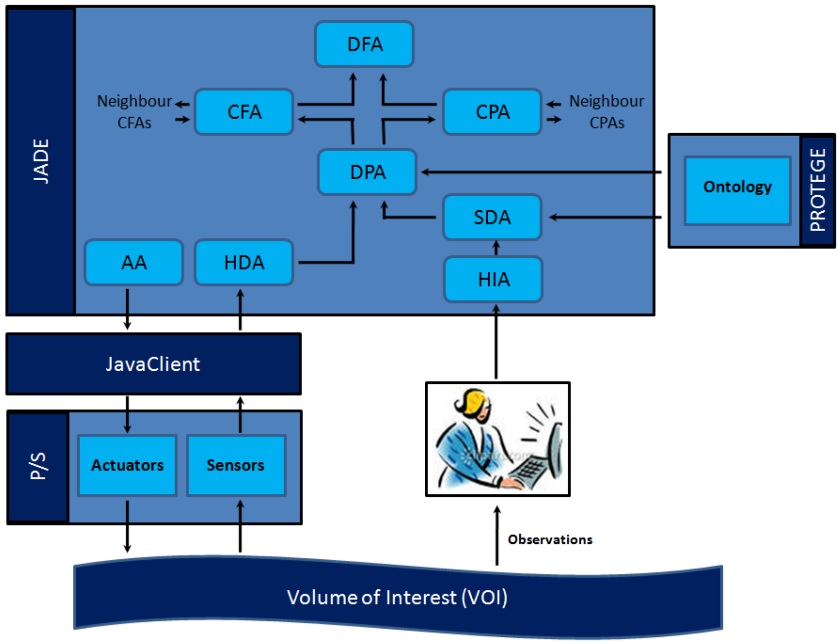

Figure 3.5: The multi-agent organization of developed fusion system.

soft and hard data and forwarding results to the consensus propagation agent (CPA) and the consensus filter agent (CFA) for aggregation, respectively, as discussed later in section 4.3.

The communication among agents is accomplished using a message-passing mechanism, called agent communication language (ACL), which complies to FIPA [5] standards. Furthermore, as P/S is compatible only with C/C++ and our agents are developed using Java in JADE, a middleware library called Javaclient is used to bridge between the two software platforms. The main advantage of the multi-agent architecture are the ease of extension and scalability, which are highly desired for the future development of our fusion system. In addition, both P/S and JADE are available for free as open source software.

3.4 Chapter Summary

In this chapter a novel soft/hard data fusion system based on the RF theory was presented. Both soft/hard data fusion and the RS theory are rather new areas of research in the

[5] Foundation for Intelligent Physical Agents

data fusion community. The developments described in this chapter support the notion of a RS theoretic approach to fusion of soft/hard data. Compared to other alternative approaches of dealing with data uncertainty (imperfection), the RS theory appears to provide the highest level of flexibility in dealing with complex unconventional data while still operating within the popular and well-studied framework of Bayesian inference. The prototype system introduced is by no means complete and is intended merely to serve as the core of our framework to be extended in the subsequent chapters.

Chapter 4

RS Theoretic Single-target Tracking Using Soft/Hard Data

4.1 Introduction

This chapter presents an extension of our prototype fusion system described earlier in chapter 3. The main contributions of this chapter can be summarized as the following. First, a scheme based on consensus propagation is developed to enable distributed aggregation of soft data requiring only local information exchange among neighbors. Most importantly, compared to other distributed fusion schemes, our CP-based aggregation has the advantage of allowing a sensor node to exclude itself from the aggregation processes, i.e. provide no informational contribution, while participating in the data propagation process across the network. This characteristic is very useful when aggregating soft data as described later in this chapter. Second, a human trust modeling approach is proposed that allows for on-the-fly estimation of agent trustworthiness and thus avoiding misleading or erroneous soft data to some extent. Extensive experimental results obtained for a single-target tracking application demonstrate the improved tracking performance and efficiency achieved by the proposed fusion system through integrating soft and hard data.

The growth in popularity of sensor networks has led to an ever-increasing demand for the development of highly scalable data fusion algorithms suitable for processing large amount of data provided by such vast networks. This has attracted a lot of research on decentralized data fusion architectures where the centralized fusion node is not required. Among the existing decentralized schemes, the fully distributed architectures are the most appealing as they rely only on local information exchange among neighbors to achieve a global outcome and hence are highly scalable. Moreover, distributed data fusion promises to yield improved efficiency, reliability, and ease of deployment. The large volume of research work published on distributed data fusion methodologies using the gossip-based [172], message passing [173], and diffusion processes [174], is an indicator of the growing interest in this subject. We propose a novel distributed data fusion approach applicable to the fusion of soft data based on the consensus propagation (CP) algorithm [175]. On the other hand, the success of a data fusion system depends on two main factors, namely, the quality of input data and the efficiency of the adopted fusion algorithm (operator). The former factor in turn is characterized by how well data is represented, how accurate the prior knowledge is, and most importantly how trustworthy the source of data is deemed to be [156]. Accordingly, the level of trust assumed for the source of data has a major role in the overall performance of a fusion system. Nonetheless, the majority of the fusion systems today are based on an optimistic presumption about the trustworthiness of data sources that they are all truthful. In context of the fusion systems dealing with soft data produced by human agents this issue is even more significant as humans are typically far more susceptible to failure and malfunction compared to machines (sensors). This observation, has motivated us to address the problem of human trust modeling in the proposed fusion system.

The rest of this chapter is organized as follows. An overview of the related work in distributed data fusion and data trustworthiness issue is provided in section 4.2. The distributed data aggregation schemes for both soft and hard data are presented in section 4.3. Section 4.4 is dedicated to the discussion of the human trust modeling approach. Our experiments with single-target tracking application are discussed in section 4.5. Finally, our conclusive remarks and potential avenues of future work are presented in section 4.6.

4.2 Background and Related Work

This section presents a summary of related work on the two areas we are contributing to, namely, distributed data fusion, and data trustworthiness modeling. A more in-depth exploration of these research areas could be found at the review studies published in [7, 176, 97].

4.2.1 Distributed Data Fusion

In a distributed data fusion (DDF) architecture each node in the sensor field operates as a local data fusion unit relying on local measurements collected from its neighbors. The fusion process is typically repeated iteratively and ideally converges to the same global outcome as a centralized data fusion scheme. This approach eliminates the need for the complicated and power-demanding process of routing data packets across the entire field, to be received and processed by the centralized fusion node (sink), and thus promises to enhance the scalability, efficiency, reliability, and ease of deployment of the fusion system. On the other hand, the main drawback of a distributed fusion architecture is the inherent lack of global structure, which would make it challenging to control the fusion process.

Early work on distributed data fusion attempts to simplify this problem by assuming a fully connected network [181, 182], i.e. every sensor can communicate with every other sensor, which is not a valid assumption for the majority of vast sensor networks of modern days. One may refer to such scheme as decentralized to distinguish it from the fully distributed systems discussed above. Modern distributed fusion systems operate within three main computational paradigms:

- Gossip-based signal processing: local statistics are exchanged iteratively till convergence to the desired global statistics [172]. The popular consensus filter [183] belongs to this category of algorithms.
- Message passing algorithms: similar in nature to the gossip-based methods and believed to be closely related theoretically. The main difference is that the exchanged messages are usually in form of tables (representing the so-called potential functions) and the message computation process is slightly different for each of the neighbor nodes. The loopy belief propagation [173] is a well-known algorithm in this category. Compared to the gossip-based methods, distributed fusion using the message passing approaches is far less studied.
- Diffusion process: a local convex combination (weighted sum) of the estimates provided by all the neighbors is computed at each node, with no iterations, to enable information diffusion across the sensor field [174]. The weights used in the diffusion process must be either pre-calculated or learned at run-time.

The gossip-based distributed algorithms are the most commonly used approach to achieve distributed fusion and thus are well studied in the community. However, they are mostly limited to average consensus filtering. Furthermore, due to their iterative nature they enforce two time-scales of operation, one to achieve consensus iteratively and another to perform fusion using the aggregated data. The diffusion processes on the other hand are not iterative and hence resolve the time-scale issue. Also, the weights used in the diffusion process may be learned and thus adapted to the requirement(s) of the fusion task at hand.

Message passing algorithms have already proven to be an efficient approach to perform distributed inference on graphs and yet are rarely deployed to develop distributed fusion systems. Further research on these new alternatives to the popular gossip-based methods is required to examine their inter-relationship and possibly come up with hybrid schemes for distributed processing that provide a well-balanced compromise. An example of such hybrid approaches is the consensus propagation [175] algorithm, which can be considered as a special case of belief propagation devised to achieve consensus among several sensor nodes, similar to the average consensus algorithm. Our distributed aggregation scheme presented for soft data is based on the consensus propagation algorithm.

4.2.2 Data Trustworthiness

The bulk of the data fusion literature is built on optimistic premises regarding the trustworthiness of fusion data. For instance, sensory data are commonly considered to be produced by highly and equally trustworthy sources. Nonetheless, in many practical cases such assumptions are too simplistic and lead to performance degradation of the fusion system. A recent trend in fusion research attempts to address this issue by explicitly modeling the data source trustworthiness. This is mostly accomplished by introducing the notion of a second level of uncertainty, i.e. uncertainty about uncertainty, represented as trustworthiness coefficients. The challenges are how to estimate these coefficients, as well as how to incorporate them into the fusion process. A number of approaches to estimate trustworthiness coefficients have been proposed which rely on domain knowledge and contextual information [157], learning through training [158], possibility theory [159], and expert judgments [160]. In addition, the problem of trustworthiness modeling has been studied within several fusion frameworks such as Dempster-Shafer theory [161], fuzzy theory [69], transferable belief model [162], and probability theory [163]. More recent work also investigates the impact of data trustworthiness on high-level data fusion [164].

In spite of the preliminary studies mentioned above, the issue of data trustworthiness is still not well understood in the fusion community, and several open questions such as interrelationship between trustworthiness coefficients, trustworthiness of heterogeneous data, a comprehensive architecture to manage data fusion algorithm and trustworthiness of data sources, and even a standard (or at least well-accepted) definition of the concept trust within the fusion process remain as a part of future research [7, 161]. Furthermore, virtually all of the studies so far target sensors providing hard data and hence are not directly applicable to modeling trustworthiness in case of soft data provided by human agent(s). Thus, our pioneering approach, discussed in the next section, is among the first attempts to address this highly unexplored issue.

4.3 Distributed Data Aggregation

In order to achieve distributed fusion of soft/hard data, we propose to first compute the global average of hard and soft data in a distributed way at each of the sensor nodes and then supply the outcome to the individual KEF of each sensor node. It is assumed that each sensor node is always provided with hard data and may (or may not) be provided with soft data. Accordingly, computing the global average of hard data is straightforward and is accomplished using the popular average consensus filter (CF) method [184]. Computing the global average of soft data, on the other hand, is more challenging as some sensor nodes may not be contributing to the data aggregation process and hence the CF approach may not be deployed. We propose to deploy a distributed inference approach called consensus propagation [175] to deal with the distributed soft data aggregation problem. The key advantage of CP is that it propagates the number of contributing sensor nodes along with the estimate of global average at any time, hence allowing a sensor node to exclude itself from the aggregation process while still propagating data.

4.3.1 Hard Data Aggregation

As mentioned earlier, the hard data z_{hard} at each sensor node provide a local estimate of the 2D position of the target, i.e. $z_{hard}^{local} = [z_x \; z_y]^T$. The local estimates are extracted in the DPA from the hard data reported by the HDA. The local estimates are then sent to the CFA in order to be used in an iterative filtering procedure. The CF is formulated based on an algebraic graph theoretic representation of the sensor network, i.e. a graph $G = \{V, E\}$ is used to show the interconnections between sensor nodes. $V = \{v_1, \ldots, v_M\}$ represents the set of graph vertices corresponding to the sensor nodes in the sensor network while $E = \{(v_i, v_j) \in V \times V \mid v_i \textit{ and } v_j \textit{ are neighbors}\}$ is the set of graph edges. Usually two sensor nodes are considered neighbors if they are situated close enough to easily communicate over a wireless channel. Assuming $z_{hard,i}^{global,k}$ to represent the global estimate of average hard data at iteration k for node i, the discrete consensus filter at node i operates as shown below:

$$Z_{hard,i}^{global,k+1} = \\ Z_{hard,i}^{global,k} + \epsilon\big[\sum_{j \in N_i}(Z_{hard,j}^{global,k} - Z_{hard,i}^{global,k}) + (Z_{hard,i}^{local} - Z_{hard,i}^{global,k})\big] \tag{4.1}$$

where $N_i = \{j \in V | (v_i, v_j) \in E\}$ is the set of neighbor sensor nodes for node i. The ϵ is the operation rate of the CF to ensure stability of the filter and is determined according to the constraint specified as the following.

$$\epsilon \leq \frac{1}{d_{max}} \tag{4.2}$$

The d_{max} is the maximum degree of graph G, where $d_i = |N_i|$ represents the degree of sensor node i. As shown in Equation 4.1, the CF updates its estimate of the global average iteratively using the local estimate, as well as the most recent estimates received from the neighbors. The iterative updates continue until reaching convergence or a predetermined maximum number of iterations.

4.3.2 Soft Data Aggregation

One can consider the fuzzy D-S state used by the KEF as a Gaussian mixture with a varying number of components. Each component j of the mixture, corresponding to an expression in the soft report, is distinguished using three parameters, namely, its weight o_j, mean z_j, and the associated covariance C_j. On the other hand, as discussed earlier, the report given by the human observer may contain up to three expressions (hypotheses) regarding the state of the target, i.e. soft data is modeled using a maximum of three components. There is also a fourth component which models the state of total ignorance about the target, i.e. $g(z) = 1$, and requires only a weight parameter. Therefore, there are a maximum of $10(= 3 \times 3 + 1)$ parameters required to represent soft data in a numerical form. The conversion from the soft qualitative into a purely quantitative form occurs at the DPA.

The global average of these ten parameters is computed using the consensus propagation method. The CP is indeed a special case of Gaussian belief propagation that can be used to compute global average in a distributed manner. Similar to CF, CP is an iterative algorithm. However CP differs from CF in that data (messages) sent to each of the neighbor sensor nodes are specific to that node, whereas CF broadcasts the same message to all neighbors at each iteration. More importantly, using CP the message sent to each neighbor sensor node contains not only the latest estimate of the desired parameter U but also the number of sensor nodes contributing to that estimation process M_U. Mathematically, for every iteration k at node i, the incoming data $[U^{ji,k-1}, M_U^{ji,k-1}]$ previously received and stored from all neighbor sensor nodes $j \in N_i$ are used (along with the local estimate U_i^{local} if applicable) to update the global estimate $U_i^{global,k}$ of the desired parameter as

$$U_i^{global,k} = \frac{\rho_i + \sum_{j \in N_i} U^{ji,k-1} M_U^{ji,k-1}}{\kappa_i + \sum_{j \in N_i} M_U^{ji,k-1}} \tag{4.3}$$

The stored incoming data are also used to compute the (neighbor-specific) outgoing messages $[U^{ij,k}, M_U^{ij,k}]$ for each of the neighbor nodes j as shown below.

$$U^{ij,k} = \frac{\rho_i + \sum_{l \in N_i/\ j} U^{li,k-1} M_U^{li,k-1}}{\kappa_i + \sum_{l \in N_i/\ j} M_U^{li,k-1}} \tag{4.4}$$

$$M_U^{ij,k} = \frac{\kappa_i + \sum_{l \in N_i/\ j} M_U^{li,k-1}}{1 + \frac{1}{\beta}(\kappa_i + \sum_{l \in N_i/\ j} M_U^{li,k-1})} \tag{4.5}$$

$$\rho_i \ = \begin{cases} U_i^{local} & \textit{node } i \textit{ is contributing} \\ 0 & \textit{otherwise} \end{cases} \tag{4.6}$$

$$\kappa_i \ = \begin{cases} 1 & \textit{node } i \textit{ is contributing} \\ 0 & \textit{otherwise} \end{cases} \tag{4.7}$$

Here, $\beta \geq 0$ is a constant that allows one to control the attenuation level of the CP. The intuition behind the attenuation process is to avoid the unbounded growth of M_U caused by the existence of loops in the network topology. It is easy to see that the larger the M_U and smaller the β, the stronger the attenuation process would be. The convergence properties of this approach are proven and discussed in detail in [175]. The outgoing messages from node i are received and stored (as incoming data) at each of the neighbor sensor nodes j to be used in the next iteration of the CP repeating the procedure described above. This iterative procedure is initialized at the CPA upon receiving the latest local estimate of the parameters from the DPA and is repeated until reaching convergence or a predefined maximum number of iterations.

As discussed earlier, some sensor nodes may not supply soft data pertinent to the parameters of a given component. We have modified the CP (as shown in 4.6 and 4.7) using the ρ_i and κ_i parameters to allow a sensor node i to control whether or not to participate in the aggregation process while still propagating the data. The idea is that by excluding itself from the computation of the number of contributing sensor nodes M_U^k and the associated parameter of interest U^k at each iteration ($\rho = 0$ and $\kappa = 0$), a non-contributing sensor node will still receive and forward messages from and to its neighbors

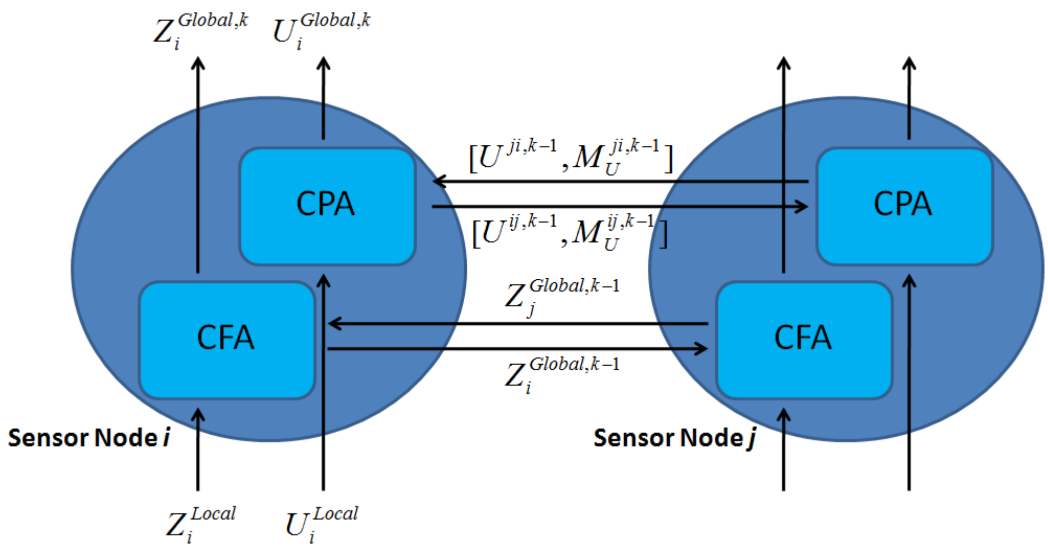

Figure 4.1: Distributed computation of global estimate of soft/hard data at each sensor node.

(propagate soft data) without affecting them by its local evidence (no contribution). At the first iteration of the CP, a non-contributing sensor node will prepare and send out its outgoing messages as $[don't\ care, 0]$. Therefore, the exclusion from the global averaging process is initiated by setting the initial number of contributing nodes to zero in the first outgoing messages. Please note that the reported estimate of the desired parameter is not important, hence treated as "don't care", as it will not be considered in the subsequent computations. Figure 4.1 demonstrates an overview of the distributed computation of global average at each sensor node i in terms of the inputs, data (message) exchange with each neighbor node j, and the produced output.

4.4 Soft Data Trustworthiness

Inspired by the work of Young and Palmer [156], we model soft data trust relying upon three core concepts, namely, belief, reliability, and credibility. The belief is defined as a conviction of the truth of a proposition usually acquired through perception. The credibility is the measure of believability of a statement, action, or source, and the ability of the observer to believe that statement based upon the consistency with other evidence. Finally, the reliability is the degree to which prior historical reports from a source have been consistent with fact.

