Agenzia nazionale per le nuove tecnologie,
l'energia e lo sviluppo economico sostenibile

MINISTERO DELLO SVILUPPO ECONOMICO

SOLUZIONE TECHNOLOGICA DI SUPPORTO ALLE DECISIONI ED INFORMATION FUSION PER GRANDI MOLI DI DATI IN CONTESTO SMART CITIES: Parte I – DEFINIZIONE DI METODOLOGIE METODI E MODELLI DELLA COMPLESSITA' PER IL DECISION MAKING PER LA SMART CITY PLATFORM

G. Iovane

UNIVERSITÀ DEGLI STUDI DI SALERNO

Report RdS/PAR2016/035 – Parte I

Soluzione Tecnologica di supporto alle decisioni ed information Fusion per grandi moli di dati in contesto smart cities:
Parte I – Definizione di metodologie, metodi e modelli della complessità per il decision making per la smart city

Gerardo Iovane - Dipartimento di Informatica (DI) – Università di Salerno

Settembre 2017

Report Ricerca di Sistema Elettrico

Accordo di Programma Ministero dello Sviluppo Economico - ENEA
Piano Annuale di Realizzazione 2016
Area: Efficienza energetica e risparmio di energia negli usi finali elettrici e interazione con altri vettori energetici
Progetto: D.6 Sviluppo di un modello integrato di smart district urbano
Obiettivo: D.6a Piattaforma ICT per la gestione di smart district
Responsabile del Progetto: Claudia Meloni ENEA

Il presente documento descrive le attività di ricerca svolte all'interno dell'Accordo di collaborazione "Soluzione Tecnologica di supporto alle decisioni ed Information fusion per grandi moli di dati in contesto smart cities"
Responsabile scientifico ENEA: Marta Chinnici
Responsabile scientifico UNISA: Gerardo Iovane

Indice

SOMMARIO 7

1 INTRODUZIONE 8

2 OUTLINE 10

3 STATO DELL'ARTE DEI SISTEMI DI INFORMATION FUSION 11

3.1 INTRODUZIONE AI SISTEMI DI FUSIONE DELLE INFOMAZIONI 11
3.2 CLASSIFICAZIONE DEI SISTEMI DI INFORMATION FUSION 13
3.2.1 Relazione tra le sorgenti 16
3.2.2 Livello di astrazione dei dati manipolati 16
3.2.3 Livello di astrazione di I/O 18
3.2.4 Tecniche di trasformazione dell'informazione 19
3.2.5 Topologia dell'architettura 19
3.2.5.1 Architettura centralizzata 19
3.2.5.2 Architettura distribuita 19
3.2.5.3 Architettura ibrida 20
3.2.5.4 Architettura Gerarchica 21
3.2.5.5 Architettura Olonica 21
3.2.5.6 Architettura Market-Based 21
3.2.6 Il processo di fusione: tipologia di informazioni trattate 23
3.2.7 Principali finalità del processo di fusione 23
3.3 MODELLI PER LA FUSIONE DELLE INFORMAZIONI 23
3.3.1 Modelli basati sull'informazione 26
3.3.1.1 Gerarchia DIKW 26
3.3.1.2 Gerarchia Rasmussen Information 27
3.3.1.3 Bosse Model 28
3.3.1.4 Il modello JDL 30
3.3.1.5 DFD Model 36
3.3.2 Modelli basati sulle attività 38
3.3.2.1 Intelligent Cycle model 38
3.3.2.2 OODA model 39
3.3.2.3 Architettura a cascata (waterfall) 41
3.3.2.4 Modello Omnibus 42
3.3.3 Modelli basati sulle regole 44
3.3.3.1 Object Oriented Model 44
3.3.3.2 Architettura Frankel-Bedworth 46
3.3.4 Modelli basati sulla situational awareness 48
3.3.4.1 Endesley Situational Awareness Model 48
3.3.4.2 Recognition Primed Decision Making 49
3.3.4.3 Fusion Situation Awareness Model 49
3.3.4.4 Perceptual Reasoning Model 50
3.3.4.5 SA Framework 51
3.3.4.6 Predictive Situation Awareness Systems 52
3.4 METODI PER LA FUSIONE DELLE INFORMAZIONI 53
3.4.1 Criticità connesse al processo di fusione delle informazioni 53
3.4.2 Principali classi di algoritmi per l'Information Fusion 55

4 ASPETTI PRESTAZIONALI DELLE TECNICHE DI INFORMATION FUSION 56

4.1 METODI PER LA FUSIONE DI DATI IMPERFETTI 56
4.1.1 Teoria della Probabilità 57
4.1.1.1 Estimatore Bayesiano 57
4.1.1.2 Metodo MAP (Maximum A Posteriori) 58

4.1.1.3 Filtro Moving Avarage 58
4.1.1.4 Metodo ML (Maximum Likelihood) 58
4.1.1.5 Metodi dei minimi quadrati 58
4.1.1.6 Metodi basati sulla simulazione Monte Carlo (MC): MC Sequenziale e Markov Chain MC 58
4.1.1.7 Particle Filter 59
4.1.1.8 Algoritmi Markov Chain Monte Carlo 59
4.1.1.9 Metropolis-Hastings Algorithm 59
4.1.1.10 Kalman Filter 59
4.1.1.11 Metodi EKF e UKF 60
4.1.1.12 Metodi Grid-Based 60
4.1.1.13 Hidden Markov Model (HMM) 60
4.1.2 Soft Computing Techniques 60
4.1.3 Fuzzy Set Theory 61
4.1.4 Fuzzy Cognitive Map 61
4.1.5 Fuzzy Ontology 61
4.1.6 Neural Network 61
4.1.7 Rough Set Theory 62
4.1.8 Possibility Theory 62
4.1.9 Dempster-Shafer Evidence Theory 62
4.1.10 Framework ibridi 63
4.1.10.1 Fuzzy Rough Set Theory 63
4.1.10.2 Fuzzy-Dempster-Shafer Theory 63
4.1.11 Random Set Theory 63
4.2 METODI PER L'ELIMINAZIONE DELLA CORRELAZIONE TRA I DATI 63
4.2.1 Metodi di Covariance Intersection 64
4.3 FUSIONE DI DATI INCONSISTENTI 64
4.3.1 Modellazione stocastica adattativa dei sensori per la riduzione degli outliers 64
4.3.2 OOSM/OOST Reducing Methods 64
4.3.3 Algoritmi Transferable Belief Model 65
4.3.4 Teoria di Dezert-Smarandache 65
4.3.5 Algoritmi di Covariance Union 65
4.4 FUSIONE DI DATI DISPARATI 65
4.4.1 Semantic Information Fusion 65
4.4.2 Abductive Reasoning 66
4.4.3 Clustering Algorithms (CA) 66

5 NUOVE FRONTIERE 67

5.1 LA FUSIONE NELL'EMERGENTE NET-CENTRIC ENVIRONMENT 67
5.2 ANALISI DEI COMPORTAMENTI IN RETE 68
5.3 APPROCCIO USER-CENTERED NELLA GESTIONE DEI SENSORI 68
5.4 TREND RECENTI PER LA FUSIONE NEGLI HIGH-LEVEL 68

6 CONCLUSIONI 71

7 RIFERIMENTI BIBLIOGRAFICI 72

Sommario

Il presente studio trae origine dall'Accordo di Programma tra il Ministero dello Sviluppo Economico (MiSE) e Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (ENEA) sulla Ricerca di Sistema Elettrico relativamente al Piano Annuale di Realizzazione 2016. Il presente documento rientra nell'ambito dell'obiettivo D.6a Piattaforma ICT per la gestione di smart districts che prevede la realizzazione di una piattaforma software in grado di raccogliere grandi moli di dati provenienti da distretti urbani. La gestione e l'estrazione delle informazioni dai dati possono aiutare nella creazione di servizi capaci di migliorare la qualità delle città e indirizzarle alla smartizzazione.
Lo studio inerente la tema in esame, è suddiviso in due rapporti tecnici.
L'obiettivo del presente rapporto è lo studio in ambito tecnico-scientifico delle metodologie di Information Fusion per il trattamento di grandi moli di dati che possano essere d'immediato e diretto impiego per il data digest nell'ambito d'interesse. In particolare, sono presentati i principali modelli e tecniche per la realizzazione di sistemi di fusione delle informazioni. Le informazioni raccolte costituiranno la base conoscitiva per la progettazione del sistema di fusione delle informazioni oggetto del progetto, fornendo altresì gli utili elementi innovativi per le diverse scelte tecnologiche da adottare all'interno del progetto.

1 Introduzione

I sistemi di fusione delle informazioni hanno suscitato notevole interesse nel tempo, offrendo sempre nuovi stimoli, vista la loro applicabilità nei domini applicativi più disparati.
Dal momento che spesso le terminologie adottate derivano dal campo di applicazione, in letteratura sono presenti diverse definizioni di Information Fusion.
L'Information Fusion è un insieme di tecniche e metodi che consentono di combinare informazioni provenienti da sorgenti diverse al fine di determinare un quadro dello scenario osservato più preciso e fornire strumenti di supporto alle decisioni ancora più affidabili. In definitiva, i sistemi di fusione delle informazioni mirano ad emulare il comportamento del cervello umano, che è in grado di fondere esperienze passate e dati raccolti dai diversi sensi al fine di acquisire conoscenza e consapevolezza dell'ambiente ed intraprendere azioni con la capacità di prevederne gli effetti. Nell'ambito di una sfida così complessa ma, al tempo stesso, affascinante, diversi passi in avanti sono stati compiuti dalla comunità scientifica, specialmente negli ultimi anni.
Le sorgenti possono differire sotto diversi punti di vista; ad esempio possono essere misure effettuate in punti diversi dello spazio (o anche in momenti diversi del tempo) su una stessa entità o su entità di differente natura. La tipologia di osservazioni che si possono effettuare su tali entità può spaziare, ad esempio, dalla misura effettuata da un comune sensore di campo all'analisi comportamentale effettuato per mezzo di tecniche automatiche di video analisi nell'ambito di un sistema di videomonitoring.
I sistemi di Information Fusion trovano applicazione nei contesti applicativi più disparati, dalle reti di sensori o la robotica fino all'elaborazione dei segnali. Molte applicazioni, specie quelle risalenti alle prime fasi del processo di Information Fusion, sono riconducibili all'area militare. Basti pensare ad esempio a sistemi di tracking per seguire le traiettorie di aerei, dove si sfrutta uno degli aspetti più interessanti legati all'Information Fusion: la capacità di acquisire rapidamente un quadro della situazione tale da consentire la predizione delle evoluzioni future del sistema.
Uno dei campi di applicazione, che ha maggiormente stimolato la ricerca negli ultimi anni, è la Homeland Security, con particolare riferimento alla lotta al crimine legata alla crescita di internet ed alla diffusione dei social network: in questi casi, infatti, gli investigatori hanno a disposizione una quantità di informazioni da elaborare elevatissima. Tale mole di dati è ingestibile se non si dispone di strumenti automatici in grado di filtrare e correlare le informazioni, cosa invece resa possibile da sistemi di fusione delle informazioni. L'obiettivo più stimolante ed al tempo stesso complesso per l'Information Fusion è quindi la possibilità di trasformare grandi quantità di informazioni in conoscenza [1]. E' evidente che quanto detto diventa ancora più significativo in cui si desidera avere un controllo intelligente di aree come è il caso delle smart city.
La progettazione di un sistema di fusione delle informazioni può ispirarsi ai diversi modelli presenti in letteratura. Tali modelli sono classificati sulla base di cosa viene considerato come elemento base del processo di fusione. L'elemento base del processo di fusione può essere il livello di astrazione delle informazioni, come nel JDL model, oppure le attività, come nell'OODA model. Si possono effettuare altre classificazioni, individuando così la categoria dei modelli basati sulle regole e la famiglia dei modelli che pongono particolare attenzione sulla situation awareness. La scelta del modello (o dei modelli) di riferimento viene in genere dettata da considerazioni su quale elemento (astrazione delle informazioni, attività, regole o consapevolezza della situazione) ricopra un ruolo fondamentale all'interno del dominio applicativo considerato.
Inoltre, la topologia dell'architettura con la quale si implementa il sistema di fusione delle informazioni gioca un ruolo importante sulle performance complessive. La ricerca negli ultimi anni si è orientata soprattutto rispetto a scenari basati su rete (net-centric environment). In tali scenari è stato dimostrato come si ottengano prestazioni migliori mediante l'utilizzo di reti oloniche e market-based. L'utilizzo di architetture oloniche porta ad una maggiore robustezza,