Accordingly, in our approach we consider belief Bel as the reported confidence in a given soft report component. In the soft data ontology described earlier in section III such confidence level is represented using the expression qualifier, i.e. *e_qualifier*. Also, the reliability Re is assumed to be provided *a priori* as a function of the earlier historical performance of an agent. Lastly, the credibility Cr is formulated and calculated on-the-fly in order to measure the consistency degree of the (human) agent reports over time.

The credibility computation approach is developed in order to satisfy several design requirements shown below:

1. Agent credibility should be directly proportional to the temporal consistency of reports, i.e. the higher the frequency and the amount of qualifier variations over time, the lower the credibility,

2. The changes in the *e_qualifier* are deemed more significant than *a(d)_qualifier* variations,

3. Credibility computations are performed on-the-fly and thus must be done efficiently,

4. The effect of the latest reports must be adjustable as desired,

5. Gradual variation of the qualifiers must be tolerated to some extent

The individual credibility cr_i $s.t.$ $i \in \{1, 2, 3\}$ for each of the three possible expressions in a given soft report are computed separately and the total credibility Cr is obtained using a simple averaging, i.e. $Cr = \frac{1}{3}\sum_{i=1}^{3} cr_i$.

The individual cr_i are computed efficiently through a recursive approach (see design requirement 3) as follows

$$cr_i^k = 1 - D_i^k \tag{4.8}$$

where cr_i^k and D_i^k represent the individual credibility estimate and the normalized exponential weighted average of qualifier variations, respectively, for expression i at iteration k. The D_i^k is computed as

$$\begin{aligned} D_i^k = \alpha \times \frac{(\frac{(1-w_{eq})}{2} \times (diff_i^{act} + diff_i^{dir}) + (w_{eq} \times diff_i^{eq}))}{DIFF_{max}} \\ +(1-\alpha) \times D_i^{k-1} \end{aligned} \tag{4.9}$$

where $w_{eq} > 0.33$ is the averaging weight assigned to the variations in the *e_qualifier* (see design requirement 2), and the absolute difference between the new and old reported expression, action, and direction qualifiers are represented as $diff_i^{eq}$, $diff_i^{act}$, and $diff_i^{dir}$, respectively. The $DIFF_{max}$ is used to normalize qualifier variations by the maximum possible value, which in our system is 3(=4-1) as we assign 4 to the highest confidence level (*certainly*) and 1 to the lowest level (*slightly*). Finally, $0 < \alpha < 1$ is the parameter controlling the influence of the latest report in the exponential averaging process, i.e. the higher the α the higher the impact of recent report (see design requirement 4). In order to enforce the last design requirement, α is set according to the variation pattern/amount in a given expression as below

$$\alpha = \begin{cases} 0.2 & diff_i^{eq} \leq 1 \; and \; diff_i^{act} \leq 1 \; and \; diff_i^{dir} \leq 1 \\ 0.8 & otherwise \end{cases} \tag{4.10}$$

to restrict penalizing the small variations in the qualifiers. Once computed, the credibility Cr of the data is combined with the reported reliability of the data source Re to provide the discounting coefficient Dis, which then determines the discounting effect applied to the reported data belief Bel as below

$$Bel_{dis} = Dis \times Bel \tag{4.11}$$

As suggested in [156], assuming source reliability parameter to take the form of the Beta distribution, i.e. $Re \sim Be(r_1, s_1)$, the single parameter combining data credibility and source reliability (Dis in our case) would also be of the Beta distribution form and $Dis \sim Be(r1 + Cr, s1 + 1 - Cr)$. Accordingly, the expected value for the discounting coefficient would be computed as

$$E(Dis) = \frac{r1 + Cr}{r1 + s1 + 1} \tag{4.12}$$

The parameters $r1$ and $s1$ distinguishing the original Beta distribution for source reliability are computed as follows

$$s1 = (n - 1)(1 - Re) \tag{4.13}$$

$$r1 = \frac{Re \times s1}{1 - Re} \tag{4.14}$$

where parameter n represents the (minimum) number of historical reports assessed to evaluate the reliability of a data source and typically $n = 8$ is used. The soft data trust modeling computations are performed by the DPA. Once discounted using the calculated discounting coefficient Dis (see 4.12), soft data expressions are then categorized based on their final belief Bel_{dis} into three categories, namely, certain-level ($0.75 < Bel_{dis} \leq 1$), almost-level ($0.5 < Bel_{dis} \leq 0.75$), and perhaps-level)$0.25 < Bel_{dis} \leq 0.5$) before being sent to the CPA.

4.5 Single-target Tracking Experiments

Although there are some preliminary works aiming at creating a standard data set to enable evaluation of soft/hard data fusion systems [150, 151], there are no publicly available data sets for this purpose at this point. Therefore, we have conducted a series of experiments ourselves, which are designed to highlight the performance enhancements that can be achieved through soft data integration into the fusion process. Although our fusion system is capable of operating, i.e. tracking target(s), while relying only on soft reports, the experiments presented in this section mostly deploy soft data as a complementary source of information to correct, improve, or update the target motion model, i.e. matrix F in 3.6. We first present the results of the preliminary experiments performed to evaluate the performance of the developed system for single-target tracking using soft/hard data. Next we discuss the results of three series of experiments designed to further evaluate the performance enhancements achieved using the proposed distributed data aggregation approach, soft data trust modeling scheme, as well as the soft data provided through the new augmented ontology, respectively.

4.5.1 Experimental Setup

The performance metric used for the experiments is the average position tracking error (ATE), i.e. the cumulative average of the Euclidean distance between the known and the estimated target position. The known target position (ground truth) is provided by the AA, while the estimated position is the output of the DFA. Each of the experimental scenarios is repeated five times in an attempt to obtain a more realistic measure of the system performance. The simulation time resolution is 1ms. The SDA and HDA supply their data to the DPA every 500ms and 100ms, respectively, i.e. $\Delta t_{soft} = 500ms$ and $\Delta t_{hard} = 100ms$. Please note that the agent may input his/her report once (if desired) and SDA will produce an appropriate soft data, according to the given ontology, and sends

Table 4.1: Soft data interpretation parameter setting

Reported data	**Corresponding parameter**	**Corresponding value(s)**
moves slow/regular/fast	Δr	0.2/0.5/0.8 (m/s)
speeding slow/regular/fast	acc	0.05/0.1/0.2 (m/s^2)
stopping slow/regular/fast	acc	-0.05/-0.1/-0.2 (m/s^2)
E/NE/N/NW/W/SW/S/SE	θ	$0/\frac{\pi}{4}/\frac{\pi}{2}/\frac{3\pi}{4}/\pi/\frac{5\pi}{4}$ $/\frac{3\pi}{2}/\frac{7\pi}{4}$ $(radians)$

it to the DPA periodically. The DPA interprets both soft and hard data and sends the extracted numerical values to the CPA and the CFA, respectively. In particular, the soft data interpretation provides the parameters shown in Table 4.1, which once aggregated are used in the DFA to determine the functions $g_j(z)$ (see 3.17) and/or matrix F (see 3.6). For target motion model update, the soft data is used only if its aggregated discounted belief Bel_{dis} is more than almost-level threshold. Please note that in each of the experimental cases the target dynamics, human agent reports, and other characteristics (e.g. agent reliability Re) are tailored to test for a specific feature of the fusion system. In all cases, the enhancement of the tracking performance is reflected by the improvement percentage of the ATE.

4.5.2 Preliminary Experiments

The basic system performance was evaluated with two categories of experiments distinguished depending on the relative role of soft data within the fusion process as described in the following.

Soft data as necessary source of information

The target is tasked to move from the departure to the destination point such that it is observable by the LRFs at the beginning and final portions of its path, but not the middle portion (see Figure 4.2). The human observer is assumed to be able to observe the target throughout the middle portion. Therefore, in order to compute a continuous track of target position, soft data must be deployed.

Table 4.2 shows the experimental results obtained for five different trials of the system with and without soft data deployment. Please note that the ATE computed when soft

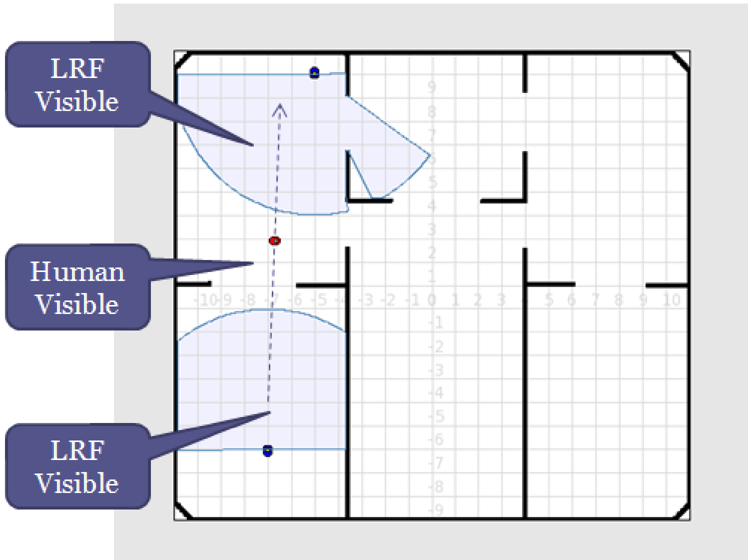

Figure 4.2: Snapshot of the simulation arena in Stage used to evaluate soft data as necessary information source: the blue circles represent the LRFs while target is designated in red

data is not included is only for the portions of the target path where it is visible to the LRF and the error associated with the invisible portion is simply ignored. The target path is set in such a way that it moves in a north-east bound direction with a speed that is considered slow to the human observer and the supplied human report is "almost robot certainly moves slow almost N". Experimental results show that using soft as well as hard data, the target track can be estimated continuously with ATE that compares favorably to the performance of LRFs alone. Indeed, the average ATE computed over the five trials is even slightly lower in the case of soft data inclusion. This can be attributed to the fact that using soft data allows for initializing the second LRF, used in the final portion of target path, with a very good estimate of target position. It is worth mentioning that several other experiments with different path planning schemes for the target were performed and resulted in a similar behavior for the system.

Soft data as complementary source of information

The experiments in this category are designed in order to evaluate the role of soft data as a complementary source of information and enhancing tracking performance. Three different test scenarios have been implemented as shown in Figure 4.3. In all scenarios

Table 4.2: Average tracking error obtained without and with soft data inclusion

Trial No.	**Soft Data Excluded**	**Soft Data Included**
1	0.177	0.155
2	0.169	0.165
3	0.172	0.160
4	0.173	0.164
5	0.171	0.156
Avg	**0.172**	**0.160**

the target is assumed to be maneuvering and soft data provided by human observer is deployed to update the target motion model, i.e. matrix F in 3.6. The soft data report is first transformed into a quantitative representation using the approach described in section 3.3.4. Next, the terms τ_x and τ_y of the target motion model are determined according to the components Δr and θ of soft data as projected onto the x and y axis, respectively, i.e.

$$F = \begin{bmatrix} 1 + \Delta r \times \cos(\theta) \times \Delta t_{soft} & 0 \\ 0 & 1 + \Delta r \times \sin(\theta) \times \Delta t_{soft} \end{bmatrix}. \tag{4.15}$$

Figure 4.3(a) shows the path the target is tasked to traverse for the first scenario. Along its path the target maneuvers twice. Each time the target changes its movement, a new soft data supplied by a human observer is used to update the matrix F. Table 4.3 provides a summary of the results obtained for five different trials. Comparing the average ATE computed for cases without and with soft data inclusion, it is clear that system performance can be enhanced by about 9% through soft data inclusion in the fusion process.

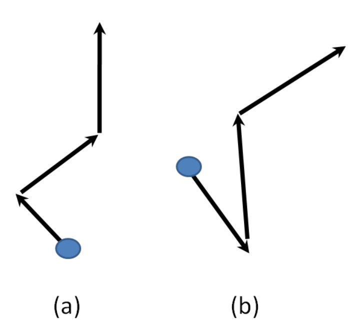

Figure 4.3: The three scenarios used to evaluate soft data as complementary information

Table 4.3: Average tracking error obtained for tracking of maneuvering target

Trial No.	**Soft Data Excluded**	**Soft Data Included**
1	0.262	0.240
2	0.268	0.246
3	0.287	0.241
4	0.261	0.247
5	0.264	0.247
Avg	**0.268**	**0.244**

The experiments for the second scenario are similar to the first scenario, with the difference that in this case the target's change of direction is more drastic (see Figure 4.3(b)) and therefore the effect of having a more accurate motion model (using soft data) should be magnified. This claim is indeed supported by the experimental results shown in Table 4.4, as the achieved performance improvement for this more challenging experiment is about 17%.

Table 4.4: Average tracking error obtained for tracking a more challenging maneuvering target

Trial No.	**Soft Data Excluded**	**Soft Data Included**
1	0.377	0.301
2	0.359	0.303
3	0.352	0.298
4	0.360	0.289
5	0.343	0.291
Avg	**0.358**	**0.296**

The third scenario is intended to show the importance of soft data as a means to improve system performance in case the prior information available regarding the target motion is wrong or misleading. As shown in Figure 4.3(c) the target is tasked to move northwards (N). Nonetheless, the initial matrix F provided to the fusion system is such that it predicts southward (S) movement for the target. In this case, the soft data provided by the human observer is used to correct for the the misleading motion model, and the obtained results are presented in Table 4.5. It can be seen that the role of soft data in this scenario is even more significant, as the achieved improvement in tracking performance approaches 30%.

Table 4.5: Average tracking error obtained for tracking with misinforming prior information

Trial No.	**Soft Data Excluded**	**Soft Data Included**
1	0.225	0.171
2	0.222	0.168
3	0.219	0.169
4	0.215	0.174
5	0.224	0.169
Avg	**0.221**	**0.170**

The experimental results presented in this section demonstrated the important role of soft data as either an essential or complementary source of information for the single-target tracking application. It was shown that incorporating soft data, the developed filter was able to achieve almost the same level of accuracy as that obtained with hard data for tracking the target within areas not visible to the typical physics-based sensors. This capability is specially conducive to military urban operations where there are many civilian/residential areas that can only be covered by human agents.

Furthermore, soft data proved to be a very useful source of information by allowing up to 30% improvement in tracking accuracy for cases where the presumed motion model of the target was allowed to be corrected (updated) using soft data supplied by a human obsever. The experiments described in the subsequent sections evaluate the advanced characteristics of the proposed data fusion framework.

4.5.3 Distributed Data Aggregation Experiments

This section presents results of the experiments performed to evaluate the performance of the developed distributed soft/hard data fusion approach as applied to the problem of single-target tracking. The experiments are performed in two categories. In both categories, soft data are used as a complementary source of information deployed to correct for the misinformed prior knowledge regarding the target motion. In the first category, an experiment is performed to validate the efficiency of the CP in computing the global average of soft data and compare it to the results obtained using the CF, and thus show the superiority of the CP. The second category of experiments aims to demonstrate the advantage of diffusing soft data through the sensor network using the proposed CP-based approach. It is shown how this enables the tracking system to deal with a misleading source of soft

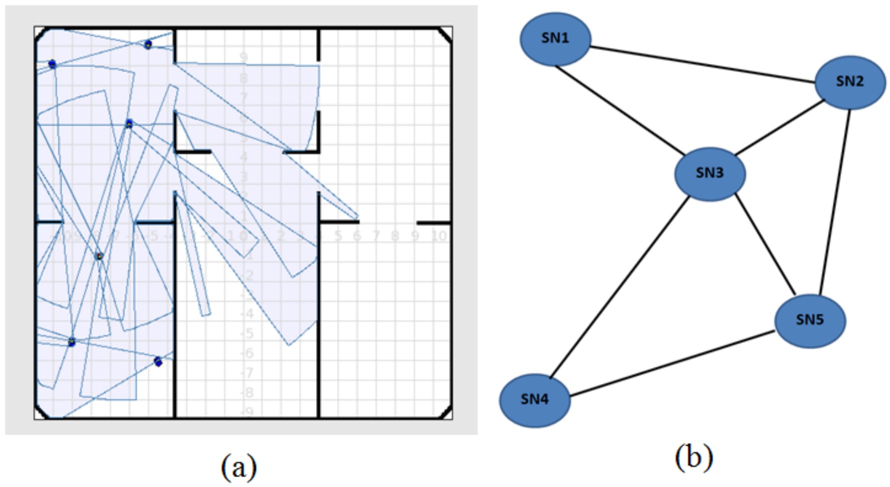

Figure 4.4: (a) A snapshot of the P/S simulation arena. (b) The graph representation of the sensor network simulated in the experiments.

data by fusing its report with other sources while relying only on local communication with neighbors.

For the experiments, the target robot is tasked (using the AA) to move according to an *a priori* known trajectory. The sensor network simulated for the experiment is comprised of five sensor nodes. Figure 4.4(a) shows a snapshot of the Stage simulation arena illustrating the positioning of human agents (blue circles) and initial position of the target (red circle). Figure 4.4(b) shows a graph representation of this network where each sensor node is depicted as a vertex and the interconnection among the nodes is depicted by the existence/non-existence of an edge.

CP vs. CF: distributed computation of global average soft data

For this experiment, it's assumed that all sensor nodes except SN3 have a source of soft data (e.g. human observer) associated with them. The target is tasked to move towards the north (upward) while the prior knowledge used to model the target motion is set to indicate a downward motion towards south. All four of the available soft data are set to be roughly correct and indicative of an upward target motion as shown in the Table 4.6.

The global average of the soft data is then used to update the matrix F and hence improve the performance of the distributed tracking system. In the first scenario, the CP approach is used to compute the global average of the four soft data reports. The

Table 4.6: CF vs. CP for distributed averaging of soft data: soft reports

Sensor node	**Soft report**
1	almost target certainly moves regular almost N
2	certainly target almost moves slow perhaps N
4	almost target perhaps moves regular almost NE
5	slightly target almost moves slow certainly NW

Table 4.7: CF vs. CP for distributed averaging of soft data: results

Trial	**CF simulated**	**CP**
1	0.232	0.153
2	0.203	0.161
3	0.228	0.148
4	0.183	0.154
5	0.198	0.145
ATE	**0.208**	**0.152**

second scenario evaluates CF applied to solve the global averaging problem. Table 4.7 shows the results obtained in terms of the average tracking error of all sensor nodes for each of the five different trials along with the final ATE. It can be seen that the final ATE produced by the CP-based distributed global averaging process provides a lower final ATE thus enhancing tracking performance. This is due to the fact that CP-based algorithm allows the exclusion of the (non-contributing) SN3 from the averaging process and yields the global target direction specified by the angle θ, which is computed over four (not five) sensor nodes and is thus closer to the desired value. On the other hand, using the CF to perform the distributed global averaging the non-contributing sensor node cannot exclude itself from the process. One might think that setting the reported value to zero for non-contributing nodes, could produce the same effect as excluding that node from the averaging process. Indeed this is how the effect of CF-based averaging is simulated at SN3 in this experiment. However, the lower quality of the tracking results obtained with this approach shows the inefficiency of such a solution.

Table 4.8: CF vs. CP for distributed averaging of soft data: soft reports

Sensor node	**First hypothesis**	**Second hypothesis**
1	perhaps target almost moves slow almost N	certainly target almost moves slow certainly S
2	almost target perhaps moves regular almost N	perhaps target certainly moves slow certainly S
3	almost target perhaps moves fast almost N	slightly target certainly moves slow certainly S
4	almost target slightly moves regular slightly NW	slightly target certainly moves regular almost S
5	certainly target perhaps moves regular almost N	perhaps target almost moves fast perhaps S

Dealing with erroneous/misleading soft data through distributed averaging

This experiment is designed and conducted to evaluate the efficiency of distributed soft data aggregation in enabling the tracking system to alleviate the impact of erroneous/misleading soft data. As in the earlier experiment, the global average soft data is used to update the incorrect prior information of the target motion model, i.e. target is tasked to move upward while the prior information indicates a downward motion. Two scenarios are considered. In scenario one, there are no soft data available and tracking is performed using the misinformed prior. In the second scenario, each of the five sensor nodes are assumed to have an associated source of soft data (human observer), producing five reports each comprised of two hypotheses (expressions) regarding the movement of the target.

As shown in table 4.8, all soft reports are set to express a higher level of certainty for the (roughly) correct hypothesis regarding the direction of target movement except for the SN1. To update the matrix F, first the global average of each of the two hypotheses are computed using the CP-based approach developed in this paper. Next, the hypothesis with the higher level of aggregated certainty is chosen to update the matrix F. Table 4.9 shows the results obtained (in terms of tracking performance) at sensor node 1 for the first and second scenarios described above. One can easily see a large percentage of improvement in tracking performance for scenario two, where the misinformed prior information is corrected for using the aggregated soft reports. This is very interesting because as discussed earlier the soft report provided to SN1 is misleading, i.e. has higher certainty associated to the wrong expression. Nonetheless, the aggregated soft report at this sensor is correctly estimated and hence once used to update the matrix F yields a drastic improvement in tracking performance.

Table 4.9: Distributed soft data fusion to correct misinformed prior: average tracking error

Trial	**Soft data excluded**	**Soft data included**
1	0.532	0.233
2	0.490	0.252
3	0.528	0.206
4	0.420	0.232
5	0.560	0.244
ATE	**0.506**	**0.233**

4.5.4 Soft Data Trust Modeling Experiments

The experiments discussed in this section are divided into two categories. The objective of the first category of experiments is to demonstrate the advantages of excluding the misleading soft reports from the fusion process. This is accomplished by using the proposed trust modeling scheme to discount the temporally inconsistent reports, i.e. reports with drastic change in qualifiers over time. On the other hand, the experiments in the second category are conducted in order to evaluate the ability of the developed fusion system to tolerate gradual change of the qualifiers. The observation is that in some tracking scenarios accurate reports could change gradually over time and if not tolerated and discounted heavily the performance would suffer. For instance, for the case where target moves towards the human agent and is thus observed more and more accurately, the confidence level expressed in the provided soft data would gradually increase.

There are two cases in the first category of experiments. In both cases target is tasked to move with slow speed towards the north (N). It's assumed that the prior knowledge regarding the target dynamics is wrong and the aggregated soft data are used to correct this prior. In case one, all agents initially provide certain-level, i.e. $e\_qualifier = certainly$ for the first report expression, and high-quality reports. However, one (SN1 in scenario one) or two (SN1 and SN2 in scenario two) of the human agents deviate from the truth and supply low-quality and eventually misleading reports over time while incurring drastic changes in the qualifiers from one report to the next. This results in a decrease in the credibility measure for these agents, which in turn causes their misleading reports to be aggregated at the almost-level (instead of certain-level) and hence effectively excludes them

Table 4.10: Avoiding misleading soft data (Case 1): initial soft data provided by all five agents at time stage one

Time stage	**SN1**	**SN2**	**SN3**	**SN4**	**SN5**
1	certainly target almost moves slow almost N OR almost target certainly moves slow certainly NE	certainly target perhaps moves slow almost N OR almost target certainly moves slow certainly SW	certainly target almost moves slow almost N OR perhaps target certainly moves slow certainly SE	certainly target slightly moves regular slightly NW OR perhaps target certainly moves regular almost SE	certainly target almost moves slow almost N OR almost target almost moves fast perhaps SW

from the aggregated reports used to update target motion model. Table 4.10 shows the initial reports provided by all agents at time stage one. The soft reports provided by SN1 at the subsequent time stages (used in scenarios one and two) are shown in Table 4.11. Please note the reports provided by other agents need not to be shown as they remain virtually unchanged throughout. Similarly, the SN2 soft data used in scenario two are also shown in Table 4.12.

The scenarios for the case two of the first category are similar to the case one scenarios. The difference is that the misleading agents reports are certain-level while the other agents in this case provide high-quality and almost-level reports. Thus, discounting the misleading reports allows for fusing them with the other high-quality almost-level reports and updating the target motion model using the obtained almost-level aggregated data. Moreover similar to the case one, the results are obtained with one (only SN1) and two (SN1 and SN2) nodes acting as the misleading agents. Figure 4.5 illustrates the enhancements in tracking performance obtained by enabling the trust modeling scheme as the improvement percentage of ATE.

Table 4.11: Avoiding misleading soft data (case 1): the SN1 soft data over time with drastic qualifier variations

Time stage	**Soft report**
2	certainly target perhaps moves slow perhaps S OR almost target perhaps moves slow slightly NE
3	certainly target almost moves slow certainly S OR slightly target perhaps moves slow certainly SE
4	certainly target almost moves slow perhaps SE OR perhaps target slightly moves slow slightly SE
5	certainly target almost moves slow certainly SE OR perhaps target certainly moves slow certainly E
6	certainly target perhaps moves slow perhaps SW OR almost target almost moves slow perhaps E

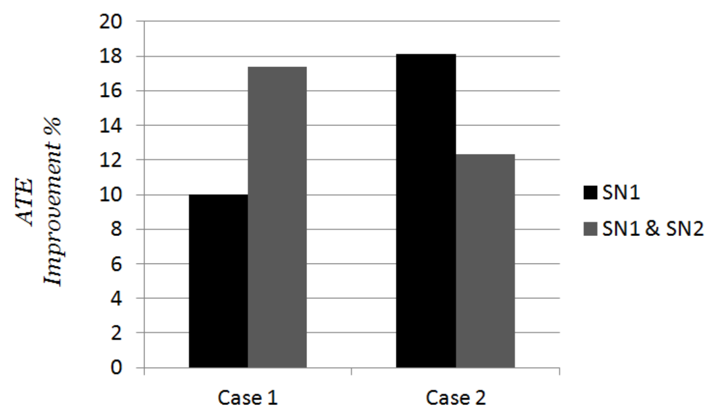

Figure 4.5: Avoiding misleading soft report(s): obtained ATE improvement percentages

Examining the results shown in Figure 4.5, we make several observations. For a single misleading agent (SN1), our approach is more effective in case two than case one. It is reasonable as without trust modeling in case two only the misleading report produced by SN1 is used to update the motion model, whereas in case one the misleading report (if not excluded using the trust modeling) is still aggregated with other high-quality reports and then used to update the target motion model. On the other hand, for the scenarios involving two misleading agents the performance improvement is more noticeable for the

Table 4.12: Avoiding misleading soft data (case 1): the SN2 soft data over time with drastic qualifier variations

Time stage	**Soft report**
2	certainly target certainly moves slow slightly E OR perhaps target slightly moves slow perhaps S
3	certainly target perhaps moves slow certainly SE OR almost target almost moves slow almost S
4	certainly target certainly moves slow perhaps SE OR perhaps target slightly moves slow perhaps S
5	certainly target perhaps moves slow certainly SW OR almost target certainly moves slow certainly S
6	certainly target almost moves slow perhaps SW OR perhaps target almost moves slow slightly SE

case one rather than case two. This could be contributed to the fact that in case one trust modeling allows us to exclude more misleading reports from the aggregation process while in case two even with trust modeling enabled both of the misleading reports are still aggregated with the other reports to compute the final report. In summary, for the case one scenarios our approach becomes more and more effective for higher number of misleading agents existing in the network whereas for scenarios of the second case the trend is reversed.

As mentioned earlier, the second category of experiments evaluate the effect of dealing with gradual change of qualifiers in the obtained performance. The effect of disabling the tolerance to gradual changes mechanism is simulated by constantly setting $\alpha = 0.8$. Similar to the first category experiments, two experimental cases each comprised of two scenarios have been conducted. In both cases, target dynamics are the same and involve three stages as follows. First, target is tasked to move with slow speed towards north west (NW). Next, target is tasked to move regular towards north east (NE). Finally, it is tasked to move fast towards south (S). For case one, all agents provide certain-level reports where some are high-quality and gradually changing and others are medium-quality and not much changing over time. The goal is to maintain the high-quality reports in the aggregation process by restricting the discounting effect of trust modeling. The experiments are performed for two scenarios where there are only one (SN1) and three (SN1, SN3, and SN5) agents providing such high-quality reports.

Table 4.13 shows the high-quality reports provided by SN1 over six time stages (scenario

Table 4.13: Tolerating gradual qualifier variations (Case 1): the SN1 soft data over time with gradual qualifier variations

Time stage	**Soft report**
1	certainly target almost moves slow perhaps NW OR almost target certainly moves slow certainly N
2	certainly target certainly moves slow almost NW OR perhaps target almost moves slow almost N
3	certainly target almost moves regular certainly NW OR almost target perhaps moves slow perhaps N
4	certainly target perhaps moves regular almost NE OR perhaps target slightly moves slow slightly NW
5	certainly target almost moves fast perhaps S OR slightly target perhaps moves slow perhaps S
6	certainly target perhaps moves fast almost S OR perhaps target slightly moves slow slightly E

one). Similarly, Table 4.14 shows the reports of SN3 and SN5 over time used in the scenario two experiments. Due to space limitations, we only show the reports provided by SN2 as an example of medium-quality reports, which are provided by other human agents (see Table 4.15). Examining the results illustrated in Figure 4.6, it is clear that by tolerating the gradual changes in qualifiers and retaining the high-quality reports in the data aggregation process, the tracking performance is improved especially for the second scenario with three agents providing high-quality and gradually changing reports the improvement achieved is rather significant (over 25%).

The second case of experiments in this category are performed assuming all reports to be almost-level. Once again, two scenarios are considered where the majority, i.e. either 3 or 4 (out of 5) of the human agents provide high-quality and gradually changing reports while the rest (SN1 and SN3) provide low-quality reports and incur abrupt change in qualifiers. Table 4.15 shows the data provided by SN1 over time. The high-quality data provided by the agents majority is produced similar to the first category experiments and is thus not shown. The fusion system is expected to retain and use the high-quality reports by limiting the discounting effect of the trust modeling scheme.

Table 4.14: Tolerating gradual qualifier variations (Case 1): the SN3 and SN5 soft data over time with gradual qualifier variations

Time stage	**SN 3 soft report**	**SN 5 soft report**
1	certainly target perhaps moves slow almost NW OR perhaps target certainly moves slow certainly S	certainly target certainly moves slow certainly NW OR slightly target almost moves fast perhaps S
2	certainly target almost moves slow certainly NW OR slightly target certainly moves slow certainly S	certainly target almost moves slow almost NW OR perhaps target almost moves fast perhaps S
3	certainly target certainly moves regular certainly NE OR perhaps target almost moves slow certainly N	certainly target perhaps moves slow perhaps NE OR slightly target almost moves fast perhaps S
4	certainly target certainly moves regular almost NE OR almost target almost moves slow almost N	certainly target slightly moves slow almost NE OR slightly target almost moves fast slightly S
5	certainly target certainly moves fast certainly S OR perhaps target almost moves slow almost N	certainly target almost moves fast almost S OR slightly target almost moves fast perhaps E
6	certainly target almost moves fast almost S OR slightly target almost moves slow almost N	certainly target certainly moves fast certainly S OR slightly target almost moves fast slightly E

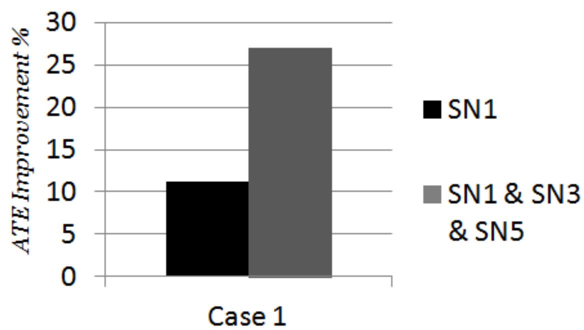

Figure 4.6: Tolerating gradual qualifier variation results: the high-quality and gradually changing soft data are retained and aggregated with other medium-quality soft data

Table 4.15: Tolerating gradual qualifier variations (Case 1): the SN2 soft data over time with occasional qualifier variations

Time stage	**Soft report**
1	certainly target perhaps moves regular almost W OR perhaps target certainly moves slow certainly S
2	certainly target perhaps moves regular perhaps W OR perhaps target certainly moves slow certainly S
3	certainly target perhaps moves regular almost E OR perhaps target certainly moves slow certainly S
4	certainly target almost moves regular almost E OR perhaps target certainly moves slow certainly S
5	certainly target almost moves regular perhaps SE OR perhaps target certainly moves slow certainly S
6	certainly target perhaps moves regular perhaps SE OR perhaps target certainly moves slow certainly S

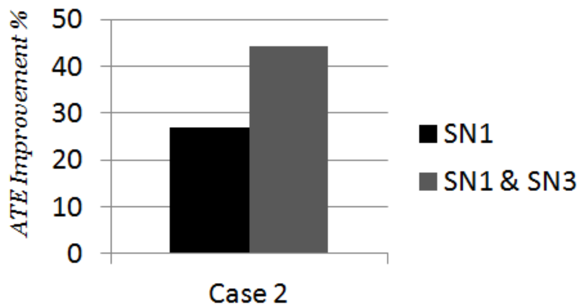

Figure 4.7: Tolerating gradual qualifier variation results: the high-quality and gradually changing soft data are retained and used solely to update the target motion model

Table 4.16: Tolerating gradual qualifier variations (Case 2): the SN1 low-quality soft data over time with abrupt variations

Time stage	**Soft report**
1	almost target perhaps moves regular perhaps N OR perhaps target certainly moves slow certainly W
2	almost target slightly moves regular certainly N OR slightly target slightly moves slow certainly W
3	almost target certainly moves slow almost N OR perhaps target slightly moves slow perhaps W
4	almost target perhaps moves fast certainly N OR slightly target certainly moves regular perhaps W
5	almost target perhaps moves fast slightly E OR perhaps target perhaps moves regular perhaps SE
6	almost target certainly moves fast perhaps E OR slightly target certainly moves fast certainly SE

Figure 4.7 shows the results obtained for two scenarios of the second case. One can see that in both scenarios the performance improvements are rather significant. This could be contributed to the fact that without tolerance to gradual change in qualifiers, the high-quality reports would not be used to update the target motion model. Furthermore, even the low quality reports could not be used either as they incur drastic change in qualifiers leading to a rapid drop in their credibility measure, which will in turn force trust in them

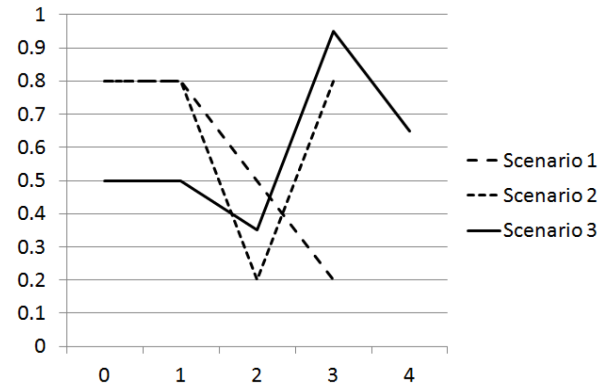

Figure 4.8: Target speed (Y axis) vs. the time stage (X axis): the speed is in m/s and each time stage lasts for $1500ms$

to go below almost level and thus make them unusable for the target motion model update. In effect, in both cases target motion model will not be updated unless the high-quality and gradually changing reports are retained and deployed through the tolerance mechanism.

4.5.5 Augmented Soft Data Experiments

The improved augmented soft data ontology allows for dealing with target dynamics involving acceleration/deceleration. The experiments described in this section evaluate the ability of our fusion system to deal with more challenging targets referred to as agile. Three increasingly more difficult test cases are performed. In each case two scenarios are considered. In one scenario, we deploy the soft data produced based on the original ontology where human agents cannot explicitly report on target acceleration/deceleration. In the other scenario, target dynamics are described more accurately using the soft data produced by the new augmented ontology. As our goal here is not to evaluate the distributed data aggregation scheme (see earlier experiments described in section x), we set all agents to produce the same report in all cases. Figure 4.8 shows how the target is tasked to move in each of the three cases. The motion direction is always set towards NE. Table 4.17 shows the soft reports provided using the original and the augmented ontology in each case. It is easy to notice that as target dynamics become more complex, the soft reports provided using the earlier ontology become more of an approximation and less representative of true target dynamics.

Figure 4.9 shows the performance improvements obtained using the new soft reports for each of the three test cases. As expected, the improvement achieved for the more challenging cases, i.e. second and third, is more noticeable. In addition, the improvement for the second case is slightly higher than the third, which may seem rather non-intuitive at first as target dynamics in the third case are more complex. However, a closer examination of the target dynamics in each of the last two cases along with the soft reports available to update the motion model in each case, would reveal that the tracking performance for the second case could actually be worse if one uses only the old ontology reports. Indeed, at both the second and third time stages the soft reports provided indicate something which is exactly to the contrary of how the target is about to behave. For instance, at time stage two the report implies that the target is going to be moving fast while in fact it will be reducing its speed at a high rate. A similar situation is repeated at the third time stage. Such a contrast between the reported target dynamics and its true behavior is not demonstrated in the third case although it involves more time stages and more frequent variations in the target dynamics.

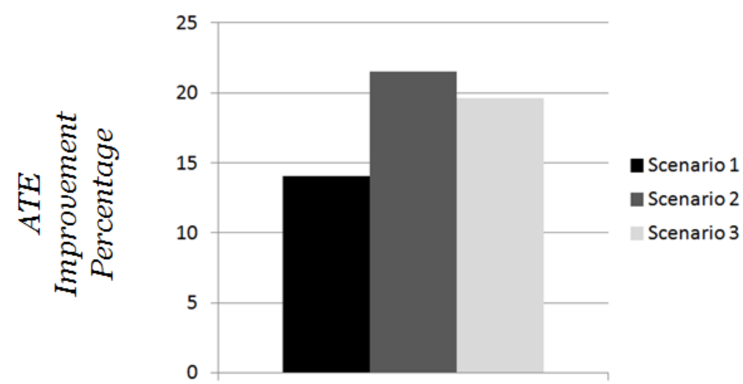

Figure 4.9: Augmented soft data experiment results

4.6 Chapter Summary

In this chapter, we addressed two important challenges pertinent to the development of soft/hard data fusion systems, namely, distributed aggregation of soft data, and human trust modeling and estimation. The experimental results obtained for a single-target tracking task demonstrated the potential of soft data to significantly improve the performance of tracking systems once fused properly with hard data.

We envision several avenues of future research for this work. The RS theory has been already shown to be capable of representing first order, second order and even composite rules [92]. This makes it interesting to study the potential of this theory to model human data (reports) in form of rule-based logical statements and incorporate them into the fusion process. A natural extension of the current single-target tracking system into multi-target tracking seems appealing. This would require solving the data association problem where both target measurements and states have the fuzzy D-S form discussed in this chapter. This has been the main objective of the human-centered multi-target tracking approach discussed in the next chapter.

Table 4.17: The soft data provided by human agents with and without the augmented ontology

	Case 1		**Case 2**		**Case 3**	
Time stage	**Old ontology**	**New ontology**	**Old ontology**	**New ontology**	**Old ontology**	**New ontology**
1	certainly target almost moves fast certainly NE	certainly target almost moves fast certainly NE	certainly target almost moves fast certainly NE	certainly target almost moves fast certainly NE	certainly target almost moves regular certainly NE	certainly target almost moves regular certainly NE
2	certainly target almost moves fast certainly NE	certainly target almost stopping regular NE	certainly target almost moves fast certainly NE	certainly target almost stopping fast NE	certainly target almost moves regular certainly NE	certainly target almost stopping slow NE
3	certainly target almost moves regular certainly NE	certainly target almost stopping regular NE	certainly target almost moves slow certainly NE	certainly target almost speeding fast NE	certainly target almost moves regular certainly NE	certainly target almost speeding fast NE
4					certainly target almost moves fast certainly NE	certainly target almost stopping regular NE

Chapter 5

Human-Centered RS Theoretic Multi-target Tracking and Classification

5.1 Introduction

A rather recent thread of research work in the data fusion community is the deployment of Random Set theory as an alternative mathematical framework to deal with target tracking problems, especially for the multi-target tracking applications RS theory is argued to be the natural choice [39, 186]. In chapter 4, we proposed a novel RS theoretic approach to enable fusion of soft human-generated data for single-target tracking applications. This chapter further develops our framework by proposing a data association algorithm applicable to soft data modeled using the RS theory, hence extending its applicability to multi-target tracking tasks. We also propose a robust RS theoretic Bayes classifier to enable classification of targets described using vague soft data. The obtained experimental results demonstrate the capacity of the RS theory as a viable approach to human-centered multi-target tracking and classification.