scalabilità ed efficienza del sistema di Information Fusion. Le architetture market-based, invece, hanno come principale vantaggio la capacità di orientare in maniera adattativa le sorgenti, al fine di ridurre le moli di dati in gioco quando non necessario.

Infine, il processo di fusione è fortemente influenzato dai metodi adoperati per la fusione dei dati in input. Tali prestazioni sono generalmente legate alla capacità dei metodi impiegati per trattare le problematiche legate alla natura dei dati. In letteratura sono presenti diversi metodi per la fusione delle informazioni, anche se generalmente non esistono algoritmi in grado di rispondere al meglio a tutte le problematiche relative ai dati. Ovviamente, a seconda dell'applicazione o al campo di impiego, possono essere combinati più metodi al fine di migliorare le prestazioni globali del sistema di Information Fusion.

Con tale premessa nella presente ricerca viene condotto uno studio sui più utili paradigmi di *Information Fusion al contesto della smart city*. In particolare, vengono presentati i principali modelli e tecniche per la realizzazione di sistemi di fusione delle informazioni, fornendo altresì gli utili elementi innovativi per le diverse scelte tecnologiche da adottare all'interno del progetto.

Il capitolo 2 offre una panoramica sintetica del documento. Nel terzo capitolo viene invece presentato lo stato dell'arte dei diversi modelli di fusione delle informazioni, con un focus particolare sulle caratteristiche dei modelli per i sistemi di *Information Fusion* presenti in letteratura. Nel capitolo **Errore. L'origine riferimento non è stata trovata.** sono invece discusse le tecniche per la fusione delle informazioni, analizzandone le prestazioni in funzione delle caratteristiche degli ingressi al sistema. Il capitolo **Errore. L'origine riferimento non è stata trovata.** si occupa di fornire le direzioni future della ricerca nell'ambito dell'*Information Fusion*.

Per concludere, il capitolo **Errore. L'origine riferimento non è stata trovata.** espone alcune considerazioni conclusive sullo studio effettuato.

2 Outline

I sistemi di fusione delle informazioni hanno suscitato notevole interesse nel tempo, offrendo sempre nuovi stimoli, vista la loro applicabilità nei domini applicativi più disparati.

Dal momento che spesso le terminologie adottate derivano dal campo di applicazione, in letteratura sono presenti diverse definizioni di *Information Fusion*. L'*Information Fusion* è un insieme di tecniche e metodi che consentono di combinare informazioni provenienti da sorgenti diverse al fine di determinare un quadro dello scenario osservato più preciso e fornire strumenti di supporto alle decisioni ancora più affidabili. In definitiva, i sistemi di fusione delle informazioni mirano ad emulare il comportamento del cervello umano, che è in grado di fondere esperienze passate e dati raccolti dai diversi sensi al fine di acquisire conoscenza e consapevolezza dell'ambiente ed intraprendere azioni con la capacità di prevederne gli effetti. Nell'ambito di una sfida così complessa ma, al tempo stesso, affascinante, diversi passi in avanti sono stati compiuti dalla comunità scientifica, specialmente negli ultimi anni.

Le sorgenti informative possono differire sotto diversi punti di vista; ad esempio possono essere misure effettuate in punti diversi dello spazio (o anche in momenti diversi del tempo) su una stessa entità o su entità di differente natura. La tipologia di osservazioni che si possono effettuare su tali entità può spaziare, ad esempio, dalla misura effettuata da un comune sensore di carica elettrica passante o temperatura all'analisi comportamentale effettuato per mezzo di tecniche automatiche di video analisi nell'ambito di un sistema di videomonitoring.

I sistemi di *Information Fusion* trovano applicazione nei contesti applicativi più disparati, dalle reti di sensori o la robotica fino all'elaborazione dei segnali. Molte applicazioni, specie quelle risalenti alle prime fasi del processo di *Information Fusion*, sono riconducibili all'area militare. Basti pensare ad esempio a sistemi di *tracking* per seguire le traiettorie di aerei, dove si sfrutta uno degli aspetti più interessanti legati all'*Information Fusion*: la capacità di acquisire rapidamente un quadro della situazione tale da consentire la predizione delle evoluzioni future del sistema.

Uno dei campi di applicazione, che ha maggiormente stimolato la ricerca negli ultimi anni, è la *Homeland Security*, con particolare riferimento alla lotta al crimine legata alla crescita di internet ed alla diffusione dei social network: in questi casi, infatti, gli investigatori hanno a disposizione una quantità di informazioni da elaborare elevatissima. Tale mole di dati è ingestibile se non si dispone di strumenti automatici in grado di filtrare e correlare le informazioni, cosa invece resa possibile da sistemi di fusione delle informazioni. L'obiettivo più stimolante ed al tempo stesso complesso per l'*Information Fusion* è quindi la possibilità di trasformare grandi quantità di informazioni in conoscenza [1]. Da ciò si comprende quanto tali tecniche possano essere utili ed efficaci nel contesto dei big data legati alla Smart City.

La progettazione di un sistema di fusione delle informazioni può ispirarsi ai diversi modelli presenti in letteratura. Tali modelli sono classificati sulla base di cosa viene considerato come elemento base del processo di fusione. L'elemento base del processo di fusione può essere il *livello di astrazione delle informazioni*, come nel JDL *model*, oppure le *attività*, come nell'OODA model. Si possono effettuare altre classificazioni, individuando così la categoria dei modelli basati sulle *regole* e la famiglia dei modelli che pongono particolare attenzione sulla *situation awareness*. La scelta del modello (o dei modelli) di riferimento viene in genere dettata da considerazioni su quale elemento (astrazione delle informazioni, attività, regole o consapevolezza della situazione) ricopra un ruolo fondamentale all'interno del dominio applicativo considerato.

Inoltre, la topologia dell'architettura con la quale si implementa il sistema di fusione delle informazioni gioca un ruolo importante sulle *performance* complessive. La ricerca negli ultimi anni si è orientata soprattutto rispetto a scenari basati su rete (*net-centric environment*). In tali scenari è stato dimostrato come si ottengano prestazioni migliori mediante l'utilizzo di reti

oloniche e *market-based*. L'utilizzo di architetture oloniche porta ad una maggiore robustezza, scalabilità ed efficienza del sistema di *Information Fusion*. Le architetture *market-based*, invece, hanno come principale vantaggio la capacità di orientare in maniera adattativa le sorgenti, al fine di ridurre le moli di dati in gioco quando non necessario.
Infine, il processo di fusione è fortemente influenzato dai metodi adoperati per la fusione dei dati in input. Tali prestazioni sono generalmente legate alla capacità dei metodi impiegati di trattare le problematiche legate alla natura dei dati. In letteratura sono presenti diversi metodi per la fusione delle informazioni, anche se generalmente non esistono algoritmi in grado di rispondere al meglio a tutte le problematiche relative ai dati. Ovviamente, a seconda dell'applicazione, possono essere combinati più metodi al fine di migliorare le prestazioni globali del sistema di *Information Fusion*.

3 Stato dell'arte dei sistemi di Information Fusion

All'interno di questo capitolo viene fornita una panoramica sullo stato dell'arte dei sistemi di fusione delle informazioni.
Dapprima vengono presentate le varie definizioni legate all'*Information Fusion*. Le diverse definizioni che il mondo scientifico ha fornito sono legate sia ad una evoluzione temporale sia ad un cambiamento riconducibile al punto di vista sotto il quale si osserva un sistema di *Information Fusion*.
Successivamente si illustrano i criteri mediante i quali è possibile classificare i sistemi di fusione delle informazioni.
Infine sono trattati i principali modelli e metodi per *l'Information Fusion* presenti in letteratura.

3.1 Introduzione ai sistemi di Fusione delle Infomazioni

I sistemi di *Information Fusion* trovano impiego in svariati contesti applicativi, a partire da applicazioni in ambito militare a quelle in ambito medico passando per la robotica, l'elaborazione dei segnali e tante altri domini applicativi. Negli ultimi anni il mondo scientifico ha rivolto notevole attenzione all'utilizzo dell'*Information Fusion* per l'analisi del Web e dei *social network* e, specie dopo l'11 settembre 2001, particolare attenzione è stata posta sull'impiego dell'*Information Fusion* nel campo della *Homeland Security*.
A causa della notevole versatilità dell'*Information Fusion*, in letteratura sono presenti diverse definizioni, che spesso adottano terminologie differenti anche se intercambiabili. Tuttavia, data la vasta gamma di domini applicativi, si rende sempre più necessaria la definizione di standard relativi a protocolli di valutazione delle prestazioni dei sistemi di fusione delle informazioni [2].