The rest of this chapter is organized as follows. The human-centered multi-target tracking approach, including the novel soft data association algorithm, is detailed in section 5.2. The preliminary multi-target tracking experiment results, demonstrating the efficiency of our soft data association algorithm, are discussed in section 5.3. The proposed RS theoretic apporach towards human-centered multi-target classification is described in section

Require: PDM, Atr, I, I_a
Ensure: $Cols$
while $I.size() > 0$ **do**
  $i = I.next()$
  {Find the index of PDM col with smallest value $dmin\_ind$ for the current target i}
  if $I_a[i] == -1$ {Agent has no opinion} **then**
    $Cols[i] = dmin\_ind$
  else
    if $|PDM[i][dmin\_ind] - PDM[i][Ia[i]]| \leq \frac{DIF\_TH}{Atr[i]}$ **then**
      $Cols[i] = dmin\_ind$
    else if $\frac{PDM[i][Ia[i]]}{Atr[i]} \leq DIF2\_TH$ **then**
      $Cols[i] = Ia[i]$
    else
      $Cols[i] = -1$ {Target unmatched!}
    end if
  end if
end while

Figure 5.1: *Comp_cols* method (local matching stage)

5.4, followed by a discussion of the obtained experimental results in section 5.5. The classification approach is then extended into the distributed paradigm through a distributed decision fusion algorithm presented in section 5.6. The related experimental results are provided in section 5.7. Finally, section 5.8 concludes this chapter and discusses several directions of future research for this work.

5.2 Human-Centered RS Theoretic Multi-target Tracking

As discussed in chapter 3, the soft data is modeled as a so-called fuzzy Dempster-Shafer state and fused using the Kalman Evidential Filter (KEF). A very recent work by Mahler develops conditional likelihood functions required to establish distance between the predicted target state and the given measurement for nonconventional cases including the fuzzy Dempster-Shafer states [187]. Leveraging these new developments, we propose a novel soft data measurement to track association algorithm tailored to the specific requirements of our original soft/hard data fusion framework considering a multi-target tracking

Require: PDM, Atr, I, I_a
Ensure: Asc_data_ind
while $I.size() > 0$ **do**
  $Cols = Comp\_cols(PDM, Atr, I, I_a)$
  $i = I.next();$
  if $Cols[i] == -1$ **then**
    $Asc\_data\_ind[i] = -1$ {mark target as unmatched}
    $I.remove(i)$
  else
    $Candid\_tr = i$ $Candid\_sd = Cols[i]$
    {Inter-target checking step}
    for $j \in I$**and**$j \neq i$ **do**
      if $Cols[j] == Candid\_sd$ **and** $PDM[j][Candid\_sd] < PDM[Candid\_tr][Candid\_sd]$ **then**
        $Candid\_tr = j$
      end if
    end for
    if $PDM[Candid\_tr][Candid\_sd] \leq DIFF2\_TH$ **then**
      $Asc\_data\_ind[Candid\_tr] = Candid\_sd$
      {Remove $Candid\_sd$ as a choice from PDM for all targets, e.g. by setting $PDM[][Candid\_sd] = LARGE\_NUM$}
      {Similarly, remove $Candid\_sd$ as a choice of the agents from I_a }
    else
      $Asc\_data\_ind[Candid\_tr] = -1$
    end if
    $I.remove(Candid\_tr)$
  end if
end while

Figure 5.2: Soft Data Association Algorithm

task. Once the data association is performed, we apply a bank of KEFs to update tracks for all targets. We performed a series of tests to evaluate the efficiency of our algorithm in terms of average tracking accuracy for several targets. Comparing with the baseline case of relying merely on human agent opinions to resolve the association problem, our results demonstrate the ability of our soft data association algorithm (SDAA) to significantly improve the tracking performance. To the best of our knowledge, this is the first work to explore a RS theoretic approach towards multi-target tracking using nonconventional human-generated data.

5.2.1 Soft Data Association Algorithm

This section presents the core enabling technology for our human-centered multi-target tracking system, i.e. a novel soft data association algorithm. Our SDAA can be considered as an augmented nearest neighbor (NN) association algorithm as it relies mainly on the pairwise distance between measurements and target tracks and assigns tracks to the measurement with the minimum distance. However in contrast to the basic NN algorithm applicable to hard data, our SDAA also takes into account the human agent opinions. Moreover, to further enhance the robustness of associations an inter-target checking (ITC) procedure is performed to resolve any potential conflicts among targets before finalizing the associations.

Before proceeding with describing our SDAA, we should define some terminology used in the pseudocode representation of this algorithm. The PDM is a $N_t \times N_m$ matrix storing the pairwise distance measure between given N_t number of targets and N_m number of measurements. The PDM computation equations are discussed later at the end of this section. The Atr is an array storing the (relative) agent trust ratios for each of the targets $i \in \{1, \dots N_t\}$ reported by the agent and is defined as

$$Atr[i] = \sum_{e=1}^{3} \frac{e\_qualifier_e}{5-e} \tag{5.1}$$

The $Atr[i] = 1$ if all the three possible expressions comprising the given target report have the highest level of certainty available, i.e $e\_qualifier_1 = \mathit{certainly}$, $e\_qualifier_2 = \mathit{almost}$ and so on (See also Figure 3.3). This parameter captures the human agent's confidence level in the expressed report regarding a target. Please note the above equation assumes the $e\_qualifier$ to be coded as $\mathit{certainly} = 4$, $\mathit{almost} = 3$, $\mathit{perhaps} = 2$, and $\mathit{slightly} = 1$. The I represents the set of all given target indices to be matched with measurements and is initialized to $\{1, \dots N_t\}$. The I_a represents the set of target indices as expressed by the term $target\_ID$ in the soft data ontology, i.e. I_a specifies the target track to measurement associations according to the human agent opinion.

In summary, there are two stages involved in the proposed SDAA, the local association stage and the global association stage. In the local stage, which is implemented as procedure $Comp\_cols()$ (See Figure 5.1), the index of the best candidate measurement for each of the target tracks is computed and returned as array $Cols$. In the local matching process the first priority is given to the closest measurement, the one with smallest PDM, which is also not too far from the one specified by the agent (See Figure 5.1 lines 7-8). The allowed difference between the measurement with minimal distance and the agent's choice is controlled by the threshold $DIFF\_TH$, which is penalized (enlarged) for $Atr < 1$, i.e. agent report not completely certain. The second priority is given to the measurement specified as the target match by the agent as long as it's distance, which is similarly penalized by Atr, is below the threshold level $DIFF2\_TH$ (See Figure 5.1 lines 9-10). In case none of the aforementioned conditions hold, target is marked as unmatched.

In the global matching stage, shown as a pseudocode in Figure 5.2, an inter-target checking (ITC) procedure is performed to check for potential conflicts among targets. Targets i and j are considered to be in conflict if they are both assigned to the same measurement in the local matching stage (See Figure 5.2 lines 8-13). The conflict is resolved by assigning the problematic measurement to the target with the smallest distance (closest match). Moreover, in order to avoid accepting a poor measurement as a match the chosen measurement is rejected if its pairwise distance to the given target is too large (See Figure 5.2 line 14). Once a measurement is assigned to a target it is no longer considered as a potential match or choice of the human agents (See Figure 5.2 line 15). This two-stage procedure is repeated until all targets are either matched or marked as unmatched. The indices of the final measurement to target associations are stored in the array Asc_data_ind.

The measurement to track distance metric used in our SDAA is developed based on the premises originally introduced by Mahler [39]. More recently he formally describes a framework to establish measurement to track association for nonconventional measurements formulated using the RS theory [187]. As discussed earlier, in our framework both soft measurement o and (predicted) target state $\mu_{t+1|t}$ are represented as fuzzy D-S states. Mathematically speaking, let $\mu_{t+1|t}$ to be comprised of focal sets of the form

$$f_k(x) = \hat{N}_{D_i}(x - x_i) \tag{5.2}$$

with the corresponding weight μ_k and similarly the measurement o to have focal sets as shown below

$$g_l(z) = \hat{N}_{C_l}(z - z_j) \tag{5.3}$$

and weights denoted as o_l.

The $PDM[i][j]$ represents the distance between the predicted state of target i and soft measurement j and is computed as below

$$PDM[i][j] = \frac{1}{\ell(o_j|\mu_i)} \tag{5.4}$$

where [39]

$$\ell(o_i|\mu_j) = \sum_{k=1}^{e}\sum_{l=1}^{d} w_{k,l}.o_l.\mu_k.\hat{N}_{C_l+HD_kH^T}(Hx_k - z_l) \tag{5.5}$$

and

$$w_{k,l} = \sqrt{\frac{det2\pi C_l}{det2\pi(C_l + HD_kH^T)}}.\frac{\sqrt{det2\pi D_k}}{\sum_{n=1}^{e}\mu_n.\sqrt{det2\pi D_n}} \tag{5.6}$$

5.3 Multi-target Tracking Experiments

As discussed earlier, there are some preliminary works aiming at creating standard datasets to enable evaluation of soft/hard data fusion systems [150, 151], there are no publicly available datasets for this purpose at this point. Therefore, we have conducted a series of multi-target tracking experiments ourselves, which are designed to highlight and evaluate several performance characteristics of the proposed soft data association algorithm.

5.3.1 General Settings

The performance metric used for the experiments is the average position tracking error (ATE), i.e. the cumulative average of the Euclidean distance between the known and the estimated target position (in meters). Each of the experimental scenarios is repeated five times in an attempt to obtain a more realistic measure of the system performance. The simulation time resolution is 1ms. The SDA supplies soft data to the DPA every

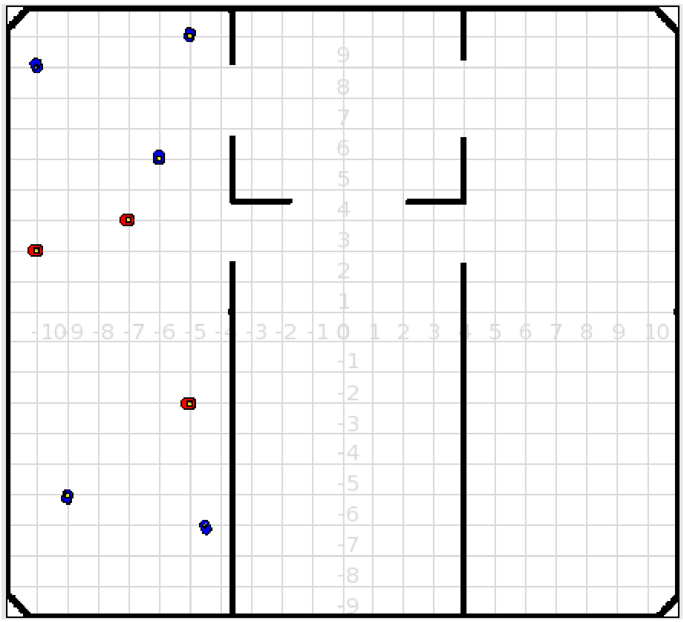

Figure 5.3: A snapshot of the Player/Stage simulation arena setup for the second experiment.

1000ms, i.e. $\Delta t_{soft} = 1000ms$. The parameter setting used to interpret soft data is shown in Table 4.1. All the threshold values are determined experimentally. The multi-target experiments are conducted with three targets (mobile robots) at this time, namely, $T1$, $T2$, and $T3$. However, due to the distributed nature of computations, it should be easy to scale up our experiments to a larger number of targets. As our main objective is to assess the soft data association algorithm, we rely merely on soft data for tracking in our experiments. Nonetheless, since the proposed framework is capable of processing both soft and hard data, one can adopt and use an appropriate hard data association algorithm (from the data fusion literature) and perform soft/hard multi-target tracking. Figure 5.3 shows a snapshot of the Stage simulation arena illustrating the positioning of human agents (blue circles) and initial position of targets (red circles) as setup for our second experiment. Furthermore, Figure 4.4 shows the connectivity map of the five sensor nodes (human agents) deployed in our experiments.

5.3.2 Removing Poor Soft Data

This scenario is designed in order to evaluate the effectiveness of the soft data association algorithm in excluding poor data from the association process and thus avoiding performance degradation. Targets are assumed to have rather distinct dynamics and are tasked to move accordingly (See Figure 5.4(a)). The soft reports provided by the human agents are organized such that for $T1$ and $T3$ all agents produce high quality reports whereas for $T2$ the (minimum) majority of agents (3 out of 5, i.e. $SN1$, $SN3$, and $SN5$ in this test) provide low quality data while the rest provide medium quality data (see Table 5.1). The low quality data must be identified and removed by the data association algorithm while the medium quality must be preserved. i.e. associated with $T2$. The baseline algorithm involves no data association algorithm, i.e. relies merely on the human agent opinion expressed as the $target\_ID$ in soft data ontology (see Figure 3.3).

Figure 5.5 shows the ATE for all three targets using both the baseline and our soft data association algorithms. Comparing the two algorithms, one could see for $T1$ and $T3$, for which the human soft data are high quality and properly associated, the results are almost identical, whereas for $T2$ the performance is significantly improved once the poor data provided by $SN1$, $SN3$, and $SN5$ are removed successfully by the SDAA. Please note the reason for tracking results for $T2$ using the SDAA being worse than the other two targets is that the available soft data (not removed) for this target provided by $SN2$ and $SN4$ are of medium quality rather than high quality (See Table 5.1). Indeed, these $T2$ reports are set to be of medium quality intentionally to ensure that the SDAA algorithm is able to associate them to the $T2$ while eliminating other low quality reports.

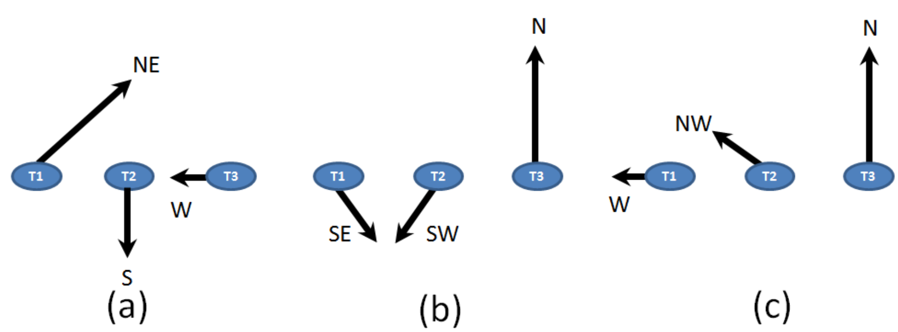

Figure 5.4: Target dynamics: (a) Removing poor data scenario (b) Resolving target confusion scenario (c) The need for ITC scenario (Hint: the longer the arrow, the faster the target speed).

Table 5.1: The soft data provided by all five agents for removing poor soft data scenario

SN1	**SN2**	**SN3**	**SN4**	**SN5**
certainly target1 perhaps moves fast perhaps NE OR perhaps target1 certainly moves regular certainly N AND certainly target2 perhaps moves fast perhaps W OR perhaps target2 certainly moves slow certainly NE AND certainly target3 perhaps moves slow perhaps W OR perhaps target3 certainly moves regular certainly SW	certainly target1 perhaps moves fast perhaps NE OR perhaps target1 certainly moves regular certainly N AND certainly target2 perhaps moves fast perhaps S OR perhaps target2 certainly moves regular certainly SE AND certainly target3 perhaps moves slow perhaps W OR perhaps target3 certainly moves regular certainly SW	certainly target1 perhaps moves fast perhaps NE OR perhaps target1 certainly moves regular certainly N AND certainly target2 perhaps moves fast perhaps W OR perhaps target2 certainly moves slow certainly NE AND certainly target3 perhaps moves slow perhaps W OR perhaps target3 certainly moves regular certainly SW	certainly target1 perhaps moves fast perhaps NE OR perhaps target1 certainly moves regular certainly N AND certainly target2 perhaps moves fast perhaps S OR perhaps target2 certainly moves regular certainly SE AND certainly target3 perhaps moves slow perhaps W OR perhaps target3 certainly moves regular certainly SW	certainly target1 perhaps moves fast perhaps NE OR perhaps target1 certainly moves regular certainly N AND certainly target2 perhaps moves fast perhaps W OR perhaps target2 certainly moves slow certainly NE AND certainly target3 perhaps moves slow perhaps W OR perhaps target3 certainly moves regular certainly SW

5.3.3 Resolving Target Confusion

This test scenario evaluates the ability of the proposed SDAA to resolve potential confusion of targets (by human observers). This could be caused for instance by targets with rather similar dynamics passing by each other at a close vicinity. The baseline algorithm deployed in this test is the same as the one described for the previous test scenario, i.e. based only on the human agent opinions. For this test the initial positioning (see Figure 5.3) and dynamics (see Figure 9(b)) for $T1$ and $T2$ are set in such a way that their trajectories get close causing confusion of the human observers, i.e. $T1$ being reported as $T2$ and vice versa by a (minimum) majority of the agents ($SN1$, $SN3$, and $SN5$) in this test. The $T3$ on

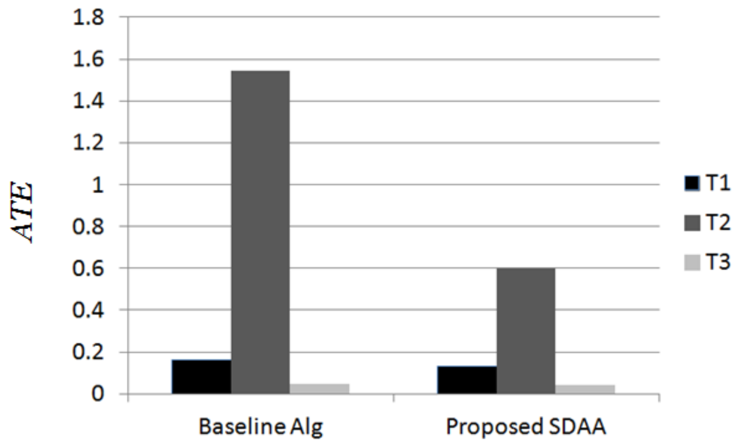

Figure 5.5: Removing poor soft data: the tracking performance using the baseline algorithm and the proposed soft data association algorithm.

the other hand traverses through a distinct trajectory and thus is assumed to be followed easily by the human observers. Table 5.2 illustrates the soft data provided by $SN1$ and $SN2$ for this test scenario over several time steps. The data provided by the $SN3$ and $SN5$ are the same as $SN1$. Similarly, $SN4$ provides a data identical to that of $SN2$ and thus is not repeated here. As shown all soft data are of high quality and correctly associated at the first two time steps. However, $SN1$ confuses $T1$ with $T2$ for the time steps 3 to 5, which if not corrected for by the SDAA leads to performance degradation as shown in Figure 5.6 diagrams. We have also included the tracking results obtained using the SDAA algorithm where the agent trust modeling scheme is disabled, i.e. always $Atr = 1$ (See Fig 5.1 and (7)). As shown the SDAA is unable to resolve the target confusion in this case. This is due to the fact that with $Atr = 1$ instead of its proper value in this experiment $Atr = 0.83$, the $DIFF\_TH$ threshold value is not penalized enough to allow for the proper measurement, the one with the smallest PDM distance, to be associated to the confused targets. This experiment illustrates the advantage of the proposed soft data association algorithm in dealing with challenging tracking scenarios involving close targets. Moreover, it signifies the importance of the agent trust modeling scheme embedded into the local matching stage of the SDAA.

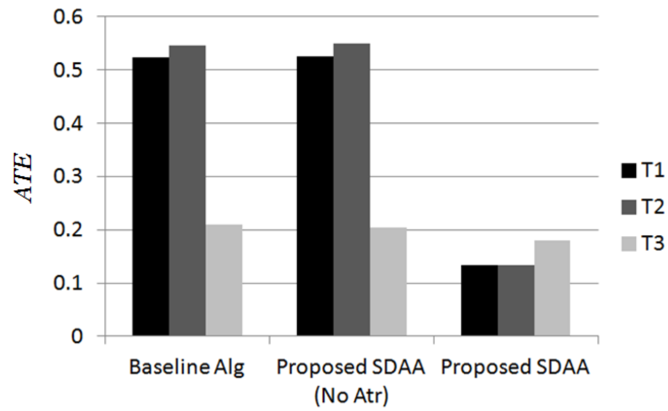

Figure 5.6: Resolving target confusion: the tracking performance using the baseline algorithm and the proposed soft data association algorithm with and without the Atr modeling.

Table 5.2: The exemplary soft data provided by $SN1$ and $SN2$ for resolving target confusion scenario

Time step	**SN1**	**SN2**
1-2	certainly target1 perhaps moves regular perhaps SE OR perhaps target1 certainly moves regular certainly E AND certainly target2 perhaps moves regular perhaps SW OR perhaps target2 certainly moves regular certainly S AND certainly target3 perhaps moves fast perhaps N OR perhaps target3 certainly moves regular certainly N	certainly target1 perhaps moves regular perhaps SE OR perhaps target1 certainly moves regular certainly E AND certainly target2 perhaps moves regular perhaps SW OR perhaps target2 certainly moves regular certainly S AND certainly target3 perhaps moves fast perhaps N OR perhaps target3 certainly moves regular certainly N
3-5	same as steps 1-2 except now target1 and target2 are swapped	same as steps 1-2
6	same as steps 1-2	same as steps 1-2

Table 5.3: The two soft data categories used for the need for ITC scenario

Category 1 data	**Category 2 data**
certainly target2 slightly moves fast slightly N AND certainly target3 slightly moves regular slightly NW AND certainly target1 slightly moves regular slightly S	certainly target1 perhaps moves slow perhaps W OR perhaps target1 certainly moves regular certainly W AND certainly target2 perhaps moves regular perhaps NW OR perhaps target2 certainly moves regular certainly W AND certainly target3 slightly moves regular slightly NE

5.3.4 The Need for ITC

There are situations where it is not satisfactory to rely merely on the distance between the predicted target state and the measurement, as well as the human observer opinion, to determine the association between target tracks and measurements. In other words, one also has to consider potential inter target conflicts and resolve them before finalizing the associations and that is the objective of the ITC stage of the proposed SDAA. The test scenario discussed in this section aims at simulating a tracking scenario where the requirement for the ITC stage is highlighted. In order to do so we arrange for the target dynamics to be rather similar (see Figure 5.4(c)). Furthermore, the soft data provided by the agents is set to belong into two categories (See Table 5.3). The first category of data are provided by the (minimum) majority of the agents, namely, $SN1$, $SN3$, and $SN5$, and require the ITC stage in order to be associated more accurately to their corresponding targets. The second category of soft data however are set to require no ITC, i.e. the associations are identical with or without the ITC procedure.

The baseline algorithm (no ITC) here performs in a sequential way, i.e. find and assign the best measurement for $T1$, then $T2$ and so on without checking for inter target conflicts and performing reassignment/reiteration steps involved in the ITC procedure. Table 5.4 shows the measurement to track associations produced by the proposed SDAA, with and without the ITC stage, for the majority of the agents providing the first category of soft data and the rest of the agents providing the second category soft data. The obtained experimental results are depicted in Figure 5.7. As shown using the ITC stage the tracking performance is highly improved especially for $T2$ and $T3$ as a result of far more accurate

Table 5.4: The need for ITC: measurement to track associations produced by the proposed SDAA

	Category 1 Data		Category 2 Data	
	No ITC	ITC	No ITC	ITC
$T1$	**report2**	**Unmatched**	**report1**	**report1**
$T2$	**report1**	**report2**	**report2**	**report2**
$T3$	**Unmatched**	**report1**	**report3**	**report3**

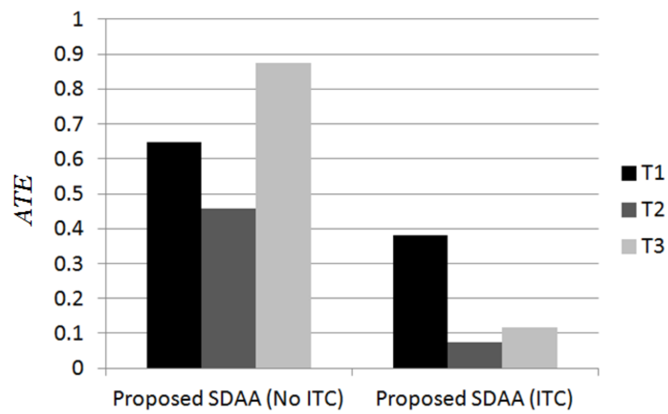

Figure 5.7: The need for ITC: the tracking performance using the proposed soft data association algorithm without and with ITC enabled.

associations being established (See Table 5.4). Please note the improvement is not as significant for $T1$. This is contributed to the fact that using ITC for the first category soft data this target is left unmatched, in contrast to the high quality matches for $T2$ and $T3$, simply because there is no such high quality matching measurement available. Nonetheless, the overall tracking performance for all targets is enhanced substantially.

5.4 Human-Centered RS Theoretic Multi-target Classification

In this section we describe an approach to further extend the proposed RS theoretic multi-target system by deploying the target state estimates over time to classify targets. The targets themselves are vaguely described in an unconventioal fashion by the soft human-supplied data. Considering that the proposed RS theoretic framework in chapter 3, deploys soft vague data to update the target state estimates over time, we assume the obtained estimation to possess/express vagueness type of data imperfection as well. In his recent book, Mahler [39] discusses fusion of an abstract type of data called AGA (ambigiously generated ambigious data) that fits our fusion scenario. In the following, we first present the developed soft data ontology used to describe different target classes, namely, spy, passenger, and threat. Next, we discuss how these vaguely described target classes along with the vague target state estimates could be deployed in a RS theoretic Bayes classifier to achieve robust target classification.

5.4.1 Modeling Target Classes Using Soft Data

Figure 5.8 shows the syntax considered for the soft data describing various target classes. In the current implementation targets are distinguished using two characteristics (actions), namely, their velocity and maneuver. However, if necessary the proposed approach could be easily extended to accomodate a larger number of target characteristics. Accordingly, in order model the vague target data generation process (aka target signature), we deploy a 2D Gaussian fuzzy membership function $\eta_{T_i}(z)$ for target T_i as

$$\eta_{T_i}(z) = N_{\sigma_{T_i}}(z - z_{T_i}). \tag{5.7}$$

The mean $z_{T_i} = [Vel_{T_i}\ \Delta\theta_{T_i}]^T$ is comprised of the nominal velocity Vel_{T_i} and maneuver $\Delta\theta_{T_i}$ of target class T_i chosen according to the target velocity and meneuver size of the provided description statement (see Figure 5.8).

In addition, the velocity and maneuver qualifier terms of the target class T_i description ($v\_qualifier$ and $m\_qualifier$, respecrtively) are deployed to determine the corresponding covariance matrix σ_{T_i}, assuming Vel_{T_i} and $\Delta\theta_{T_i}$ to be independent, as shown below (see also Table 5.5).

$$\sigma_{T_i} = \begin{bmatrix} var(vel_{T_i}) & 0 \\ 0 & var(\Delta\theta_{T_i}) \end{bmatrix} \tag{5.8}$$

In order to obtain the data pertinent to target characteristics of interest disucssed above, the original target state representation model $x = [X\ Y]^T$ discussed in chapter 3 is augmented as $x = [X\ Y\ \dot{X}\ \dot{Y}]$. The augmented target state includes an estimate of target

Description =< target_class >< *action1* >< *v_qualifier* >< *velocity* >< *action2* >< m_qualifier >< *mnv_size* >

< *_qualifier > ∈ {*slightly, perhaps, almost, certainly*}

< target_class > ∈ {*spy, passenger, threat*}

< *action1* > ∈ {*moves*}

< *action2* > ∈ {*maneuvers*}

< *velocity* > ∈ {*slow, regular, fast*}

< *mnv_size* > ∈ {*small, medium,* large}

Figure 5.8: Syntax considered for the soft target description data.

Table 5.5: Soft target class data interpretation parameter setting

Reported data	**Corresponding parameter**	**Corresponding value(s)**
moves slow/regular/fast	Vel_{T_i}	0.2/0.5/0.8 (m/s)
maneuvers small/medium/large	$\Delta\theta_{T_i}$	10/15/25 $degrees$
(v_qualifier) certainly/almost/perhaps/slightly	$var(vel_{T_i})$	0.2/0.4/0.6/0.8 (m/s)
(m_qualifier) certainly/almost/perhaps/slightly	$var(\Delta\theta_{T_i})$	5/10/15/20 $degrees$

velocity along the x and y axes ($\dot{X}$ and $\dot{Y}$, respectively) and thus the associated motion and measurement models are also modified accordingly. Using the new target state, the desired target characteristics, i.e. velocity Vel and maneuver size $\Delta\theta$ are calculated over time as shown in the equations below.

$$Vel_t = \sqrt{\dot{X}_t^2 + \dot{Y}_t^2} \tag{5.9}$$

$$\theta_t = \arctan \frac{\dot{Y}_t}{\dot{X}_t} \tag{5.10}$$

$$\Delta\theta_t = |\theta_t - \theta_{t-1}| \tag{5.11}$$

Once computed, the vague target characteristic measurement $z_t = [Vel_t \ \Delta\theta_t]^T$ is modeled as a 2D fuzzy Gaussian membership function $g(z_t)$ defined below

$$g(z_t) = N_{\sigma_o}(z - z_t) \tag{5.12}$$

where the covariance matrix σ_o is denoted as

$$\sigma_o = \begin{bmatrix} var(vel_t) & 0 \\ 0 & var(\Delta\theta_t) \end{bmatrix}. \tag{5.13}$$

5.4.2 Robust RS Theoretic Bayes Classiifer

Using a RS theoretic representation of vague measurement model $g(z_t)$ and the vague target data generation model $\eta_{T_i}(z)$, Mahler [39] defines the generalized likelihood function for such AGA data as below

$$Pr(\Theta_t|T_i) = Pr(\Theta_t \cap \Sigma_{T_i} \neq \emptyset) \tag{5.14}$$

Where Θ_t and Σ_{T_i} denote the random set representations of the vague measurement model and vague target data generation model, respectively. Accordingly, it is shown that for fuzzy AGA data the above equation takes the following form

$$Pr(g(z)|\eta_{T_i}(z)) = sup\ min_z\{g(z), \eta_{T_i}(z)\} \tag{5.15}$$

Which for the specific case of $g(z)$ and $\eta_{T_i}(z)$ being Gaussians, is shown to reduce to the following

$$Pr(g(z_t)|\eta_{T_i}(z)) = exp(-\frac{(z_t - z_{T_i})^2}{2(\sigma_o + \sigma_{T_i})^2}) \tag{5.16}$$

Deploying this likelihood function within the Bayesian framework, we develop a RS theoretic Bayes filter as

$$Pr(T_i|g(z_t), \ldots, g(z_0)) = \frac{Pr(g(z_t)|T_i) \ldots Pr(g(z_t)|T_i))}{\sum_T Pr(g(z_t|T) \ldots Pr(g(z_t)|T))} \tag{5.17}$$

where target measurements are assumed to be independent and the target class priors are considered to be uniform. For classificatin purposes, the maximum likelihood principle is applied to deduce the target class as shown below.

$$\hat{T} = Cls_g = argmax_{T_i} Pr(T_i | g(z_t), \ldots, g(z_0)) \tag{5.18}$$

It can be noted that the aforementioned target classifier is designed such that the more the number of target readings, the better the classification performance and reliability would become through temporal "fusion" of new target data. We have defined a classification confidence measure Cf shown below to reflect upon this gradual improvement in the classification reliability in our experiments.

$$Cf(\hat{T}) = \frac{Pr(\hat{T}|g)}{\sum_{T_i} Pr(T_i|g)} \tag{5.19}$$

Indeed, the experimental results presented in the subsequent section support this notion.

5.5 Multi-target Classification Experiments

In this section we first discuss the modifications made to the multi-agent organization of our system (see Figure 3.5) in order to implement the vague target classifier and then present the results of the preliminary experiments performed to evaluate the performance of the developed target classifier.

5.5.1 Augmented Multi-agent Organization

As shown in Figure 5.9, the target classification algorithm is performed by the classifier agent (CA), which operates using the target state estimates provided by the DFA and vague target descriptions supplied by the DPA. Please note similar to the soft target data, the vague target description data is collected by the HIA, formatted by the SDA (according to the syntax depicted in Figure 5.8) and is preprocessed/interpretted at the DPA (using the semantics shown in Table 5.5).

5.5.2 Preliminary Experiments

The experiments in this section are comprised of two test scenarios whereby the soft data describing target classes express a low and high level of vagueness, respectively. The

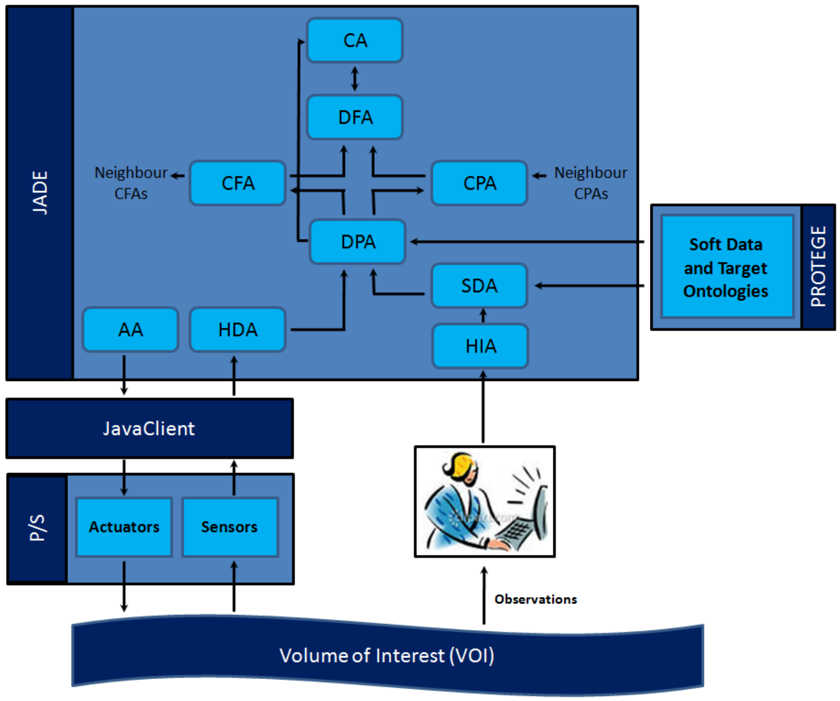

Figure 5.9: The augmented multi-agent organization of developed fusion system.

objective is first to assess the ability of the proposed classifier to deploy vague target description data and achieve robust classification and second to observe how higher levels of vagueness in target description affects performance of the classifier. Following is the common settings used in both test scenarios. Both cases involve five targets $T1 \ldots T5$ and three target classes, namely, *spy*, *passenger*, and *threat* (presuming a surveillance application). $T1$, $T2$, and $T5$ are of passenger type while $T3$ and $T4$ belong to the threat and spy classes, respectively. Tables 5.6 and 5.7 present the vague target description data deployed in the experiments. Comparing the target descriptions, one can note that in case of high vagueness, the statement provided by the human agent has less certain qualifiers corresponding to highly overlapping random set representations of the vague target areas. Consequently, one would expect the target classification task in this case to be more challenging due to such overlapping areas. This is confirmed by the experimental results presented later on in this section.

Figures 5.10 and 5.11 demonstrate the target data, namely, target velocity and maneuver size, respectively. This data is deployed in both experiments, i.e. fed to the CA over time. As shown ten readings are collected over time for each of the targets. We have

Table 5.6: Soft data description of target classes with low vagueness

Target Class	**Description Data**
Spy	target1 moves certainly slow maneuvers certainly small
Passenger	target2 moves certainly regular maneuvers certainly medium
Threat	target3 moves certainly fast maneuvers probably large

Table 5.7: Soft data description of target classes with high vagueness

Target Class	**Description Data**
Spy	target1 moves certainly slow maneuvers certainly small
Passenger	target2 moves almost regular maneuvers almost medium
Threat	target3 moves probably fast maneuvers slightly large

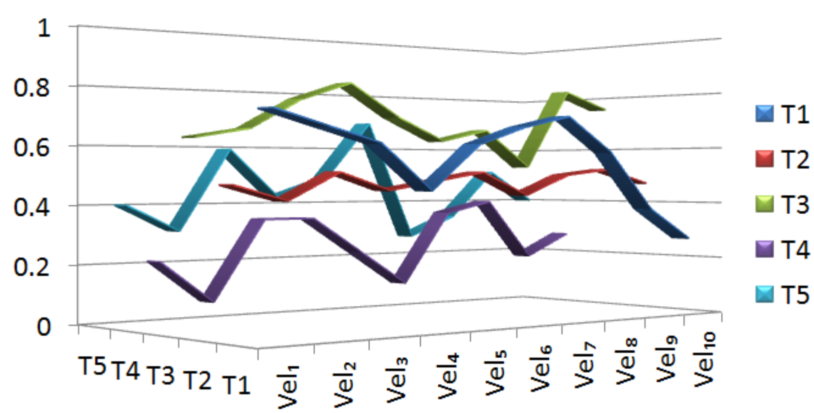

Figure 5.10: The target velocity estimate provided over time (in m/s).

arranged the target data such that readings for the $T1$ and $T3$ are contaminated with a higher level of noise while the rest of targets data are less noisy , i.e. closer to what is expected from the specific target class. Our objective is to evaluate the robustness of the proposed classifier with respect to noisy target data.

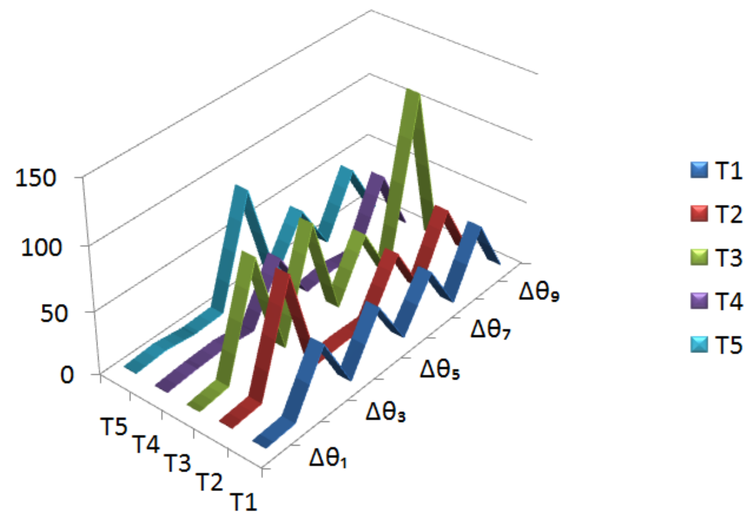

Figure 5.11: The target maneuver estimate provided over time (in degrees).