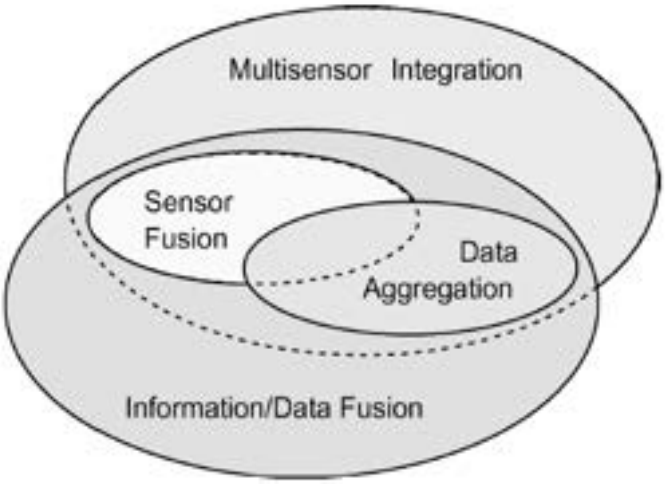

Figura 1. Information Fusion : relazioni nella terminologia [3]

In [3] viene illustrato come i concetti di *Data Fusion* e *Information Fusion*, spesso presenti in letteratura, possano in realtà essere usati in maniera indistinta. In particolare la Figura 1 sintetizza le relazioni che intercorrono tra i vari termini presenti in letteratura.
Risulta evidente che la fusione dei dati (o informazioni, dato che i termini possono essere usati in maniera interscambiabile) provenienti da sensori costituisce un sottoinsieme dell'*Information/Data Fusion*: in questa circostanza di parla di *Sensor Fusion* o *Multisensor Fusion*, dove i sensori possono non essere necessariamente della stessa tipologia.
L'aggregazione dei dati, in alcuni casi provenienti da sensori, consente di ridurre la ridondanza degli stessi e rendere il contenuto informativo più efficiente. In ragione di ciò la *Data Aggregation* può essere vista come un sottoinsieme della Information Fusion, con alcuni elementi in comune con il sottoinsieme della *Sensor Fusion*.
Infine, viene presentato il concetto di *Multisensor Integration*. La caratteristica principale della *Multisensor Integration* è l'impiego della fusione delle informazioni al fine di prendere delle decisioni che siano più attendibili di quelle che si potrebbero prendere se si analizzassero le fonti informative singolarmente. Anche la *Multisensor Integration* si può quindi vedere come un insieme che include inoltre gli insiemi *Multisensor/Sensor Fusion* e *Data Aggregation* in quanto, al fine di intraprendere azioni per interagire con l'ambiente, vengono adoperati tra l'altro dati provenienti da sensori e/o loro aggregazioni. La Multisensor Integration riveste una rilevanza particolare nell'ambito del presente progetto vista la ampia possibilità di impiego di sensori di varia natura.
In letteratura sono presenti diverse definizioni di *Information Fusion*, legate spesso all'ambito in cui si applicano. La prima più autorevole definizione che si trova in letteratura è stata fornita dal DoD (Dipartimento della Difesa) USA all'interno del *Data Fusion Subpanel* del *Joint Directors of Laboratories* (JDL) [4] nel 1991. Il JDL definisce l'*Information Fusion* come un processo multilivello e multiforme in grado di gestire in maniera automatica il rilevamento, l'associazione, la correlazione, la stima e la combinazione di dati e informazioni provenienti da sorgenti differenti.
In seguito Klein ha generalizzato questa definizione, asserendo che i dati in ingresso al sistema di fusione delle informazioni possono provenire sia da sorgente singola che da sorgenti multiple [5].
Queste prime definizioni per *l'Information Fusion* hanno carattere generale e possono essere utilizzate quindi in diversi campi applicativi.
Nel 1999 Wald ha spostato il focus della definizione sul *framework* impiegato per la fusione delle informazioni. In particolare, egli ebbe ad affermare che la fusione dei dati è costituita da un

framework formale, all'interno del quale sono espressi i significati e gli strumenti per l'elaborazione di dati provenienti da diverse sorgenti [6]. Inoltre, nella definizione di Wald, i dati provenienti dalla stessa sorgente ma ottenuti in istanti differenti sono considerati a tutti gli effetti come provenienti da sorgenti differenti.
Un'altra importante definizione fu data da Dasarathy nel 2001 [7]; sulla base di tale definizione, la fusione delle informazioni è vista come un insieme di teorie, tecniche e strumenti in grado di sfruttare la sinergia presente all'interno delle informazioni, raccolte da sorgenti differenti, allo scopo di prendere decisioni o intraprendere azioni. Tali azioni risultano migliori sotto diversi aspetti qualitativi e quantitativi rispetto a quelle che si attiverebbero se si considerassero le sorgenti individualmente.
Nel 2002 Waltz e altri [8] hanno presentato una definizione di *Information Fusion* che evidenzia le relazioni tra informazione e conoscenza. In particolare il processo di fusione viene visto come un processo di creazione della conoscenza. All'interno di questo processo possono essere individuate tre entità principali:

- **Data**
- **Information**
- **Knowledge**

L'entità **Data** è costituita da osservazioni (omogenee o eterogenee) che vengono allineate, correlate e combinate all'interno di insiemi organizzati ed indicizzati. Tali insiemi denotano così l'**Information** e possono essere riportati su modelli mirati a descrivere l'aspetto ed il comportamento del dominio sotto osservazione (detto **Knowledge**).
Nel corso di questa seconda annualità, al fine di consentire lo sviluppo di una piattaforma ENEA software (si veda D6.a2) capace di accogliere e gestire grandi moli di dati, il processo di conoscenza del il dato con le sue tre entità (sopra indicate) diviene il principale attore dello studio presentato nel presente documento.
In definitiva l'*Information Fusion* appare come un processo che elabora dati provenienti da diverse sorgenti al fine di ottenere informazioni sostanzialmente più affidabili. La definizione di Jie Yu (2007) [9] pare riassumere tale concetto, definendo l'*Information Fusion* come il processo atto ad:

- *associare, correlare e combinare dati e informazioni che provengono da*
- *singoli o multipli sensori o sorgenti al fine di ottenere*
- *stime raffinate di parametri, caratteristiche, eventi e comportamenti delle entità osservate all'interno del dominio di interesse.*

La fusione delle informazioni produce diversi vantaggi ma il beneficio più evidente consiste in un arricchimento del contenuto informativo. Alcuni dei miglioramenti ottenibili sono:

- riduzione dell'ambiguità e quindi dell'incertezza della scelta finale;
- riduzione della vulnerabilità del sistema;
- aumento della precisione del sistema;
- miglioramento della robustezza del sistema;
- miglioramento dell'affidabilità del sistema;
- estensione della copertura nei domini spaziale e temporale.

L'*Information Fusion* produce un miglioramento della capacità di ottenere informazioni attraverso l'attivazione di un processo che riduce sensibilmente le ambiguità che si possono avere. Tali ambiguità sono particolarmente evidenti quando le sorgenti informative sono differenti e forniscono informazioni che possono in alcuni casi contraddirsi a vicenda.

La robustezza di un sistema che adotta un *framework* di fusione delle informazioni risulta maggiore a causa del contributo di diverse sorgenti, permettendo così al sistema di rimediare alla mancanza di dati provenienti da una o più sorgenti. Questo migliora anche la confidenza con la quale si possono stimare caratteristiche del sistema in esame, dal momento che dati non attendibili vengono confrontati con rilevazioni differenti.
L'information fusion ci fornisce il principale strumento per approfondire la conoscenza dei dati presenti nei DB dei contesti verticali che presentano i diversi casi d'uso (riportati nel report D6).

3.2 Classificazione dei sistemi di Information Fusion

I sistemi di fusione delle informazioni possono essere classificati sotto diversi punti di vista; in Figura 2, vengono illustrati alcuni dei modi in cui tale classificazione è possibile. I sistemi di *Information Fusion* possono essere basati su:

- natura delle relazioni tra le sorgenti;
- livello di astrazione dei dati elaborati;
- livello di astrazione degli ingressi e delle uscite del sistema;
- metodo di trasformazione dell'informazione;
- topologia dell'architettura;
- tipologia di informazioni trattate;
- obiettivo del processo di fusione.

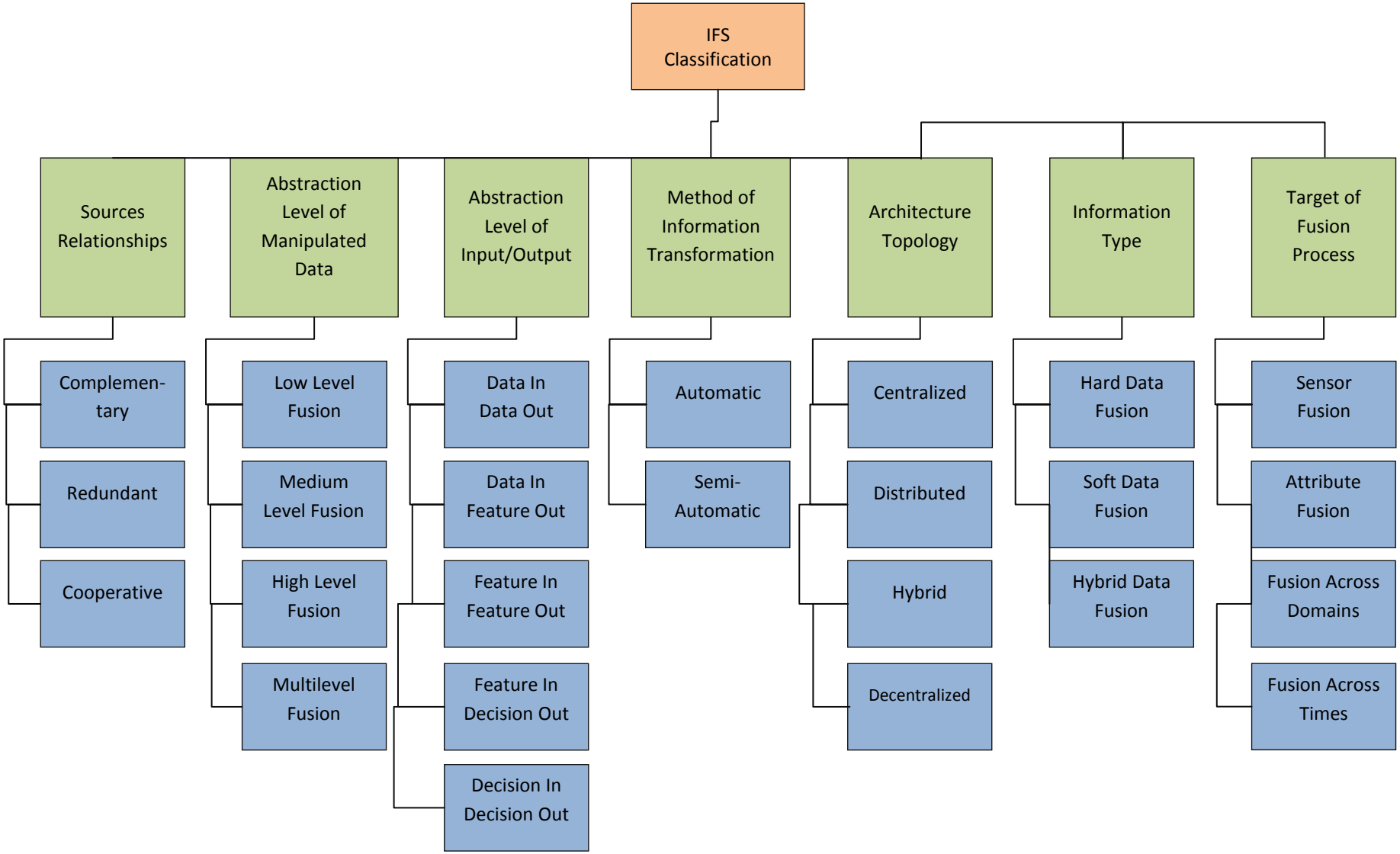

Figura 2. Classificazione dei diversi sistemi di Information Fusion

3.2.1 Relazione tra le sorgenti

I sistemi di fusione dell'informazione possono essere classificati in funzione delle relazioni che intercorrono tra le sorgenti [3], [10], [11].