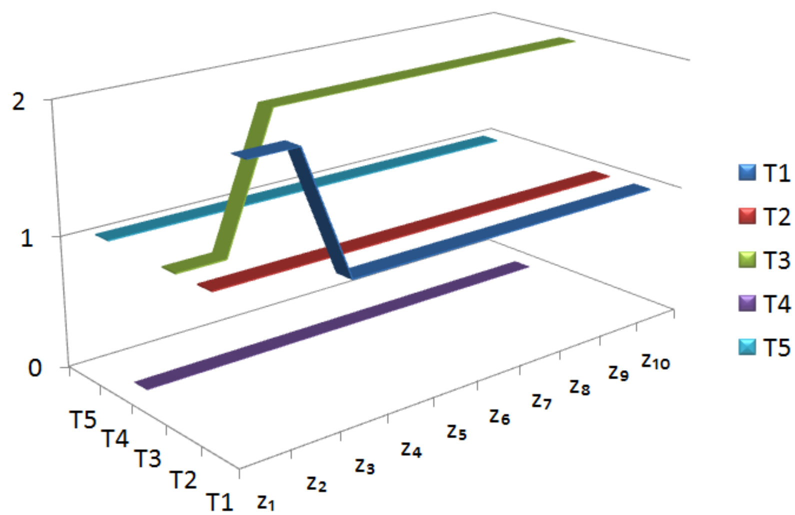

Figure 5.12: Low vagueness target description experiment: the estimated target class over time (class 0: spy, class 1: passenger, class 2: threat).

Figures 5.12 and 5.13 depict the obtained experimental results, in terms of estimated target class and classification confidence, respectively, for the test involving the low vagueness target descriptions. As shown the target class estimation is correct for all targets except $T1$ and $T3$, which could be contributed to the high level of noise for these targets data. Nonetheless, after collecting only three measurements the classifier is able to deduce the correct target class for both $T1$ and $T3$ demonstrating its robustness to noise. Furthermore, as expected the classification confidence improves rather steadily over time with more target data being supplied to the classifier. Similarly, the obtained experimental results for the test involving the high vagueness target descriptions are presented in Figures 5.14 and 5.15. Once again, the estimated target classes are initially invalid for both $T1$ and $T3$ and are corrected after receiving the fourth and third measurements, respectively. As expected in this case, it takes a larger number of target readings for the classifier to deduce the correct target classes for the targets with noisy data. The larger number of measurements, which is still reasonably small, is justified by the fact that the highly overlapping target areas make the classification task more difficult, i.e. less confident. This is also reflected in the estimated classification confidence results presented in Figure 5.15. Comparing this diagram with its counterpart in the former experiment (Figure 5.13), one can see that although in both cases the trend is increasing over time, the maximum confidence level achieved in the latter case is lower.

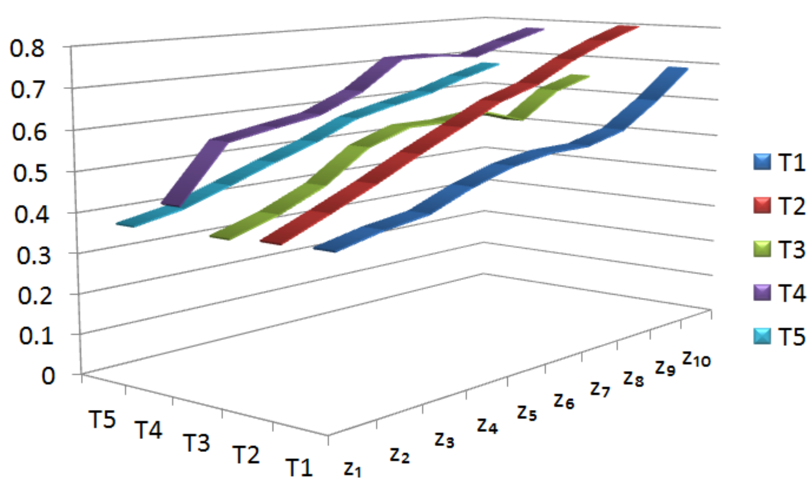

Figure 5.13: Low vagueness target description experiment: the estimated classification confidence over time.

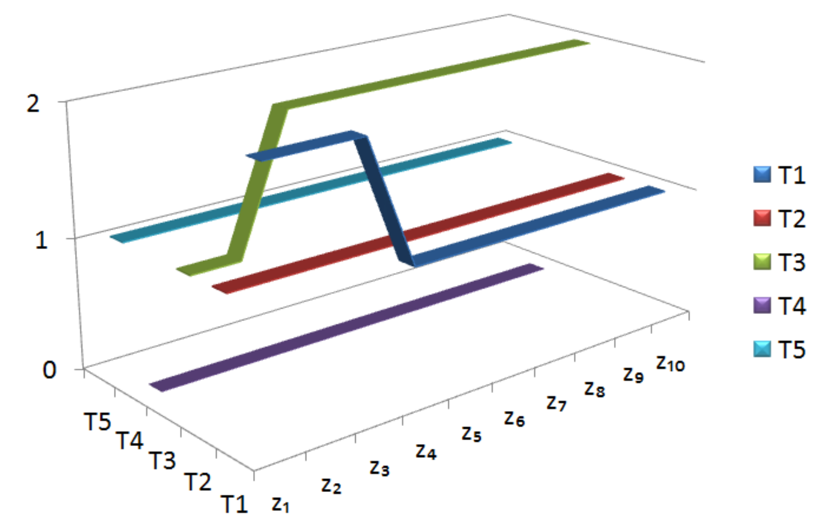

Figure 5.14: High vagueness target description experiment: the estimated target class over time (class 0: spy, class 1: passenger, class 2: threat).

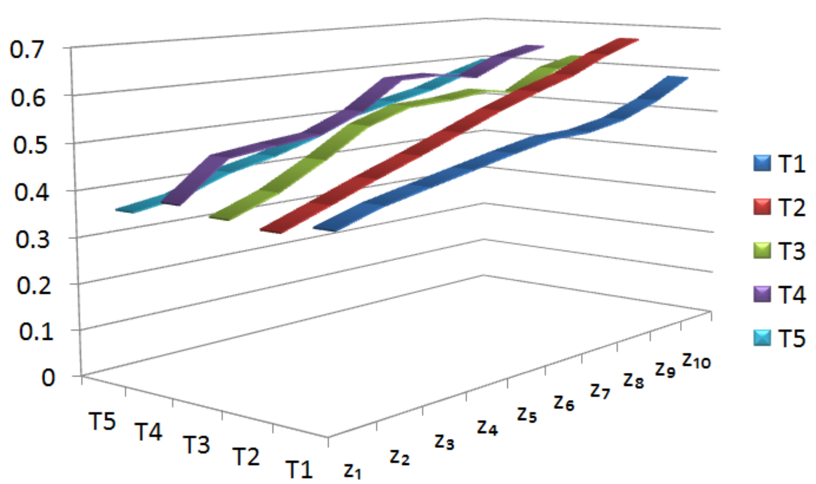

Figure 5.15: High vagueness target description experiment: the estimated classification confidence over time.

5.6 Distributed Multi-target Classification

The distributed decision fusion approach underlying the proposed distributed target classification method is described in this section. The aim is to enhance the classification performance through fusion of local classifier results while the distributed nature of computations leads to improved scalability. As discussed earlier, the proposed approach differs from previous work on distributed target classification in being truly distributed, i.e. no need for a fusion center, and also not relying only on binary local decisions.

In order to achieve the distributed fusion of local results, we propose to compute the weighted average of the local likelihood measures $Pr^i(T_c|Z_t)$ (see 5.17) for each of the detected targets i and given target classes T_c (Please note Z_t denotes the set of target readings gathered by time t). This is accomplished using the distributed aggregation scheme called consensus propagation algorithm and thus is truly distributed, i.e. requires merely local information exchange among sensor nodes. Assuming each of the local likelihood measures as a parameter, there are a total of $N^t \times N^c$ parameters to be aggregated, where N_t and N_c represent the number of detected targets and the given number of target classes, respectively.

5.6.1 Distributed Decision Fusion Algorithm

The CP is indeed a special case of Gaussian belief propagation that can be used to compute global average in a distributed manner. Similar to the popular consensus filter [184] (CF), CP is an iterative algorithm. However CP differs from CF in that data (messages) sent to each of the neighbor sensor nodes are specific to that node, whereas CF broadcasts the same message to all neighbors at each iteration. More importantly, using CP the message sent to each neighbor sensor node contains not only the latest estimate of the desired parameter U but also the number of sensor nodes contributing to that estimation process M_U. Mathematically, for every iteration k at node i, the incoming data $[U^{ji,k-1}, M_U^{ji,k-1}]$ previously received and stored from all neighbor sensor nodes $j \in N_i$ are used (along with the local estimate U_i^{local} if applicable) to update the global estimate $U_i^{global,k}$ of the desired parameter as

$$U_i^{global,k} = \frac{\rho_i + \sum_{j \in N_i} U^{ji,k-1} M_U^{ji,k-1}}{\kappa_i + \sum_{j \in N_i} M_U^{ji,k-1}} \tag{5.20}$$

The stored incoming data are also used to compute the (neighbor-specific) outgoing messages $[U^{ij,k}, M_U^{ij,k}]$ for each of the neighbor nodes j as shown below.

$$U^{ij,k} = \frac{\rho_i + \sum_{l \in N_i/\ j} U^{li,k-1} M_U^{li,k-1}}{1 + \sum_{l \in N_i/\ j} M_U^{li,k-1}} \tag{5.21}$$

$$M_U^{ij,k} = \frac{\kappa_i + \sum_{l \in N_i/\ j} M_U^{li,k-1}}{1 + \frac{1}{\beta}(\kappa_i + \sum_{l \in N_i/\ j} M_U^{li,k-1})} \tag{5.22}$$

$$\rho_i \ = \begin{cases} U_i^{local} & node\ i\ is\ contributing \\ 0 & otherwise \end{cases} \tag{5.23}$$

$$\kappa_i \ = \begin{cases} Cf_i & node\ i\ is\ contributing \\ 0 & otherwise \end{cases} \tag{5.24}$$

The Cf_i represents the confidence level assumed by node i regarding its soft data and indeed determines the corresponding weight for the likelihood measures provided by node i. The confidence measure Cf could be provided *a priori*, based on the historical performance for instance, or estimated on-the-fly. In either case, weighted averaging according to this measure allows more reliable data to have more significant impact on the final aggregated likelihoods, hence improving the classification performance.

Once aggregated, the likelihood measures $Pr^i_{aggr}(T_c|Z_t)$ are deployed along with the maximum likelihood principle to infer the target class as shown below.

$$\hat{T}_{aggr} = argmax_{T_i} Pr^i_{aggr}(T_i|g(z_t), \ldots, g(z_0)) \tag{5.25}$$

5.7 Distributed Multi-target Classification Experiments

The test scenarios discussed in this section were designed and conducted in order to illustrate the classification performance enhancements achieved through the distributed decision fusion scheme. In terms of implementation, the multi-agent organization depicted in Fig. 5.9 was deployed with a key distinction. The target reading data was no longer aggregated as the outcome decisions (likelihoods) of each sensor node where aggregated later on, i.e. the CPA was used to aggregate likelihood measures instead of target data. The estimated likelihood measures at CA were sent to the CPA where they were aggregated and the final target class for each target was estimated according to the algorithm discussed in the previous section.

Table 5.8: Soft data description of target classes with very high vagueness

Target Class	**Description Data**
Spy	target1 moves certainly slow maneuvers certainly small
Passenger	target2 moves almost regular maneuvers almost medium
Threat	target3 moves slightly fast maneuvers slightly large

5.7.1 Preliminary Experimental Results

The experiments involved two series of tests were the majority of the total five (sensor) nodes, i.e. $SN2$, $SN4$, and $SN5$, were provided with high quality target description data whereas the rest of them, i.e. $SN1$ and $SN3$, had low quality target description data. In both tests we looked at the number of target readings required by the minority sensor nodes to achieve correct target classification for all detected targets. The former test scenario aimed at evaluating the distributed decision fusion through simple averaging while the latter involved a more challenging target classification task that required distributed decision fusion with weighted averaging to yield satisfactory results.

Tables 5.6 and 5.7 show the target description data provided to the sensor nodes in the

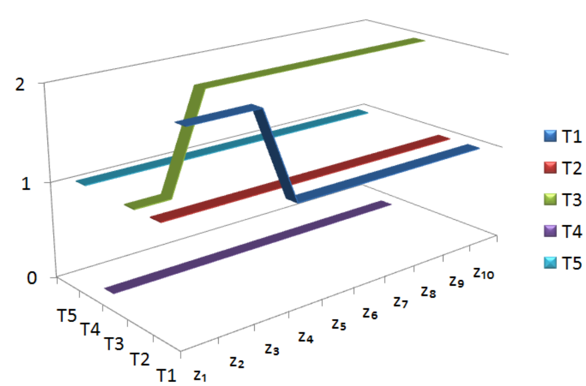

Figure 5.16: Distributed decision fusion experiment: the estimated target class over time for $SN1$ with no decision fusion (class 0: spy, class 1: passenger, class 2: threat).

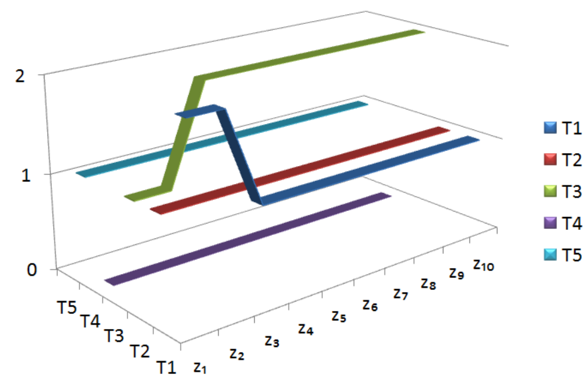

Figure 5.17: Distributed decision fusion experiment: the estimated target class over time for $SN1$ with decision fusion (class 0: spy, class 1: passenger, class 2: threat).

first test. The obtained experimental results for the first test scenario and sensor node $SN1$ are depicted in Figs. 5.16 and 5.17. As shown, deploying the distributed decision fusion scheme, $SN1$ is able to infer perfect target classification after three target readings instead of four. This could be contributed to decision fusion allowing for $SN1$ likelihood measures to be aggregated with those of high quality majority sensor nodes, hence deducing the correct target class faster.

The second test is more challenging due to the low quality data for the class threat being even more vague as shown in Table 5.8. Figs. 5.18, 5.19, and 5.20 show the obatined experimental results for the second test scenario with no distributed decision fusion, and distributed decision fusion with simple and weighted averaging schemes, respectively. The simple averaging was implemented with Cf set to one for all sensor nodes while for the weighted averaging case the majority sensor nodes, with higher quality target description data, were assigned a Cf twice as of the minority ones. Once again, deploying the distributed decision fusion results in a faster recognition, i.e. less number of target readings before perfect classification is achieved. However, in this case the improvements are more significant, i.e. reduction in target readings from nine to five, due to the test scenario being more challenging and also the weighted averaging scheme is shown to yield the superior performance, i.e. requiring only three target readings.

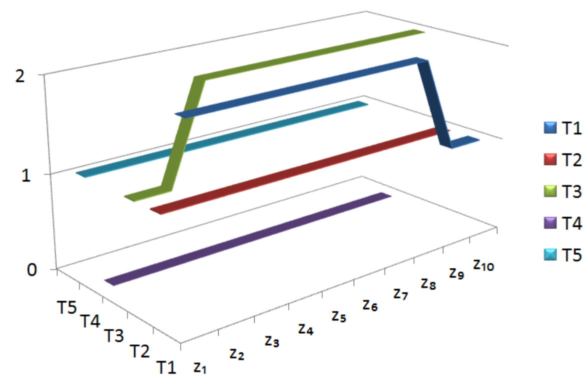

Figure 5.18: Distributed weighted decision fusion experiment: the estimated target class over time for $SN1$ with no decision fusion (class 0: spy, class 1: passenger, class 2: threat).

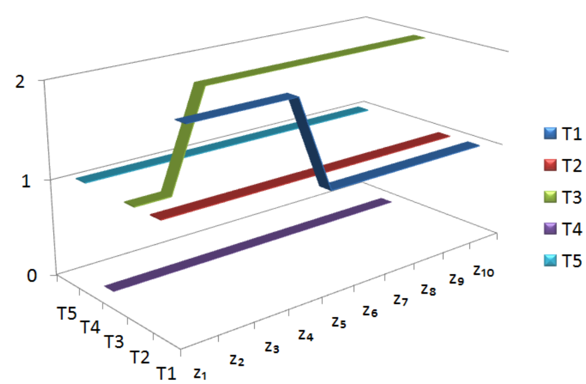

Figure 5.19: Distributed weighted decision fusion experiment: the estimated target class over time for $SN1$ with decision fusion (class 0: spy, class 1: passenger, class 2: threat).

5.8 Chapter Summary

A novel approach to enable tracking of multiple targets using soft human-generated data was presented in this chapter. The problem of measurement-to-track association for unconventional soft data, expressed using a RS theoretic formulation, was addressed and tackled efficiently as demonstrated by our experimental results. The proposed soft data association algorithm was shown to be capable of removing poor data from the association process, enhancing the target discrimination power in confusing situations, and improving tracking

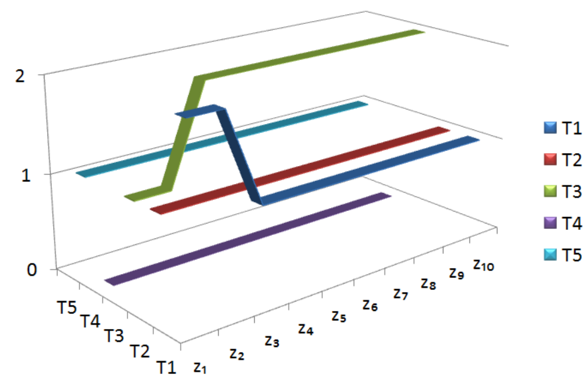

Figure 5.20: Distributed weighted decision fusion experiment: the estimated target class over time for $SN1$ with weighted decision fusion (class 0: spy, class 1: passenger, class 2: threat).

accuracy using the built-in inter-target checking mechanism. Currently, the number of targets is assumed to be known *a priori* at the current implementation. Future work will explore using soft data to estimate/update the number of targets over time dynamically.

Furthermore, deploying a random set theoretic representation of vague target description data and target data, a new classification method applicable to target classes described using soft human-generated data, which are inherently vague, was discussed. The preliminary experimental results showed the capability of the proposed classifier to perform robust target classification. Subsequently, the classification approach was enhanced by incorporating the classification vote of (potentially) all soft sensors through performing a distributed decision fusion stage before finalizing the estimate of the target class. Several experimental results demonstrated the improved classification performance yielded through the proposed distributed decision fusion algorithm.

Chapter 6

Concluding Remarks and Future Work

This chapter presents closing arguments for this dissertation. A summary of contributions, conclusive remarks, and potential directions for future research work are presented.

6.1 Highlights of the Thesis

This dissertation has considered the problem of a unified representational and computational framework to enable fusion of soft, as well as hard data. Adopting a random set theoretic approach, several novel, robust, and computationally efficient algorithms are developed to realize a soft/hard data fusion framework for the target tracking task. The proposed framework is applicable to both single-target and multi-target tracking applications and provides promising results as shown in several experimental scenarios.

For the case of single-target tracking, several contributions were established. First, assuming a linear-Gaussian target dynamics an extension of the popular Kalman filter within the RS theory, termed KEF, was adopted as the underlying data fusion mechanism. Second, a soft data ontology was developed to enable representation of soft data using the RS theoretic formulation of the KEF. Accordingly, the data modeling schemes for both soft and hard data were presented. Third, to deal with the trust-related issues regarding the soft human-generated data a human trust modeling approach was developed to enable discounting of undependable soft data on-the-fly. Fourth, distributed data aggregation algorithms for both soft and hard data were developed to enhance upon the scalability and robustness of the proposed RS theoretic framework.

For the multi-target tracking case, a key contribution was established with the introduction of a measurement-to-track association algorithm applicable to soft data modeled within the RS theory framework as fuzzy D-S state(s). The obtained experimental results demonstrated the ability of proposed soft data association algorithm to exclude poor data from the association process, improve the target discrimination power in confusing tracking scenarios, and lastly improve tracking accuracy by to avoiding poor associations using a built-in inter-target checking mechanism. Moreover, a novel distributed target classification approach applicable to targets classes described with soft human-generated data was presented. The preliminary experimental results showed the robustness of the proposed classifier with respect to both target data noise, and highly vague target description data.