In particolare, possono essere quindi suddivisi in:

- complementari;
- ridondanti;
- cooperativi.

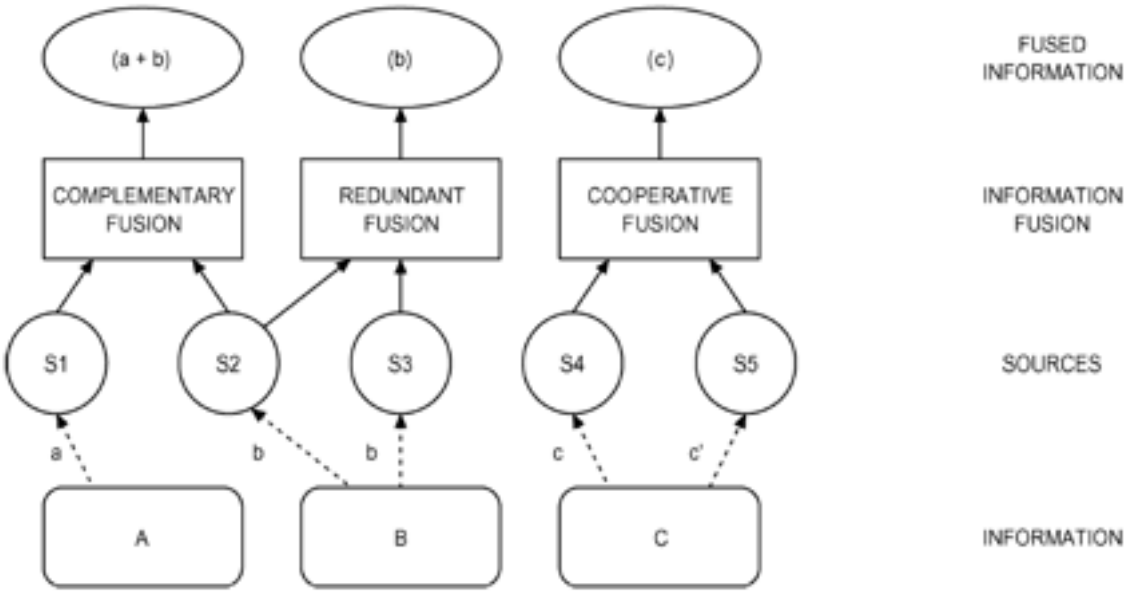

Figura 3. Information Fusion: Classificazione basata sulla relazione tra le sorgenti

Si parla di fusione *complementare* quando gli ingressi sono osservazioni complementari di una stessa realtà e, di conseguenza, la fusione consente di ottenere informazioni più complete e dettagliate.

Si indica la fusione come *ridondante* quando prende in ingresso da sorgenti indipendenti lo stesso pezzo di informazione al fine di produrre un'informazione in genere più accurata delle originali.

Infine, il processo di fusione si definisce come *cooperativo* quando le informazioni sul campo provenienti da sorgenti indipendenti vengono fuse per avere informazioni in genere più complesse.

3.2.2 Livello di astrazione dei dati manipolati

I sistemi di *Information Fusion* possono essere classificati a seconda del livello di astrazione dei dati manipolati [12], [13], [14], [15]. Le classificazioni possibili sono:

- *Low-Level Fusion,*
- *Medium-Level Fusion,*
- *High-Level Fusion,*
- *Multilevel Fusion.*

Si parla di *Low-Level Fusion* quando i dati che vengono manipolati sono prodotti dai sensori (in tal caso si parla appunto di *signal* o *measurement*) e si utilizzano quindi dati grezzi non processati. Il risultato di questa elaborazione è costituito da un insieme di dati più raffinati rispetto ai primi. A valle di questa elaborazione l'eventuale ridondanza è stata ridotta mediante aggregazione e sono state risolte altre problematiche, come ad esempio la presenza di dati spuri.

Nella *Medium-Level Fusion* gli attributi o le caratteristiche di una o più entità vengono elaborate al fine di ottenere una mappa delle caratteristiche utili allo scopo dell'applicazione. In questo processo di fusione i

dati manipolati possono essere ad esempio la forma o la posizione di un oggetto e si indica questo livello di astrazione anche con i termini *feature level* o *attribute level.*
All'interno della *High-Level Fusion* i dati che vengono fusi sono in realtà decisioni o simboli: la loro fusione produce quindi decisioni o simboli più precise ed affidabili.
Infine, la *Multilevel Fusion* rappresenta un sistema di fusione in cui i dati che vengono manipolati appartengono a diversi livelli di astrazione.
La fusione delle informazioni ai livelli più alti di astrazione non è molto consolidata in letteratura. Uno sforzo maggiore in questa direzione consentirebbe notevoli miglioramenti nella fase di supporto alla decisione, riducendo rischi e falsi allarmi specie nei casi in cui le ipotesi possono essere affette da maggiore impredicibilità. Esempi concreti possono essere casi in cui bisogna individuare possibili gruppi di persone pericolose o effettuare operazioni di salvataggio [16]. In quest'ultimo caso i soccorritori sono chiamati a prendere decisioni sotto condizioni di forte stress ed è quindi molto importante avere a disposizione strumenti che supportino la loro capacità cognitiva. Tale scopo può essere raggiunto se il sistema di fusione garantisce una forte capacità di correlare mediante regole informazioni appartenenti a categorie differenti. In particolare in [16] gli autori propongono un metodo automatico di elaborazione di dati incerti che suggerisce e valuta simultaneamente diverse ipotesi riguardo la presenza o meno di vittime.

3.2.3 Livello di astrazione di I/O

Gli ingressi e le uscite possono essere viste secondo differenti livelli di astrazione. In ingresso e in uscita al processo di fusione si possono quindi avere (con livello di astrazione in ordine crescente) *Data*, *Feature* o *Decision* [17]. Le combinazioni possibili di questi livelli di astrazione sono:

- DAI-DAO: i dati non processati in ingresso sono fusi producendo dati in uscita anch'essi non processati.
- DAI-FEO: i dati non processati in ingresso vengono fusi al fine di ottenere in uscita degli attributi dello scenario osservato.
- FEI-FEO: gli attributi in ingresso vengono trasformati all'interno del processo di fusione in attributi generalmente più raffinati.
- FEI-DEO: gli attributi in ingresso consentono al processo di fusione di produrre delle decisioni che potranno essere seguite.
- DEI-DEO: le decisioni in ingresso possono essere ulteriormente elaborate al fine di ottenere altre decisioni maggiormente affidabili.

Le categorie di *Information Fusion* ottenibili in funzione al livello di astrazione degli ingressi e delle uscite possono essere mappate su quelle ottenibili dalla classificazione fatta in funzione dei dati manipolati. Molte analogie infatti risultano da un confronto tra questa classificazione e quella presentata nel paragrafo 3.2.2, anche se nella classificazione in funzione del livello di astrazione degli ingressi e delle uscite vi è un livello di granularità più fine. La principale differenza è che si può specificare i livelli di astrazione basandosi non solo sugli input ma anche sugli output del processo di fusione riducendo possibili ambiguità.

3.2.4 Tecniche di trasformazione dell'informazione

I processi di fusione delle informazioni possono essere classificabili inoltre in automatici e in semi-automatici. All'interno dei sistemi di fusione semi-automatici l'operatore è presente all'interno dell'architettura e influenza con le proprie scelte l'output del processo. Classicamente, quando i dati da manipolare o da produrre sono a basso livello di astrazione, il processo di fusione è completamente automatico. Ai livelli più alti invece l'uomo può essere visto anch'esso come una sorgente che influenza il processo di fusione.

3.2.5 Topologia dell'architettura

I sistemi per la fusione delle informazioni possono essere classificate anche dal punto di vista di come le informazioni vengono scambiate fra i vari moduli [18]. L'architettura impiegata influenza notevolmente le *performance* del sistema di *Information Fusion* e va quindi scelta opportunamente in funzione delle caratteristiche del dominio applicativo in esame.

3.2.5.1 Architettura centralizzata

Nelle architetture centralizzate vi è un unico centro di fusione al quale arrivano i dati dai sensori. All'interno del centro di fusione avvengono poi i processi di allineamento e fusione dei dati.
Il vantaggio principale di una architettura centralizzata è la possibilità di ridurre i costi legati al numero di sensori e soprattutto al numero dei centri di elaborazione delle informazioni. Tuttavia vi è una grossa mole di dati grezzi che deve essere inviata al centro, il che è particolarmente critico in applicazioni in cui si vuole tenere basso il traffico, come nel caso delle *Wireless Sensor Network* (WSN); il flusso di dati dai sensori al centro di fusione può in aggiunta causare colli di bottiglia che abbassano notevolmente le prestazioni del sistema di fusione.
Inoltre, il sistema risulta particolarmente sensibile ai *failure* dei link di comunicazione tra i sensori e il centro e, ovviamente, soprattutto alla criticità relativa alla vulnerabilità del centro di fusione.
In genere un sistema di fusione dei dati centralizzato è meno flessibile rispetto ad altre topologie e quindi, se più sensori vengono aggiunti al sistema, si deve in qualche modo ripensare il sistema eventualmente sotto diversi punti di vista (banda richiesta, potenza, etc.).

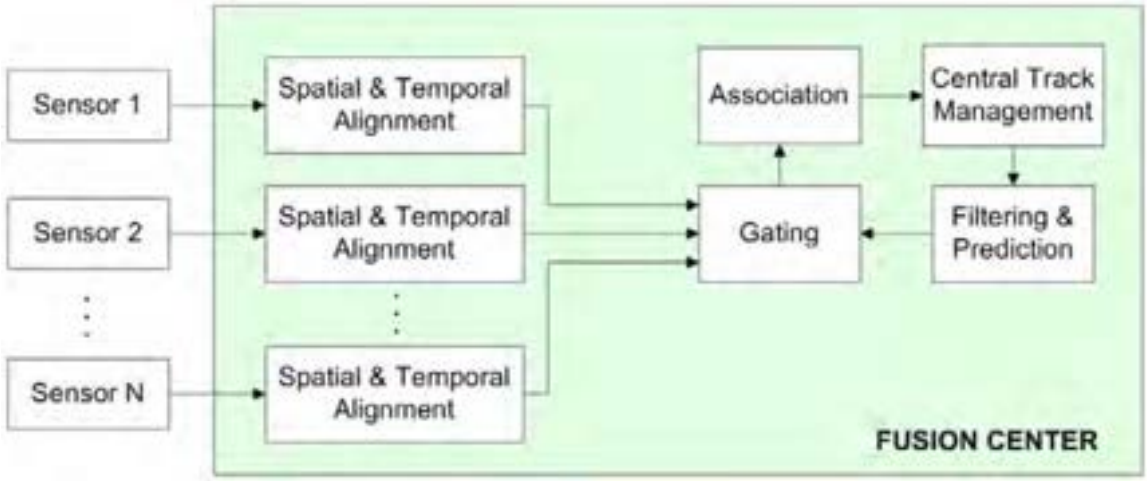

Figura 4. Architettura centralizzata di un sistema di Fusione delle Informazioni

3.2.5.2 Architettura distribuita

Nelle architetture distribuite vi sono *n* nodi a ciascuno dei quali arriva un dato insieme di dati grezzi. In tali nodi i dati vengono pretrattati producendo dei dati più raffinati o attributi. Questi vengono poi inviati al

centro di fusione nel quale avviene un'ulteriore elaborazione. Un'architettura distribuita in genere può auto-adattarsi a modifiche relative alla sua struttura.
Le architetture distribuite consentono quindi un uso più efficiente del centro di fusione in quanto vi è una progressiva riduzione della mole di dati da elaborare e inoltre vi è una tolleranza maggiore ai *failure.*
Tali vantaggi sono comunque ottenuti a discapito di un aumento delle risorse e quindi dei costi necessari alla realizzazione di un'architettura distribuita.

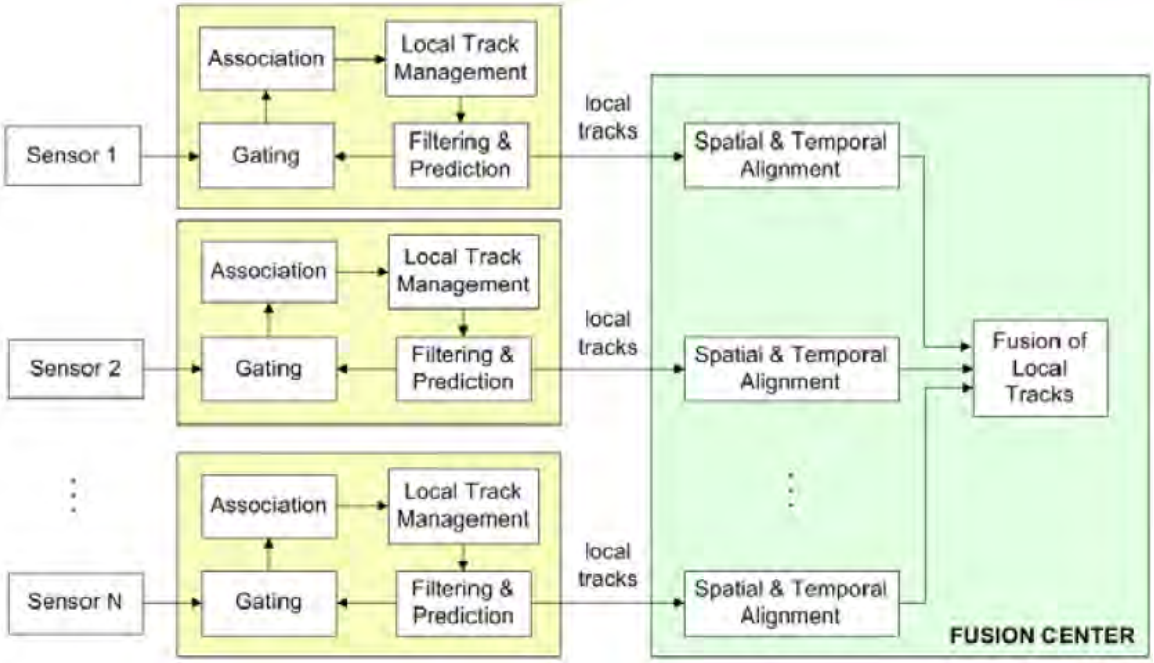

Figura 5. Architettura distribuita di un sistema di Fusione delle Informazioni

3.2.5.3 Architettura ibrida

Nei sistemi ibridi vi è un unico centro di fusione al quale i dati possono arrivare sia in forma grezza che elaborata da nodi periferici come nel caso delle architetture distribuite.
Gli svantaggi dell'architettura centralizzata (ad esempio eccessiva complessità nell'algoritmo di fusione) e dell'architettura distribuita (ad esempio ridondanza nelle informazioni) possono essere ridotti adottando una soluzione ibrida che ne metta a fattor comune i vantaggi [18]. La soluzione ibrida sembrerebbe ben coniugarsi con una soluzione basata su *network* (*net-centric environment*).

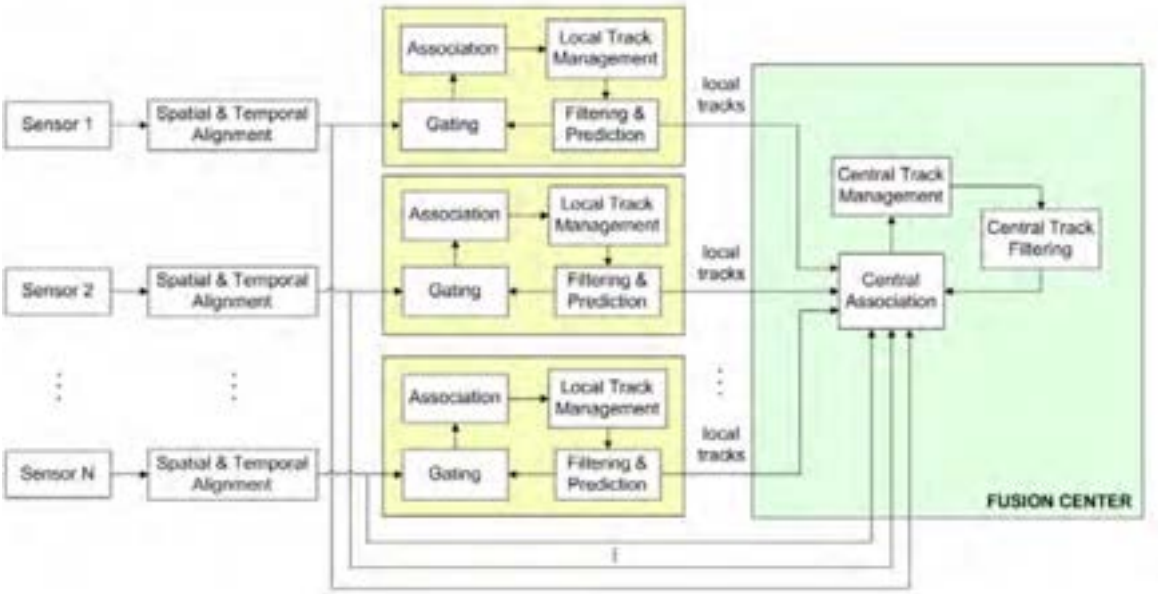

Figura 6. Architettura ibrida di un sistema di Fusione delle Informazioni

3.2.5.4 Architettura Gerarchica

In un'architettura gerarchica viene seguito un approccio *top-down* nei confronti della scomposizione dei *task* relativi al processo di fusione, producendo una riduzione del carico computazionale e dello scambio di informazioni tra le varie componenti.
Lo svantaggio principale di questa architettura è una scarsa flessibilità rispetto ad eventuali cambi di requisiti per il processo di fusione.

3.2.5.5 Architettura Olonica

L'architettura olonica si sta affermando negli ultimi anni e il sistema di fusione è costituito da una rete di piattaforme di sensori ciascuna con il proprio sistema di fusione (olone). Tali piattaforme interagiscono tra loro senza che vi sia un unico centro di fusione.
I singoli oloni sono capaci di organizzarsi in maniera autonoma in funzione dei cambiamenti che la rete di oloni può subire (proprietà di scalabilità e *self-organization*); ciascun olone è in grado di lavorare ai propri obiettivi cooperando con gli altri, garantendo minore vulnerabilità rispetto ai *failure*.
L'architettura olonica può anche essere indicata con il termine *decentralizzata* (DDF), in quanto non è presente un centro di fusione. In [19] viene descritto come i modelli grafici forniscono una interessante e accurata descrizione dei sistemi di DDF che portano ad una più profonda e più chiara comprensione delle operazioni all'interno di un sistema DDF.
Uno svantaggio legato a questo tipo di architettura è la possibilità di comportamenti non ben predicibili a priori, specie per sistemi complessi.

3.2.5.6 Architettura Market-Based

L'architettura *Market-Based* rappresenta un approccio interessante per la gestione dei sensori e del processo di fusione complessivo, che si ispira alle tecniche di *e-commerce*.
In [20] viene in particolare proposto un approccio di tipo *market-based* nella scelta dei *tradeoff* per la gestione dei sensori Figura 7. Questo approccio consente di produrre solo le informazioni rilevanti in funzione dello scopo e della situazione corrente. In questo modo si riduce di molto la possibilità di produrre più dati di quelli necessari, così come spesso accade nei sistemi tradizionali di fusione.
Gli autori in [20] hanno quindi sviluppato un *framework* che applica alla gestione di una rete di sorgenti le seguenti tecniche proprie di un approccio *market-based*:

- meccanismi di vendita all'asta per allineare lo *scheduling* con gli obiettivi della missione;
- tecniche di approssimazione per gestire la complessità computazionale;
- meccanismi di *pricing* per far rispettare vincoli sulle risorse;
- meccanismi di *surplus sharing* per ridurre l'impatto del comportamento strategico.

Tra gli svantaggi associabili alle architetture *market-based* vi è una potenziale elevata complessità legata al processo decisionale per la gestione delle risorse.