Last but not least, this dissertation presented a critical review of data fusion state of the art methodologies. We introduced a new data centric taxonomy of data fusion methodologies, and explored challenging aspects and associated theoretical frameworks and algorithms existing in each of the categories. It is our hope for this work to serve as a review contribution of advances in the breadth of research on sensor data fusion, and to provide the data fusion community with a picture of the contemporary state of fusion literature.

6.2 Concluding Remarks

The successful design, development, and application of the RS theoretic approaches towards soft/hard data fusion in this dissertation have demonstrated that, random set theory is capable of providing viable and efficient methods to enable fusion of soft and hard data in a unified framework. In particular, the proposed algorithms for the single-target and multi-target tracking tasks provided a proof-of-concept framework where the powerful representational and computational abilities of the random set theory allowed for the intricate forms of data imperfection expressed by human-generated soft data to be properly modeled and processed.

The target tracking framework proposed in this dissertation may be deemed as the first step towards expanding the applicability domain of linear-Gaussian data fusion techniques based on the well-known Kalman filter from the hard conventional data into the unconventional soft data. The proposed human trust modeling scheme is among the first works to explore the trust-related issues regarding the unconventional human-generated data, contributing to the rather recent literature work on the data reliability in information fusion systems [7]. Moreoever, the distributed data aggregation approach developed in this dissertation is the first distributed data processing method based on the consensus propagation algorithm. It enhances the literature on distributed data fusion in terms of flexibility by separating the data propagation and data aggregation processes and thus allowing the non-contributing sensor nodes to still participate in the data propagation process, which is very conducive when aggregating soft data provided by human observers. Lastly, the distributed classification algorithm introduced for target classes described with vague soft data is among the first of such classifiers and also complements the proposed RS theoretic multi-target tracking framework.

6.3 Future Research Directions

Since the time around the inception of this dissertation, several works targeting soft, as well as soft/hard, data fusion systems have appeared [150, 151, 152, 165, 188] further consolidating the position of the soft/hard data fusion as an emerging and active area of study in the data fusion community. The existing literature mostly includes a Dempster-Shafer theoretic approach towards fusion of soft/hard, as well as preliminary efforts to develop standard datasets to be deployed to evaluate soft/hard data fusion systems. In spite of these preliminary studies, RS theory has not yet been explored as a viable solution. Very recently bishop et al. [171] have proposed a RS theoretic approach to enable state estimation where the inputs (or measurements), which are deployed to reduce the uncertainty in the state of a target, are soft data expressed in the natural language propositions form. Nonetheless, their work neither considers modeling both soft and hard data in a unified framework nor it is applicable to the multi-target tracking application.

In the following, we present several interesting directions for future work moving beyond the current scope of research efforts:

- The RS theory has been shown to be capable of representing first order, second order and even composite rules [92]. This makes it appealing to explore the potential of this theory to model human data provided in form of rule-based logical statements and incorporate them into the fusion process. A similar case could be imagined for modeling of the abundance of data available on the Web,
- The multi-target soft data ontology could be extended to enable human agents to supply richer reports regarding the target dynamics such as the appearance/disappearance of a target, as well as merging, and spawning of target(s),

- The number of targets is assumed to be known *a priori* at the current implementation. Future work could explore using soft data to dynamically estimate/update the number of targets over time,

- Current multi-target experiments are restricted to three targets. One could perform experiments with a larger number of targets and sensor nodes to further evaluate the efficiency of the proposed distributed data aggregation scheme in terms of its scalability. In addition, the current multi-target experiments have been conducted using soft data only, although the proposed framework is capable of processing both soft and hard for tracking. Future multi-target tracking experiments with both soft and hard data could further verify the applicability range of the proposed RS theoretic framework,

- The current framework assumes target with linear-Gaussian dynamics. This is not a limitation of the RS theory and is indeed imposed by the KEF in order to ensure all equations have a nice closed form. Future studies could attempt to tackle the problem of soft/hard data fusion for non-linear target dynamics. A potential solution could be based on RS theoretic formulation of the Bayes filter where the resultant likelihood and prior functions are approximated using popular techniques such as sequential Monte Carlo (particle filtering),

- Finally, the current experiments are performed in simulation environment only. This has had some undesired consequences such as restricting the length of experiments, which does not exceed a minute. This is caused by the fact that having several sensor nodes each comprised of several asynchronously communicating agents results in a large number of concurrent agents running on a single computer competing for computational resources with each other, as well as the P/S simulation platform itself. In our experiments, we observed that this issue led to a rather unstable system performance for simulations lasting more than a minute and accordingly we were forced to limit the collection (sampling) time for soft data to 1000 milliseconds or so. To achieve such sampling rates in our experiments, soft human reports were stored (as priory given) and then fed to the system through the HIA at the desired time instances. Although an unrealistic assumption, this issue could be easily tackled once the system implementation is ported onto a physical robotic platform, which thanks to the hardware abstraction layer provided by the Player server must be a straightforward procedure. This would allow individual sensor nodes to run their agents on separate computational resources, e.g. processing unit of robots, and thus almost eliminate the resource competition phenomenon and the related system unstability issue.

References

[1] I. Bloch, Information combination operators for data fusion: a comparative review with classification, IEEE Trans. on SMC: Part A 26 (1) (1996) 52-67.

[2] D.L. Hall, J. Llinas, An introduction to multisensor fusion, Proc. of the IEEE 85 (1) (1997) 6-23.

[3] D. Smith, S. Singh, Approaches to multisensor data fusion in target tracking: a survey, IEEE Trans. on Knowledge and Data Engineering 18 (12) (2006) 1696-1710.

[4] A.A. Goshtasby, S. Nikolov, Image fusion: advances in the state of the art, Information Fusion 8 (2) (2007) 114-118.

[5] I. Corona, G. Giacinto, C. Mazzariello, F. Roli, C. Sansone, Information fusion for computer security: state of the art and open issues, Information Fusion 10 (4) (2009) 274-284.

[6] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, S. Hübner, Ontology-based integration of information - a survey of existing approaches, in: Proc. of the IJCAI - Workshop on Ontologies and Information Sharing, 2001, pp. 108-117.

[7] G.L. Rogova, V. Nimier, Reliability in information fusion: literature survey, in: Proc. of the International Conference on Information Fusion, 2004, pp. 1158-1165.

[8] J.T. Yao, V.V. Raghavan, Z. Wu, Web information fusion: a review of the state of the art, Information Fusion 9 (4) (2008) 446-449.

[9] F.E. White, Data Fusion Lexicon, Joint Directors of Laboratories, Technical Panel for C3, Data Fusion Sub-Panel, Naval Ocean Systems Center, San Diego, 1991.

[10] L.A. Klein, Sensor and Data Fusion Concepts and Applications, second ed., Society of Photo-optical Instrumentation Engineers (SPIE), Bellingham, WA, 1999.

[11] H. Boström, S.F. Andler, M. Brohede, R. Johansson, A. Karlsson, J. van Laere, L. Niklasson, M. Nilsson, A. Persson, T. Ziemke, On the definition of information fusion as a field of research, Informatics Research Centre, University of Skövde, Tech. Rep. HS-IKI-TR-07-006, 2007.

[12] E.L. Walts, Data fusion for C3I: a tutorial, in: Command, Control, Communications Intelligence (C3I) Handbook, EW Communications Inc., Palo Alto, CA, 1986, pp. 217-226.

[13] A.N. Steinberg, C.L. Bowman, F.E. White, Revisions to the JDL data fusion model, in: Proc. of the SPIE Conference on Sensor Fusion: Architectures, Algorithms, and Applications III , 1999, pp. 430-441.

[14] J. Llinas, C. Bowman, G. Rogova, A. Steinberg, E. Waltz, F.E. White, Revisiting the JDL data fusion model II, in: Proc. of the International Conference on Information Fusion, 2004, pp. 1218-1230.

[15] B.V. Dasarathy, Decision Fusion, IEEE Computer Society Press, Los Alamitos, CA, 1994.

[16] I.R. Goodman, R.P.S. Mahler, H.T. Nguyen, Mathematics of Data Fusion, Kluwer Academic Publishers, Norwell, MA, 1997.

[17] M.M. Kokar, J.A. Tomasik, J. Weyman, Formalizing classes of information fusion systems, Information Fusion 5 (3) (2004) 189-202.

[18] M. Kumar, D.P. Garg, R.A. Zachery, A generalized approach for inconsistency detection in data fusion from multiple sensors, in: Proc. of the American Control Conference, 2006, pp. 2078-2083.

[19] P. Smets, Analyzing the combination of conflicting belief functions, Information Fusion 8 (4) (2007) 387-412.

[20] R.P.S. Mahler, "statistics 101" for multisensor, multitarget data fusion, IEEE Aerospace and Electronic Systems Magazine 19 (1) (2004) 53-64.

[21] R. Joshi, A.C. Sanderson, Multisensor Fusion: A Minimal Representation Framework, World Scientific, 1999.

[22] X.L. Dong, L. Berti-Equille, D. Srivastava, Truth discovery and copying detection in a dynamic world, Journal Proc. of the VLDB Endowment 2 (1) (2009) 562-573.

[23] Y. Zhu, E. Song, J. Zhou, Z. You, Optimal dimensionality reduction of sensor data in multisensor estimation fusion, IEEE Trans. on Signal Processing 53 (5) (2005) 1631-1639.

[24] B.L. Milenova, M.M. Campos, Mining high-dimensional data for information fusion: a database-centric approach, in: Proc. of International Conference on Information Fusion, 2005, pp. 638-645.

[25] P. Krause, D. Clark, Representing Uncertain Knowledge: An Artificial Intelligence Approach, Kluwer Academic Publishers, Norwell, MA, 1993.

[26] P. Smets, Imperfect Information: Imprecision and Uncertainty, in: A. Motro, P. Smets (Eds.), Uncertainty Management in Information Systems: From Needs to Solutions, Kluwer Academic Publishers, Norwell, MA, 1997, pp. 225-254.

[27] G.J. Klir, M.J. Wierman, Uncertainty-Based Information: Elements of Generalized Information Theory , second ed., Physica-Verlag HD, New York, 1999.

[28] D. Dubois, H. Prade, Formal Representations of Uncertainty, in: D. Bouyssou, D. Dubois, M. Pirlot, H. Prade (Eds.), Decision-Making Process-Concepts and Methods, ISTE & Wiley, London, 2009, pp. 85-156.

[29] M.C. Florea, A.L. Jousselme, E. Bosse, Fusion of imperfect information in the unified framework of random sets theory: application to target identification, Defence R & D Canada, Valcartier, Tech. Rep. ADA475342, 2007.

[30] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, Rough sets: a tutorial, in: S.K. Pal, A. Skowron (Eds.), Rough Fuzzy Hybridization: A New Trend in Decision Making, Springer, Singapore, 1999, pp. 1-98.

[31] F.K.J. Sheridan, A survey of techniques for inference under uncertainty, Artificial Intelligence Review 5 (1-2) (1991) 89-119.

[32] H.F. Durrant-Whyte, T.C. Henderson, Multisensor Data Fusion, in: B. Siciliano, O. Khatib (Eds.), Handbook of Robotics, Springer, 2008, pp. 585-610.

[33] L.A. Zadeh, Fuzzy sets, Information and Control 8 (3) (1965) 338-353.

[34] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1) (1978) 3-28.

[35] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Norwell, MA, 1992.

[36] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[37] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (2-3) (1990) 191-209.

[38] J. Yen, Generalizing the Dempster-Shafer theory to fuzzy sets, IEEE Trans. on SMC 20 (3) (1990) 559-570.

[39] R.P.S. Mahler, Statistical Multisource-Multitarget Information Fusion, Artech House, Boston, MA, 2007.

[40] G. Welch, G. Bishop, An introduction to the Kalman filter, Department of Computer Science, University of North Carolina, Tech. Rep. TR-95-041, 1995.

[41] S.J. Julier, J.K. Uhlmann, A new extension of the Kalman filter to nonlinear systems, in: International Symposium on Aerospace/Defense Sensing, Simulation and Controls, 1997, pp. 182-193.

[42] L.D. Stone, T.L. Corwin, C.A. Barlow, Bayesian Multiple Target Tracking, Artech House Inc., Norwood, MA, 1999.

[43] A. Doucet, N. de Freitas, N. Gordon, Sequential Monte Carlo Methods in Practice (Statistics for Engineering and Information Science), Springer, New York, 2001.

[44] B.A. Berg, Markov Chain Monte Carlo Simulations and Their Statistical Analysis, World Scientific, Singapore, 2004.

[45] D. Crisan, A. Doucet, A survey of convergence results on particle filtering methods for practitioners, IEEE Trans. on Signal Processing 50 (3) (2002) 736-746.

[46] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear/non-gaussian Bayesian state estimation, IEE Proc.-F 140 (2) (1993) 107-113.

[47] M.K. Pitt, N. Shephard, Filtering via simulation: auxiliary particle filters, Journal of the American Statistical Association 94 (446) (1999) 590-599.

[48] N. Metropolis, S. Ulam, The Monte Carlo method, Journal of American Statistics Association 44 (247) (1949) 335-341.

[49] W.K. Hastings, Monte Carlo sampling methods using markov chains and their applications, Biometrika 57 (1) (1970) 97-109.

[50] G. Casella, E.I. George, Explaining the Gibbs sampler, The American Statistician 46 (3) (1992) 167-174.

[51] J.S. Rosenthal, Parallel computing and Monte Carlo algorithms, Far East Journal of Theoretical Statistics 4 (2000) 207-236.

[52] M.K. Cowles, B.P. Carlin, Markov chain Monte Carlo convergence diagnostics: a comparative review, Journal of American Statistics Accosication 91 (434) (1996) 883-904.

[53] A.P. Dempster, A generalization of Bayesian inference (with discussion), Journal of the Royal Statistical Society Series B 30 (2) (1968) 205-247.

[54] T.D. Garvey, J.D. Lowrance, M.A. Fischler, An inference technique for integrating knowledge from disparate sources, in: Proc. of the International Joint Conference on Artificial Intelligence, 1981, pp. 319-325.

[55] B.R. Bracio, W. Horn, D.P.F. Moller, Sensor fusion in biomedical systems, in: Proc. of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1997, pp. 1387-1390.

[56] J.A. Barnett, Computational methods for a mathematical theory of evidence, in: Proc. of the International Joint Conference on Artificial Intelligence, 1981, pp. 868-875.

[57] J. Gordon, E.H. Shortliffe, A method for managing evidential reasoning in a hierarchical hypothesis space, Artificial Intelligence 26 (3) (1985) 323-357.

[58] A. Benavoli, B. Ristic, A. Farina, M. Oxenham, L. Chisci, An approach to threat assessment based on evidential networks, in: Proc. of the International Conference on Information Fusion, 2007, pp. 1-8.

[59] H.H.S. Ip, H. Tang, Parallel evidence combination on a SB-tree architecture, in: Proc. of the Australian and New Zealand Conference on Intelligent Information Systems, 1996, pp. 31-34.

[60] M. Bauer, Approximation algorithms and decision making in the Dempster-Shafer theory of evidence - an empirical study, International Journal of Approximate Reasoning 17 (2-3) (1997) 217-237.

[61] T. Denoeux, A.B. Yaghlane, Approximating the combination of belief functions using the fast möbius transform in a coarsened frame, International Journal of Approximate Reasoning, 31 (1-2) (2002) 77-101.

[62] M. Oxenham, The effect of finite set representations on the evaluation of Dempster's rule of combination, in: Proc. of the International Conference on Information Fusion, 2008, pp. 1-8.

[63] P.P. Shenoy, G. Shafer, Axioms for Probability and Belief-Function Propagation, in: R.R. Yager, L. Liu (Eds.), Classic Works of the Dempster-Shafer Theory of Belief Functions, Springer, Berlin/Heidel- berg, 2008, pp. 499-528.

[64] D. Dubois, H. Prade, Possibility theory and data fusion in poorly informed environments, Control Engineering Practice 2 (5) (1994) 811-823.

[65] P.J. Escamilla-Ambrosio, N. Mort, Hybrid Kalman filter-fuzzy logic adaptive multisensor data fusion architectures, in: Proc. of the IEEE Conference on Decision and Control, 2003, pp. 5215-5220.

[66] J.Z. Sasiadek, P. Hartana, Sensor data fusion using Kalman filter, in: Proc. of the International Conference on Information Fusion, 2000, pp. WED5/19-WED5/25.

[67] H. Zhu, O. Basir, A novel fuzzy evidential reasoning paradigm for data fusion with applications in image processing, Soft Computing Journal - A Fusion of Foundations, Methodologies and Applications 10 (12) (2006) 1169-1180.

[68] D. Dubois, H. Prade., Possibility Theory: An Approach to Computerized Processing of Uncertainty, Plenum Press, 1988.

[69] D. Dubois, H. Prade, Possibility theory in information fusion, in: Proc. of the International Conference on Information Fusion, 2000, pp. PS6-PS19.

[70] S. Destercke, D. Dubois, E. Chojnacki, Possibilistic information fusion using maximal coherent subsets, IEEE Trans. on Fuzzy Systems 17 (1) (2009) 79-92.

[71] D. Dubois, H. Prade, When upper probabilities are possibility measures, Fuzzy Sets and Systems 49 (1) (1992) 65-74.

[72] H. Borotschnig, L. Paletta, M. Prantl, A. Pinz, Comparison of probabilistic, possibilistic and evidence theoretic fusion schemes for active object recognition, Computing 62 (4) (1999) 293-319.

[73] S. Benferhat, C. Sossai, Reasoning with multiple-source information in a possibilistic logic framework, Information Fusion 7 (1) (2006) 80-96.

[74] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (1) (2007) 3-27.

[75] L. Yong, X. Congfu, P. Yunhe, A new approach for data fusion: implement rough set theory in dynamic objects distinguishing and tracing, in: Proc. of the IEEE International Conference on SMC, 2004, pp. 3318-3322.

[76] J.F. Peters, S. Ramanna, A. Skowron, J. Stepaniuk, Z. Suraj, Sensor fusion: a rough granular approach, in: Proc. of the IFSA World Congress and 20th NAFIPS International Conference, 2001, pp. 1367-1371.

[77] W. Haijun, C. Yimin, Sensor data fusion using rough set for mobile robots system, in: Proc. of the IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications, 2006, pp. 1-5.

[78] R.R. Yager, Generalized probabilities of fuzzy events from fuzzy belief structures, Information Sciences 28 (1) (1982) 45-62.

[79] O. Basir, F. Karray, Z. Hongwei, Connectionist-based Dempster-Shafer evidential reasoning for data fusion, IEEE Trans. on Neural Networks 6 (6) (2005) 1513-1530.

[80] D.S. Yeung, C. Degang, E.C.C. Tsang, J.W.T. Lee, X. Wang, On the generalization of fuzzy rough sets, IEEE Trans. on Fuzzy Systems 13 (3) (2005) 343-361.

[81] T. Guan, B. Feng, Rough fuzzy integrals for information fusion and classification, Lecture Notes in Computer Science 3066 (2004) 362-367.

[82] D.G. Kendall, Foundations of a Theory of Random Sets, in: E.F. Hardings, D.G. Kendall (Eds.), Stochastic Geometry, J. Wiley, 1974, pp. 322-376.

[83] V. Kreinovich, Random Sets Unify, Explain, and Aid Known Uncertainty Methods in Expert Systems, in: J. Goutsias, R.P.S. Mahler, H.T. Nguyen (Eds.), Random Sets: Theory and Applications, Springer-Verlag, 1997, pp. 321-345.

[84] S. Mori, Random Sets in Data Fusion. Multi-Object State-Estimation as a Foundation of Data Fusion Theory, in: J. Goutsias, R.P.S. Mahler, H.T. Nguyen (Eds.), Random Sets: Theory and Applications, Springer-Verlag, 1997, pp. 185-207.

[85] R.P.S. Mahler, Random sets: unification and computation for information fusion - a retrospective assessment, in: Proc. of the International Conference on Information Fusion, 2004, pp. 1-20.