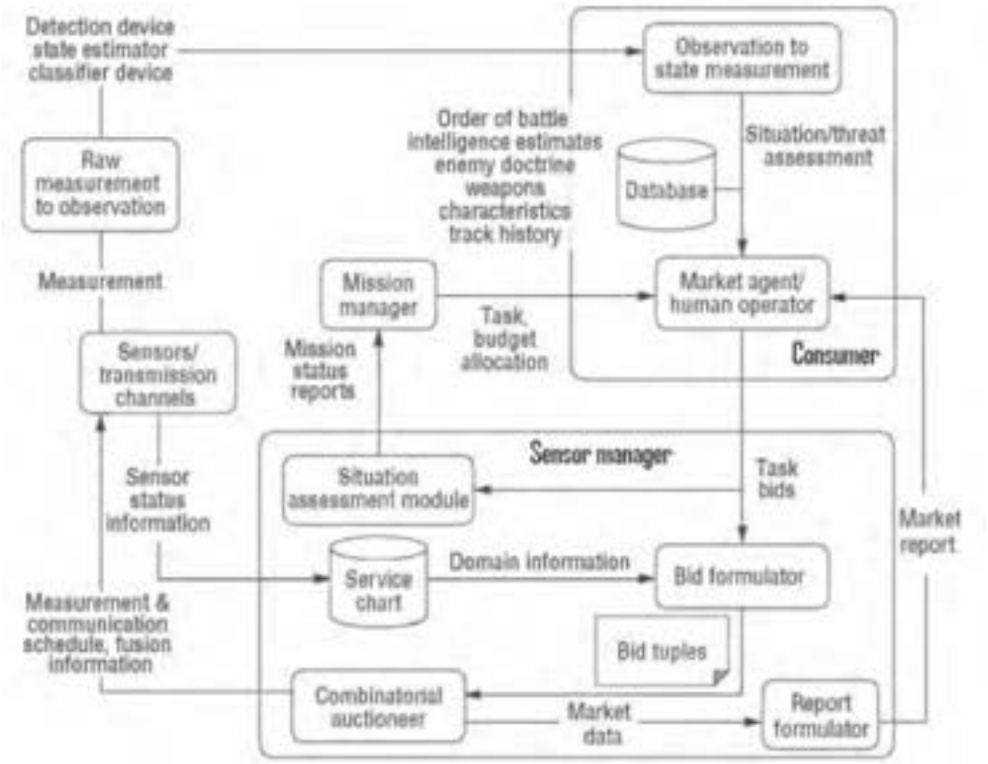

Figura 7. Market Architecture per la gestione dei diversi sensori

3.2.6 Il processo di fusione: tipologia di informazioni trattate

Il progetto descritto in [21] ha lo scopo di consentire la fusione di dati *hard* e *soft*, dati rilevati da sensori, prodotti dall'uomo e più vicini alle sue forme espressive. Tale lavoro prende soprattutto spunto dalle recenti esperienze di combattimento in guerra, a valle delle quali si cerca di mettere a fattor comune informazioni provenienti dai report testuali dei partecipanti alle azioni militari (*soft data*) ai dati provenienti da satelliti (*hard data*). In [21] si argomenta il fatto di non dover necessariamente fondere i dati il più vicino alla loro fonte all'interno dell'architettura del sistema. In alcuni casi è lecito asserire che più passaggi vengono fatti sui dati prima che questi vengano fusi e più si può perdere contenuto informativo. Questo non viene preso in particolar conto nel caso di fusione di dati soft e hard in quanto:

- la fusione di *hard data* è molto ben consolidata e quindi il contenuto informativo non viene perso;
- la fusione di *soft data* è stata invece molto meno affrontata in letteratura e ha al suo interno maggiore incertezza.

Per queste due ragioni fondere *hard* e *soft data* troppo presto può introdurre maggior rischio e si preferisce quindi separare i due processi prima di fondere le informazioni. Nell'ambito della *soft data fusion* si spinge sull'evoluzione di:

- *Graphical-Science-Based Methods;*
- *Fuzzy Divergence* e altre tecniche di analisi semantica;
- Metodi *Energy-logic Assessment Modeling.*

L'argomento della *Soft Data Fusion* viene trattato in [22], ponendo l'attenzione soprattutto su come i grafi siano strumenti essenziali per la rappresentazione di una *situation* all'interno dei livelli più alti dell'architettura, in quanto facilmente comprensibili per l'uomo; inoltre viene proposto un *framework* semantico per il *Soft Data Fusion*.
In [23] gli autori hanno proposto un lavoro sull'*Inexact Graph Matching* per la fusione *real time*, dove gli elementi informativi vengono estratti da testi scritti in linguaggio naturale e rappresentati come terne RDF (l'RDF è un modello basato su grafi proposto dal W3C).

3.2.7 Principali finalità del processo di fusione

In [24] gli autori hanno presentato una tassonomia per classificare i sistemi di *Information Fusion* in funzione di qual è lo scopo del processo di fusione.
Nel caso in cui l'obiettivo del processo di fusione è la produzione di osservazioni su un fenomeno più attendibili rispetto ai dati provenienti in ingresso dai sensori si parla di *Sensor Fusion.* Vi è quindi una forte analogia rispetto alla *Low-Level Data Fusion* e alla classificazione DAI-DAO.
Si indica invece con *Attribute Fusion* un sistema di fusione delle informazioni dove, a partire da misure di attributi, si vogliono produrre osservazioni più precise e complesse rispetto alle caratteristiche di un fenomeno; anche in questo caso vi sono delle forti analogie con altre classificazioni (*Medium-Level* e FEI-FEO).
Nella *Fusion across domains* la misura di un attributo viene effettuata da più sensori in maniera complementare rispetto all'intervallo misurabile. In questo caso si possono cogliere delle analogie con i sistemi di fusione complementari.
Nella *Fusion across times* si cerca di avere delle informazioni più accurate rispetto a un fenomeno andando a fondere informazioni in tempo reale con altre informazioni raccolte su base storica.

3.3 Modelli per la Fusione delle Informazioni

All'interno di questa sezione vengono discusse le caratteristiche dei principali modelli per l'*Information Fusion* e sono delineate le similitudini tra alcuni di essi. L'attenzione è stata posta esclusivamente sui principali modelli presenti in letteratura e non su una trattazione esaustiva dei modelli, poiché molto spesso si tratta di varianti atte alla soluzione di situazioni peculiari legate all'applicazione da realizzare.

In particolare, i modelli vengono classificati in quattro principali famiglie, andando ad estendere la classificazione utilizzata da in [3].
I modelli sono categorizzati, come illustrato Figura 8, in:

- Modelli basati sull'informazione;
- Modelli basati sulle regole;
- Modelli basati sulle attività;
- Modelli basati sulla *situational awareness.*

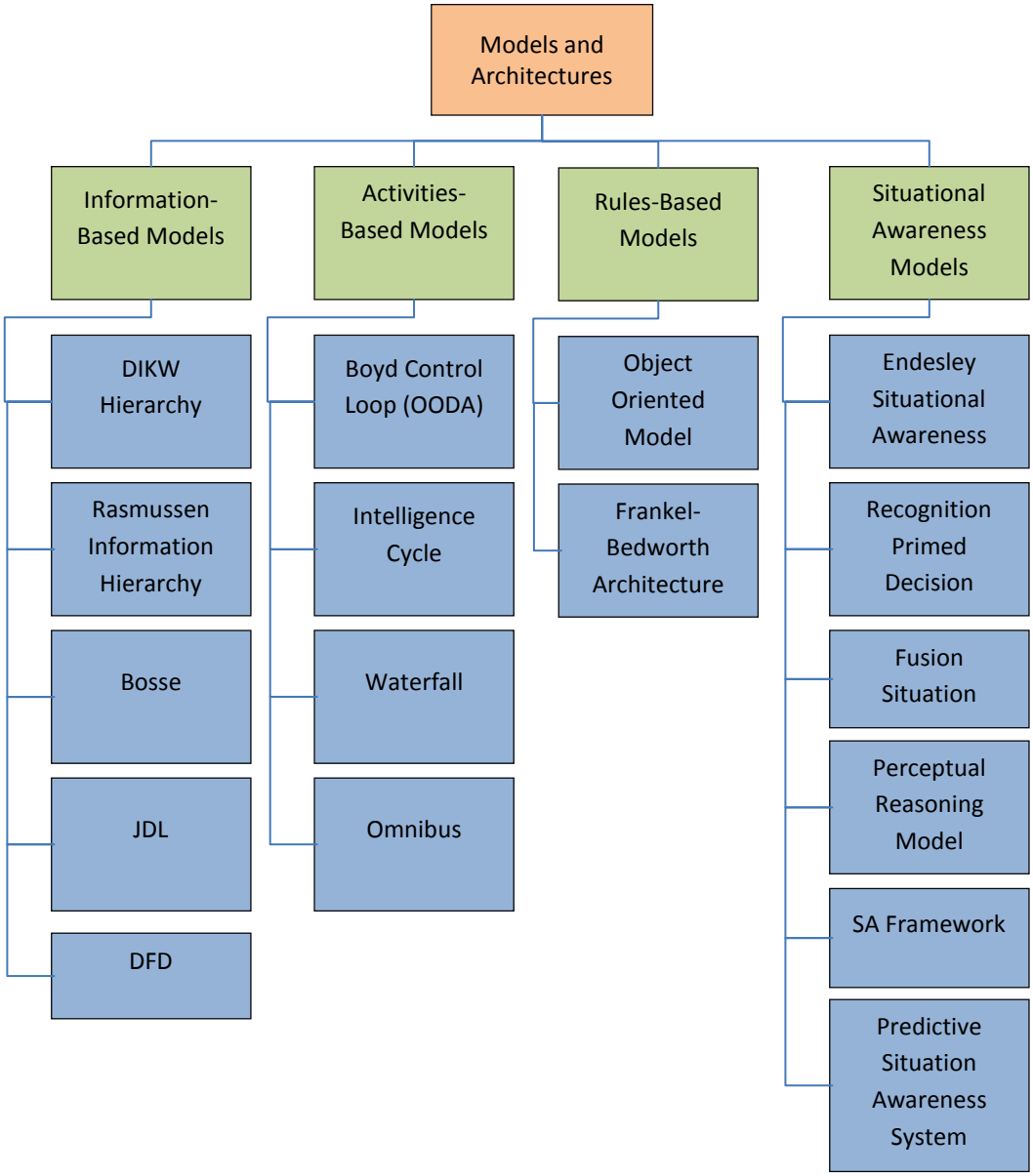

Figura 8. Modelli per la Fusione delle Informazioni

3.3.1 Modelli basati sull'informazione

La classificazione dei modelli basata sul livello di astrazione delle informazioni individua le seguenti categorie:

- DIKW *Hierarchy*;
- Rasmussen *Information Hierarchy*;
- Bosse *Model*;
- JDL *Model*;
- DFD *Mode*l.

Tali modelli sono trattati all'interno dei prossimi paragrafi e la loro analisi fornisce dei validi strumenti per la comprensione dei requisiti per la realizzazione di un sistema di *Information Fusion nel contesto delle smart city e più nello specifico del Sistema Elettrico fermo restante la validità generale e la ampia applicabilità anche ad ambiti diversi*.
Tra le categorie di modelli che si descrivono in questa sezione particolare importanza assume il JDL *model* in quanto è largamente trattato nel mondo scientifico ed inoltre è alla base di altri modelli.
Tuttavia i modelli basati sull'informazione (compreso quindi anche il JDL *model*) non forniscono alcuna indicazione in relazione alla sequenza di esecuzione della fusione.

3.3.1.1 Gerarchia DIKW

L'architettura gerarchica *Data Information Knowledge and Wisdom* [25] è immaginabile come una piramide composta da strati.
Alla base vi sono i dati che sono generalmente di basso livello e privi di elevato contenuto informativo.
Al crescere del livello di astrazione si attraversano il secondo strato, relativo alle informazioni, e il terzo strato, relativo alla conoscenza. Lo strato delle informazioni racchiude un'interpretazione semantica dei dati e le relazioni che intercorrono tra i dati grezzi e il contesto applicativo. Lo strato della conoscenza invece è costituito dalla consapevolezza riguardo allo scenario osservato in funzione delle esperienze pregresse, delle idee e dei principi (in genere soggettivi).
Infine il vertice è composto dalla saggezza, ovvero la cognizione di cosa sia giusto o meno.

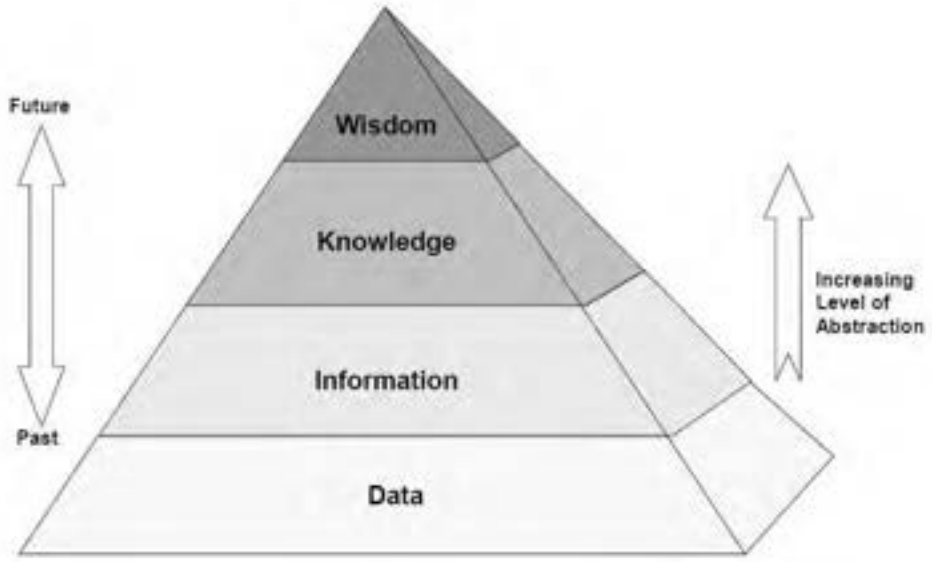

Figura 9. La piramide della conoscenza

I dati risultano essere quindi gli elementi fondanti delle informazioni che, a loro volta, costituiscono le unità base della conoscenza. La conoscenza è infine di fatto la base per la consapevolezza di ciò che riguarda l'ambiente in esame.

3.3.1.2 Gerarchia Rasmussen Information

Il modello del processo di elaborazione umana dell'informazione formulato da Rasmussen si articola su tre strati ordinati in base al livello di complessità dell'elaborazione. Alla base si trovano i processi basati sulle competenze dell'uomo e che vengono svolti in maniera automatica senza porvi grande attenzione. Lo strato intermedio è costituito dai processi che si fondano su regole ovvero su procedure ben stabilite mentre il livello di elaborazione più complesso è rivolto alla gestione di situazioni nuove e complesse.

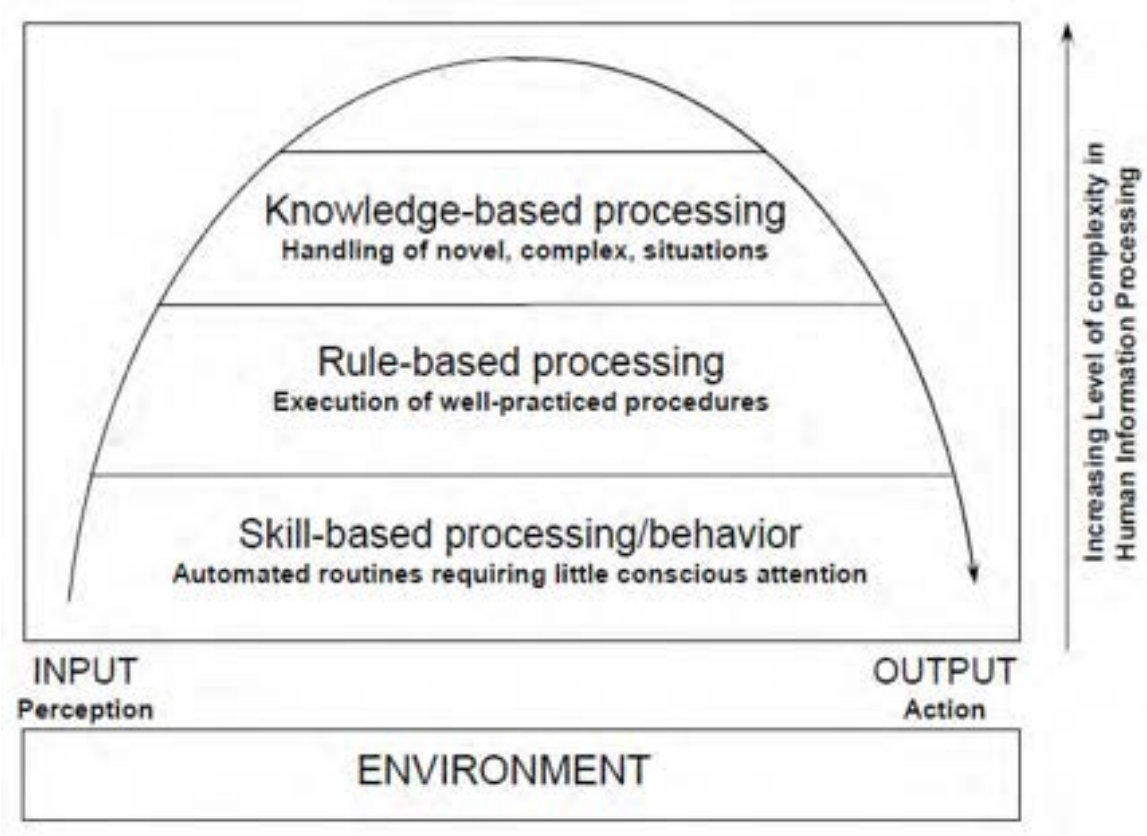

Figura 10. Il processo di comprensione umano secondo Rasmussen

3.3.1.3 Bosse Model

Un sistema di *Information Fusion* secondo il modello proposto da Bosse e altri (2006) è paragonabile ad una *box funzionale* che costituisce il vero motore del processo di fusione [26].
Gli ingressi sono ovviamente costituiti da dati, misure e informazioni rilevate dal campo. Il processo di fusione viene poi condizionato da segnali di controllo intesi come conoscenze pregresse, modelli e controlli nonché dai veri obiettivi che l'IFS si pone di raggiungere.
Gli output della *box funzionale* sono ulteriori segnali di controllo che possono agire ancora sul processo di fusione e le soluzioni che rispondono all'esigenza del sistema di fusione dell'informazione.
Il modello di Bosse mette in evidenzia la complessità di un sistema di *Information Fusion*, mostrando come le prestazioni del *fusion engine* non solo dipendono dagli ingressi, dalle uscite, dai controlli e dagli obiettivi ma anche fortemente dalla qualità degli ingressi.

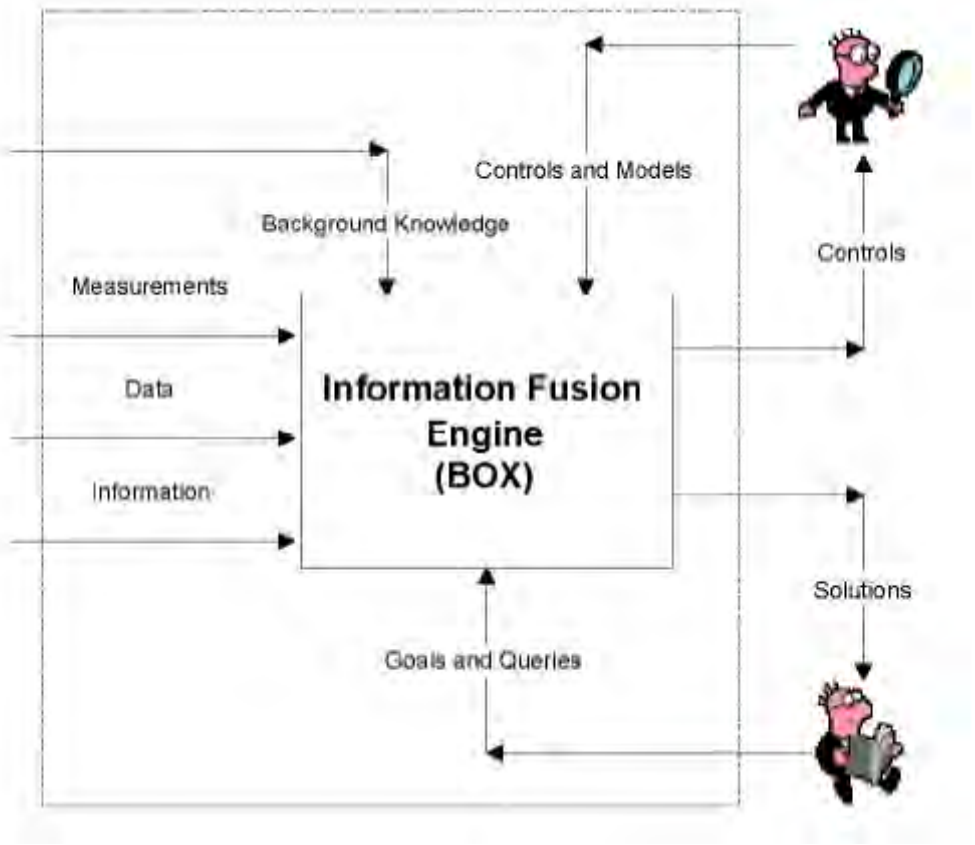

Figura 11. Rappresentazione di un IFS

I componenti alla base di un sistema di fusione delle informazioni secondo il modello di Bosse sono riassumibili negli insiemi riportati in Tabella 1.

Tabella 1. Insiemi di base del modello Bosse

Insieme	Descrizione
I	Costituisce l'insieme di tutte le possibili categorie di ingressi I_j
O	Rappresenta l'insieme di tutte le possibili categorie di uscite O_k
F	Include di tutte le funzioni di fusione f che potrebbero caratterizzare il motore di fusione
G	Rappresenta l'insieme di tutti i possibili obiettivi (*goals*) e/o quesiti

	(*queries*). G_{jk} è l'insieme degli obiettivi/quesiti istanziati per produrre l'uscita O_k a partire dall'ingresso I_j
C	Rappresenta l'insieme di tutti i possibili controlli inclusa la conoscenza in background. C_{jk} è l'insieme dei parametri di controllo, di modelli di fusione e di conoscenza in background istanziati per produrre l'output O_k a partire dall'input I_j

In Figura 12. Rappresentazione di una funzione di fusione sono schematizzate le relazioni che intercorrono tra gli elementi degli insiemi elencati in Tabella 1. In particolare, il fulcro di questa vista semplificata risulta essere la funzione di fusione *f*. Le funzioni di fusione possono essere espresse come:

$$f: \boldsymbol{I} \rightarrow \boldsymbol{O} / f(I_j)|(C_{jk} G_{jk}) = O_k$$

Inoltre, compare il termine *Q* che esprime una funzione di qualità che può essere applicata sia sugli ingressi che sulle uscite del sistema.