[86] R.P.S. Mahler, PHD filters of higher order in target number, IEEE Trans. on Aerospace and Electronic Systems 43 (4) (2007) 1523-1543.

[87] B.T. Vo, B.N. Vo, A. Cantoni, The cardinalized probability hypothesis density filter for linear gaussian multi-target models, in: Proc. of the Annual Conference on Information Sciences and Systems, 2006, pp. 681-686.

[88] B.T. Vo, B.N. Vo, A. Cantoni, Analytic implemenations of the cardinalized probability hypothesis density filter, IEEE Trans. on Signal Processing 55 (7) (2007) 3553-3567.

[89] B.N. Vo, W.K. Ma, The gaussian mixture probability hypothesis density filter, IEEE Trans. on Signal Processing 54 (11) (2006) 4091-4104.

[90] B.N. Vo, S. Singh, A. Doucet, Sequential Monte Carlo methods for multi-target filtering with random finite sets, IEEE Trans. on Aerospace and Electronic Systems 41 (4) (2005) 1224-1245.

[91] K. Panta, B.N. Vo, S. Singh, A. Doucet, Probability hypothesis density versus multiple hypothesis tracking, in: Proc. of the SPIE Conference on Signal Processing, Sensor Fusion and Target Recognition, 2004, pp. 284-295.

[92] R.P.S. Mahler, A survey of PHD filter and CPHD filter implementations, in: Proc. of the SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition, 2007, pp. 65670O1-65670O12.

[93] D. Angelosante, E. Biglieri, M. Lops, Multiuser detection in a dynamic environment: joint user identification and parameter estimation, in: Proc. of the IEEE International Symposium on Information Theory, 2007, pp. 2071-2075.

[94] M. Maehlisch, R. Schweiger, W. Ritter, K. Dietmayer, Multisensor vehicle tracking with the probability hypothesis density filter, in: Proc. of the International Conference on Information Fusion, 2006, pp. 1-8.

[95] R.P.S. Mahler, Sensor management with non-ideal sensor dynamics, in: Proc. of the International Conference on Information fusion, 2004, pp. 783-790.

[96] B. Khaleghi, A. Khamis, F. Karray, Random finite set theoretic based soft/hard data fusion with application for target tracking, in: Proc. of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, 2010, pp. 50-55.

[97] B. Khaleghi, A. Khamis, F. Karray, and S. N. Razavi, Multisensor data fusion: A review of the state-of-the-art, Information Fusion (to appear), DOI: 10.1016/j.inffus.2011.08.001.

[98] B. Khaleghi and F. Karray, Distributed random set theoretic soft/hard data fusion: target tracking application, in: Proc. of IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing (to appear).

[99] X. Mingge, H. You, H. Xiaodong, S. Feng, Image fusion algorithm using rough sets theory and wavelet analysis, in: Proc. of the International Conference on Signal Processing, 2004, pp. 1041-1044.

[100] S.J. Julier, J.K. Uhlmann, A non-divergent algorithm in the presence of unknown correlation, in: Proc. of the American Control Conference, 1997, pp. 2369-2373.

[101] A. Makarenko, A. Brooks, T. Kaupp, H.F. Durrant-Whyte, F. Dellaert, Decentralised data fusion: a graphical model approach, in: Proc. of the International Conference on Information Fusion, 2009, pp. 545-554.

[102] P.S. Maybeck, Stochastic Models, Estimation and Control, Academic Press, New York, 1979.

[103] J.K. Uhlmann, S.J. Julier, H.F. Durrant-Whyte, A culminating advance in the theory and practice of data fusion, filtering and decentralized estimation, Covariance Intersection Working Group, Tech. Rep., 1997.

[104] C.Y. Chong, S. Mori, K.C. Chang, Distributed multitarget multisensor tracking, in: Y. Bar-Shalom (Ed.), Multitarget-Multisensor Tracking: Advanced Applications, Artech House, Norwood, MA, 1990, pp. 247-295.

[105] K.C. Chang, C.Y. Chong, S. Mori, On scalable distributed sensor fusion, in: Proc. of the International Conference on Information Fusion, 2008, pp. 1-8.

[106] A.D. Marrs, C.M. Reed, A.R. Webb, H.C. Webber, Data incest and symbolic information processing, United Kingdom Defence Evaluation and Research Agency, Tech. Rep., 1999.

[107] S.P. McLaughlin, R.J. Evans, V. Krishnamurthy, Data incest removal in a survivable estimation fusion architecture, in: Proc. of the International Conference on Information Fusion, 2003, pp. 229-236.

[108] L.Y. Pao, M. Kalandros, Algorithms for a class of distributed architecture tracking, in: Proc. of the American Control Conference, 1997, pp. 1434-1438.

[109] S.P. McLaughlin, R.J. Evans, V. Krishnamurthy, A graph theoretic approach to data incest management in network centric warfare, in: Proc. of the International Conference on Information Fusion, 2005, pp. 1162-1169.

[110] T. Brehard, V. Krishnamurthy, Optimal data incest removal in bayesian decentralized estimation over a sensor network, in: IEEE International Conference on Acoustics, Speech and Signal Processing, 2007, pp. III173-III176.

[111] W. Khawsuk, L.Y. Pao, Decorrelated state estimation for distributed tracking of interacting targets in cluttered environments, in: Proc. of the American Control Conference, 2002, pp. 899-904.

[112] M.B. Hurley, An information theoretic justification for covariance intersection and its generalization, in: Proc. of the International Conference on Information Fusion, 2002, pp. 505-511.

[113] L. Chen, P.O. Arambel, R.K. Mehra, Estimation under unknown correlation: covariance intersection revisited, IEEE Trans. on Automatic Control 47 (11) (2002) 1879-1882.

[114] W. Niehsen, Information fusion based on fast covariance intersection filtering, in: Proc. of the International Conference on Information Fusion, 2002, pp. 901-904.

[115] D. Franken, A. Hupper, Improved fast covariance intersection for distributed data fusion, in: Proc. of the International Conference on Information Fusion, 2005, pp. 154-160.

[116] A.R. Benaskeur, Consistent fusion of correlated data sources, in: Proc. of IEEE Annual Conference of the Industrial Electronics Society, 2002, pp. 2652-2656.

[117] Y. Zhou, J. Li, Robust decentralized data fusion based on internal ellipsoid approximation, in: Proc. of World Congress of the International Federation of Automatic Control, 2008, pp. 9964-9969.

[118] W.J. Farrell, C. Ganesh, Generalized chernoff fusion approximation for practical distributed data fusion, in: Proc. of the International Conference on Information Fusion, 2009, pp. 555-562.

[119] Z. Djurovic, B. Kovacevic, QQ-plot approach to robust Kalman filtering, International Journal of Control 61 (4) (1995) 837-857.

[120] S.J. Wellington, J.K. Atkinson, R.P. Sion, Sensor validation and fusion using the Nadaraya-Watson statistical estimator, in: Proc. of the International Conference on Information Fusion, 2002, pp. 321-326.

[121] P.H. Ibarguengoytia, L.E. Sucar, V. Vadera, Real time intelligent sensor validation, IEEE Trans. on Power Systems 16 (4) (2001) 770-775.

[122] J. Frolik, M. Abdelrahman, P. Kandasamy, A confidence-based approach to the self-validation, fusion and reconstruction of quasiredundant sensor data, IEEE Trans. on Instrumentation and Measurement 50 (6) (2001) 1761-1769.

[123] M. Kumar, D.P. Garg, R.A. Zachery, A method for judicious fusion of inconsistent multiple sensor data, IEEE Sensors Journal 7 (5) (2007) 723-733.

[124] M. Kumar, D. Garg, R. Zachery, Stochastic adaptive sensor modeling and data fusion, in: Proc. of the SPIE Conference on Smart Structures and Materials: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems, 2006, pp. 100-110.

[125] U. Orguner, F. Gustafsson, Storage efficient particle filters for the out of sequence measurement problem, in: Proc. of the International Conference on Information Fusion, 2008, pp. 1-8.

[126] S.S. Blackman, R. Popoli, Design and Analysis of Modern Tracking Systems, Artech House, 1999.

[127] Y. Bar-Shalom, Update with out-of-sequence measurements in tracking: exact solution, IEEE Trans. on Aerospace and Electronic Systems 38 (3) (2002) 769-777.

[128] M. Mallick, S. Coraluppi, C. Carthel, Advances in asynchronous and decentralized estimation, in: Proc. of the IEEE Aerospace Conference, 2001, pp. 4/1873-4/1888.

[129] K. Zhang, X.R. Li, Y. Zhu, Optimal update with out-of-sequence measurements for distributed filtering, in: Proc. of the International Conference on Information Fusion, 2002, pp. 1519-1526.

[130] K. Zhang, X.R. Li, Y. Zhu, Optimal update with out-of-sequence measurements, IEEE Trans. on Signal Processing 53 (6) (2005) 1992-2004.

[131] Y. Bar-Shalom, H. Chen, M. Mallick, One-step solution for the multistep out-of-sequence measurement problem in tracking, IEEE Trans. on Aerospace and Electronic Systems 40 (1) (2004) 27-37.

[132] Y. Anxi, L. Diannong, H. Weidong, D. Zhen, A unified out-of-sequence measurements filter, in: Proc. of the IEEE International Radar Conference, 2005, pp. 453-458.

[133] M. Mallick, J. Krant, Y. Bar-Shalom, Multi-sensor multi-target tracking using out-of-sequence measurements, in: Proc. of the International Conference on Information Fusion, 2002, pp. 135-142.

[134] S.R. Maskell, R.G. Everitt, R. Wright, M. Briers, Multi-target out-of-sequence data association: tracking using graphical models, Information Fusion 7 (4) (2006) 434-447.

[135] S. Challa, J.A. Legg, Track-to-track fusion of out-of-sequence tracks, in: Proc. of the International Conference on Information Fusion, 2002, pp. 919-926.

[136] A. Novoselsky, S.E. Sklarz, M. Dorfan, Track to track fusion using out-of-sequence track information, in: Proc. of the International Conference on Information Fusion, 2007, pp. 1-5.

[137] L.A. Zadeh, Review of "a mathematics theory of evidence", AI Magazine 5 (3) (1984) 81-83.

[138] M.C. Florea, A.L. Jousselme, E. Bosse, D. Grenier, Robust combination rules for evidence theory, Information Fusion 10 (2) (2009) 183-197.

[139] R.R. Yager, On the Dempster-Shafer framework and new combination rules, Information Sciences 41 (2) (1987) 93-137.

[140] P. Smets, The combination of evidence in the transferable belief model, IEEE Trans. on Pattern Analysis and Machince Intelligence 12 (5) (1990) 447-458.

[141] J. Dezert, Foundations for a new theory of plausible and paradoxical reasoning, Information Security 9 (2002) 13-57.

[142] E. Lefevre, O. Colot, P. Vannoorenberghe, Belief function combination and the conflict management, Information Fusion 3 (2) (2002) 149-162.

[143] F. Voorbraak, On the justification of Dempster's rule of combination, Artificial Intelligence 48 (2) (1991) 171-197.

[144] R. Haenni, Are alternatives to Dempster's rule of combination real alternatives?: comments on "about the belief combination and the conflict management problem" - Lefevre et al, Information Fusion 3 (3) (2002) 237-239.

[145] J. Dezert, A. Martin, F. Smarandache, Comments on "a new combination of evidence based on compromise" by K. Yamada, Fuzzy Sets and Systems 160 (2009) 853-855.

[146] J.K. Uhlmann, Covariance consistency methods for fault-tolerant distributed data fusion, Information Fusion 4 (3) (2003) 201-215.

[147] S. Maskell, A Bayesian approach to fusing uncertain, imprecise and conflicting information, Information Fusion 9 (2) (2008) 259-277.

[148] F. Rheaume, A.R. Benaskeur, Forward prediction-based approach to target-tracking with out-of-sequence measurements, in: Proc. of the IEEE Conference on Decision and Control, 2008, pp. 1326-1333.

[149] D.L. Hall, M. McNeese, J. Llinas, T. Mullen, A framework for dynamic hard/soft fusion, in: Proc. of the International Conference on Information Fusion, 2008, pp. 1-8.

[150] M.A. Pravia, R.K. Prasanth, P.O. Arambel, C. Sidner, C.Y. Chong, Generation of a fundamental data set for hard/soft information fusion, in: Proc. of the International Conference on Information Fusion, 2008, pp. 1-8.

[151] M.A. Pravia, O. Babko-Malaya, M.K. Schneider, J.V. White, C.Y. Chong, Lessons learned in the creation of a data set for hard/soft information fusion, in: Proc. of the International Conference on Information Fusion, 2009, pp. 2114-2121.

[152] K. Premaratne, M.N. Murthi, J. Zhang, M. Scheutz, P.H. Bauer, A Dempster-Shafer theoretic conditional approach to evidence updating for fusion of hard and soft data, in: Proc. of the International Conference on Information Fusion, 2009, pp. 2122-2129.

[153] A. Auger, J. Roy, Expression of uncertainty in linguistic data, in: Proc. of the International Conference on Information Fusion, 2008, pp. 1-8.

[154] D.L. Hall, M.D. McNeese, D.B. Hellar, B.J. Panulla, W. Shumaker, A cyber infrastructure for evaluating the performance of human centered fusion, in: Proc. of the International Conference on Information Fusion, 2009, pp. 1257-1264.

[155] P.J. Escamilla-Ambrosio, N. Mort, Hybrid Kalman filter-fuzzy logic adaptive multisensor data fusion architectures, in: Proc. of the IEEE Conference on Decision and Control, 2003, pp. 5215-5220.

[156] S. Young, J. Palmer, Pedigree and confidence: Issues in data credibility and reliability, in: Proc. of the International Conference on Information Fusion, 2007, pp. 1-8.

[157] V. Nimier, Introducing contextual information in multisensor tracking algorithms, in: Proc. of the International Conference on Processing and Management of Uncertainty in Knowledge-Based Systems, Advances in Intelligent Computing, 1994, pp. 595-604.

[158] B. Yu, K. Sycara, Learning the quality of sensor data in distributed decision fusion, in: Proc. of the International Conference on Information Fusion, 2006, pp. 1-8.

[159] F. Delmotte, L. Dubois, P. Borne, Context-dependent trust in data fusion within the possibility theory, in: Proc. of the IEEE International Conference on SMC, 1996, pp. 538-543.

[160] S.A. Sandri, D. Dubois, H.W. Kalfsbeek, Elicitation, assessment and pooling of expert judgments using possibility theory, IEEE Trans. on Fuzzy Systems 3 (3) (1995) 313-335.

[161] R. Haenni, S. Hartmann, Modeling partially reliable information sources: a general approach based on Dempster-Shafer theory, Information Fusion 7 (4) (2006) 361-379.

[162] Z. Elouedi, K. Mellouli, P. Smets, Assessing sensor reliability for multisensor data fusion within the transferable belief model, IEEE Trans. on SMC - Part B 34 (1) (2004) 782-787.

[163] E.J. Wright, K.B. Laskey, Credibility models for multi-source fusion, in: Proc. of the International Conference on Information Fusion, 2006, pp. 1-7.

[164] A. Karlsson, Dependable and generic high-level information fusion - methods and algorithms for uncertainty management, School of Humanities and Informatics, University of Skovde, Sweden, Tech. Rep. HS-IKI-TR-07-003, 2007.

[165] K. Sambhoos, J. Llinas, and E. Little, Graphical methods for real-time fusion and estimation with soft message data, in: Proc. of the International Conference on Information Fusion, 2008, pp. 1-8.

[166] G. Ferrin, L. Snidaro, S. Canazza, and G. L. Foresti, Soft data issues in fusion of video surveillance, in: Proc. of International Conference on Information Fusion, 2008, pp. 1-8.

[167] E. Pollard, A. Plyer, B. Pannetier, F. Champagnat, and G. L. Besnerais, GM-PHD filters for multi-object tracking in uncalibrated aerial videos, in: Proc. of International Conference on Information Fusion, 2009, pp. 1171-1178.

[168] A. M. Ali, R. E. Hudson and K. Yao, Tracking of random number of targets with random number of sensors using random finite set theory, in: Proc of IEEE International Conference on Acoustics, Speech and Signal Processing, 2009, pp. 2217-2220.

[169] R. P. S. Mahler, Multitarget bayes filtering via first-order multitarget moments, IEEE Trans. on Aerospace and Electronic Systems, vol. 39, no. 4, pp. 1152-1178, 2003.

[170] R. P. S. Mahler, A theory of PHD filters of higher order in target number, in: Proc. of SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition, 2006.

[171] A. N. Bishop and B. Ristic, Fusion of natural language propositions: Bayesian random set framework, in: Proc. of International Conference on Information Fusion, 2011, pp. 1-8.

[172] A. G. Dimakis, S. Kar, J. F. Moura, M. G. Rabbat, and A. Scaglione, Gossip algorithms for distributed signal processing, Proc. of the IEEE, vol. 98, no, 11, pp. 1847-1864, 2010.

[173] D. Bickson, O. Shental, and D. Dolev, Distributed Kalman filter via Gaussian belief propagation, in: Proc. of Allerton Conference, 2008, pp. 628-635.

[174] F. S. Cattivelli and A. H. Sayed, Diffusion strategies for distributed Kalman filtering and smoothing, IEEE Trans. on Autom. Control, vol. 55, no. 9, pp. 2069-2084, 2010.

[175] C. C. Moallemi and B. van Roy, Consensus propagation, IEEE Trans. on Info. Theory, vol. 52, no. 1, pp. 4753-4766, 2006.

[176] J. Sijs, M. Lazar, P. van den Bosch, and Z. Pepp, An overview of non-centralized Kalman filters, in: Proc. of the IEEE International Conference on Control Applications, 2008, pp. 739-744.

[177] R. P. S. Mahler, “statistics 101” for multisensor, multitarget data fusion, IEEE Aerospace and Electronic Systems Magazine, vol. 19, no. 1, pp. 53-64, 2000.

[178] Y. Bar-Shalom, and T. E. Fortmann, Tracking and data association, Academic Press, 1988.

[179] S. S. Blackman, Multiple hypothesis tracking for multiple target tracking, IEEE Aerospace and Electronic Systems Magazine, vol. 19, no. 1, pp. 5-18, 2004.

[180] H. W. Sorenson and D. L. Alspach, Recursive Bayesian estimation using Gaussian sums, Automatica, vol. 7, pp. 465-479, 1971.

[181] J. L. Speyer, Computation and transmission requirements for a decentralized linear-quadratic-gaussian control problem, IEEE Trans. on Automatic Control, vol. 24, no. 2, pp. 266-269, 1979.

[182] B. S. Y. Rao, H. F. Durrant-Whyte, and J. A. Sheen, A fully decentralized multi-sensor system for tracking and surveillance, Int. Journal of Robotics Research, vol. 12, no. 1, pp. 20-44, 1993.

[183] R. Olfati-Saber, Distributed Kalman filter with embedded consensus filters, in: Proc. of IEEE Conf. on Decision and Control, 2005, pp. 8179-8184.

[184] R. Olfati-Saber, J. A. Fax, and R. M. Murray, Consensus and cooperation in networked multi-agent systems, Proc. of the IEEE, vol. 95, no. 1, pp. 215-233. 2007.

[185] D. L. Hall and J. M. Jordan, Human-Centered Information Fusion, Boston, MA, USA: Artech House, 2010.

[186] B. T. Vo, Random Finite Sets in Multi-Object Filtering, PhD Thesis, School of Electrical, Electronic and Computer Engineering, The University of Western Australia, 2008.

[187] R. P. S. Mahler, Measurement-to-track association for nontraditional measurements, in: Proc. of the IEEE International Conference on Information Fusion, 2011, pp. 1-8.

[188] T. L. Wickramarathne, K. Premaratne, M. N. Murthi, M. Scheutz, S. Kubler, and M. Pravia, Belief theoretic methods for soft and hard data fusion, Proc. of IEEE ICASSP, 2011, pp. 2388-2391.