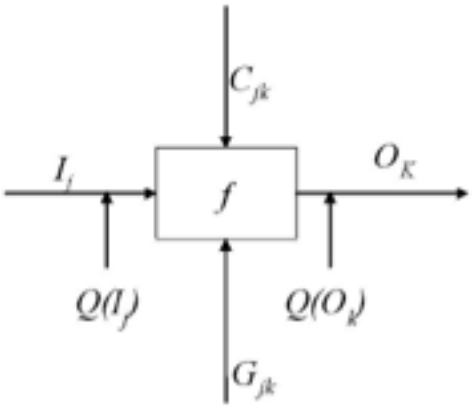

Figura 12. Rappresentazione di una funzione di fusione

Il problema di fusione può anche essere rappresentato all'interno di una matrice come una coppia ingresso-uscita (I_j, O_k) del tipo:

$$\begin{array}{c|c} & (O_k, G_k) \\ \hline (I_j, C_j) & \cdots \quad f(I_j)|(C_{jk}, G_{jk}) = O_k \quad \cdots \end{array}$$

Figura 13. Matrice di rappresentazione

3.3.1.4 Il modello JDL

Il *Joint Directors of Laboratories* (JDL) *Model* è tra i modelli più diffusi per la fusione delle informazioni [27], [25]. Il JDL *model* nasce all'interno delle attività del Dipartimento della Difesa degli Stati Uniti d'America nel 1985 e si è evoluto molto negli anni, dando spunto anche alla creazione di nuovi modelli.
Il JDL *model* ha lo scopo di [28]:

- fornire un punto di riferimento per discussioni relazionate all'*Information Fusion;*
- facilitare la comprensione dei tipi di problemi per i quali utilizzare tecniche di *Information Fusion*;
- standardizzare le caratteristiche dei problemi relazionati alla fusione;
- fornire un *framework* per automatizzare le soluzioni;
- classificare i differenti tipi di processi di fusione.

Gli elementi alla base del JDL *model* sono:

- *Sources*;
- *Database Management System*;
- *Human Computer Interface*;
- *Process Levels*.

In Figura 14 viene illustrato uno dei primi modelli JDL *model* realizzato da White nel 1988 [10].

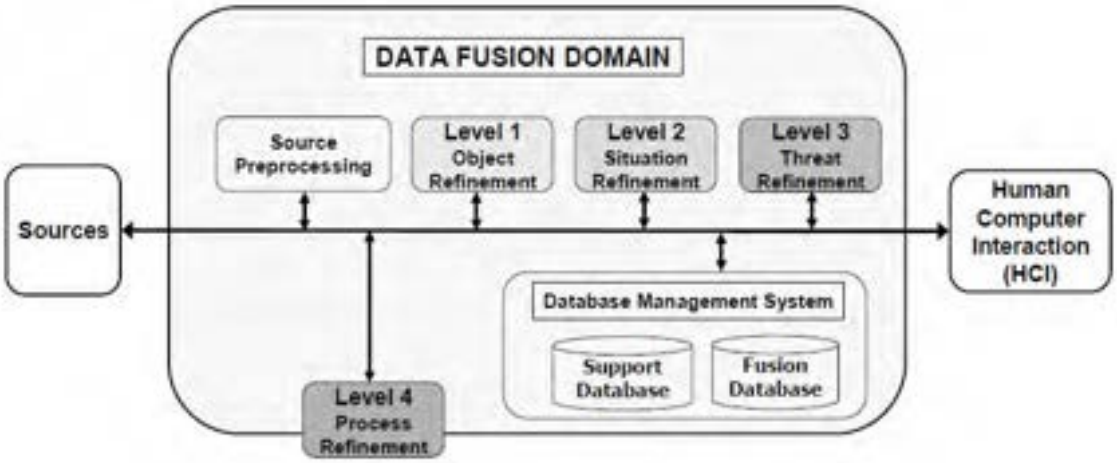

Figura 14. Il modello di fusione dei dati JDL (White, 1988)

Le sorgenti (*Sources*) forniscono ovviamente le informazioni da elaborare in ingresso al sistema di *Information Fusion*. Come è già stato detto in precedenza all'interno di questo documento le sorgenti possono essere di diverso tipo:

- dati provenienti da sensori;
- conoscenze pregresse;
- database;
- input forniti dall'utente.

L'elaborazione di queste informazioni è realizzata anche grazie all'ausilio di un *Database Management System* che consente ad esempio di memorizzare le conoscenze pregresse a supporto del processo di fusione.
Ovviamente il processo di fusione non può dirsi concluso senza dei moduli di interazione con l'esterno ed in particolare con l'utente. Questi moduli sono individuati dalla *Human Computer Interface* (HCI) o *Human Computer Interaction* come indicato in **Errore. L'origine riferimento non è stata trovata.**. L'HCI non solo consente di presentare il risultato del processo di fusione all'utente ma, allo stesso tempo, permette all'utente di fornire dei dati in ingresso.

Il processo di fusione viene poi suddiviso in livelli a seconda delle operazioni da compiere. Nelle diverse definizioni del JDL *model* il numero di livelli è passato dal valore iniziale di cinque fino ad arrivare, in una delle ultime versioni, a sette. I livelli del processo di fusione sono ovviamente comunicanti tra loro e relazionati opportunamente.

I livelli possono essere sintetizzati come in Tabella 2.

Tabella 2. I diversi livelli di processo nel modello JDL

Livello	Descrizione
Livello 0	Il Livello 0 si occupa del preprocessamento e della fusione di dati inerenti lo stato di un fenomeno a livello di segnale (basso livello di astrazione).
Livello 1	Stima e predice lo stato di entità sulla base di inferenze ricavate dalle osservazioni catturate. Inoltre si occupa di stimare l'identità di entità, dell'allineamento dei dati, della rilevazione di correlazione tra i dati, etc.
Livello 2	Stima e predice lo stato delle entità sulla base delle relazioni tra entità ed eventi.
Livello 3	Si occupa della valutazione degli effetti delle azioni pianificate in varie situazioni (simulazione, *multi-perspective assessment*, etc.).
Livello 4	Gestisce in maniera adattativa l'acquisizione dei dati e la loro elaborazione al fine di migliorare le performance del sistema di *Information Fusion*.
Livello 5	Viene indicato come *Cognitive Refinement* in quanto si concentra sul fattore umano nel sistema di fusione; in particolare determina in maniera adattativa chi richiede informazioni e chi può avere accesso alle informazioni e inoltre recupera e presenta le informazioni in maniera personalizzata, rispetto a chi le dovrà usare, al fine di supportare meglio il processo decisionale.

Livello 6	Gestisce in maniera adattativa la missione alla quale il processo di fusione è funzionale.

Come già detto, nelle prime versioni del modello JDL i livelli andavano dal Livello 0 al Livello 4 (così come in Tabella 2). La prima estensione del numero di livelli si è avuta nel 2000 ad opera di Hall e Tate [29] con l'aggiunta del concetto di *Cognitive Refinement*. All'interno del loro lavoro, il Livello 5 (appunto indicato come *Cognitive Refinement*) ebbe origine in quanto si avvertiva la necessità di concentrare l'attenzione della ricerca non solo unicamente a livello di *data processing*. Infatti, fino ad allora, la concentrazione era molto incentrata sui dati provenienti dai sensori e sulla loro rappresentazione, mentre si era trascurato il ciclo di supporto alle decisioni.
Blasch e Plano (2002) [30] sottolinearono anch'essi che i sistemi di fusione delle informazioni fino ad allora avevano trattato esclusivamente l'elaborazione automatica dei dati, senza effettuare alcuna considerazione sul ruolo dell'utente all'interno del processo di fusione. Blasch e Plano rinominarono il Livello 5 in *User Refinement*, considerandolo come una determinazione adattativa delle caratteristiche di chi richiede informazioni e di chi ha accesso a queste. Inoltre, secondo gli autori, questo livello poteva essere visto anche come il recupero e la presentazione automatica e adattativa dei dati al fine di supportare le decisioni da parte degli utenti.
Successivamente anche McMullen e Hall (2004) [31] trattarono il Livello 5 all'interno del proprio lavoro e in Figura 16 viene riportata la decomposizione grafica del JDL Model basata sulla tassonomia di Hall e McMullen del 2004 [31] relativa ai cinque livelli (esclusi il livello 0 e il livello 6) del modello.
Il modello JDL secondo il lavoro di McMullen e Hall è illustrato in Figura 15; in particolare, in figura viene mostrato come il modello JDL proposto affronta la problematica della sicurezza in rete, evidenziando come le singole componenti del modello contribuiscono al processo di fusione.

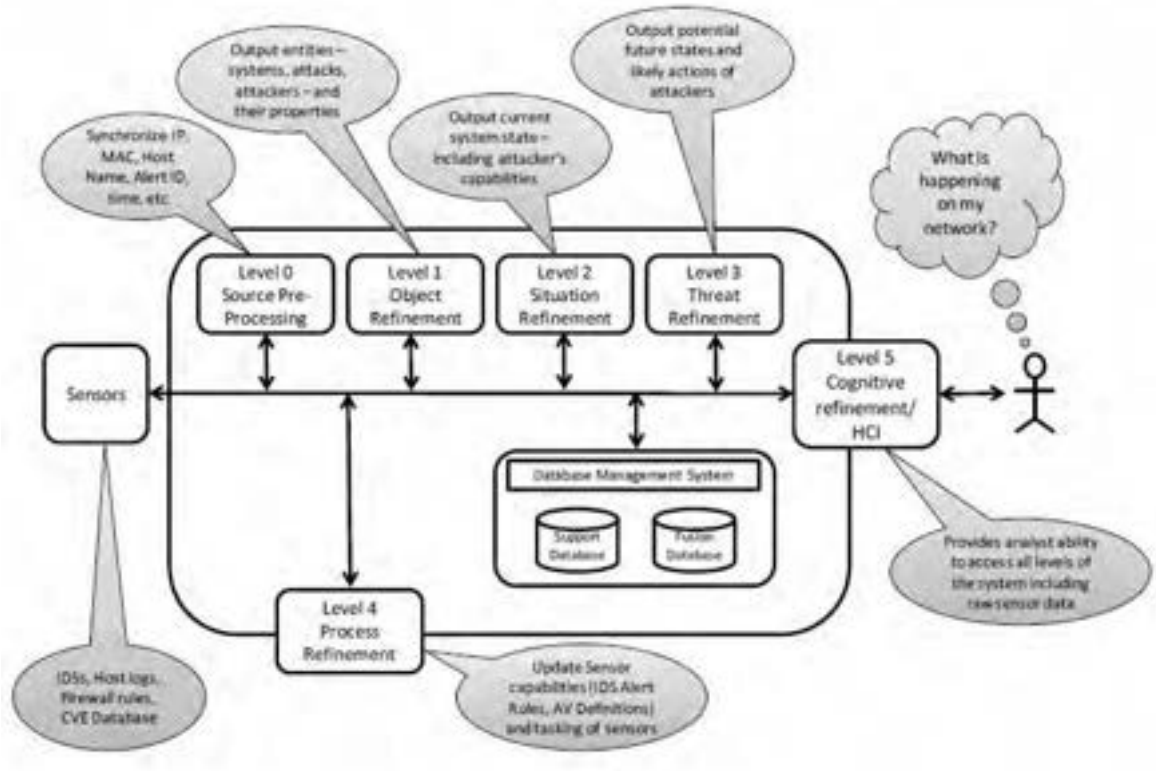

Figura 15. Revisione del modello JDl ad opera di Hall e McMullen nel contesto della cyber security

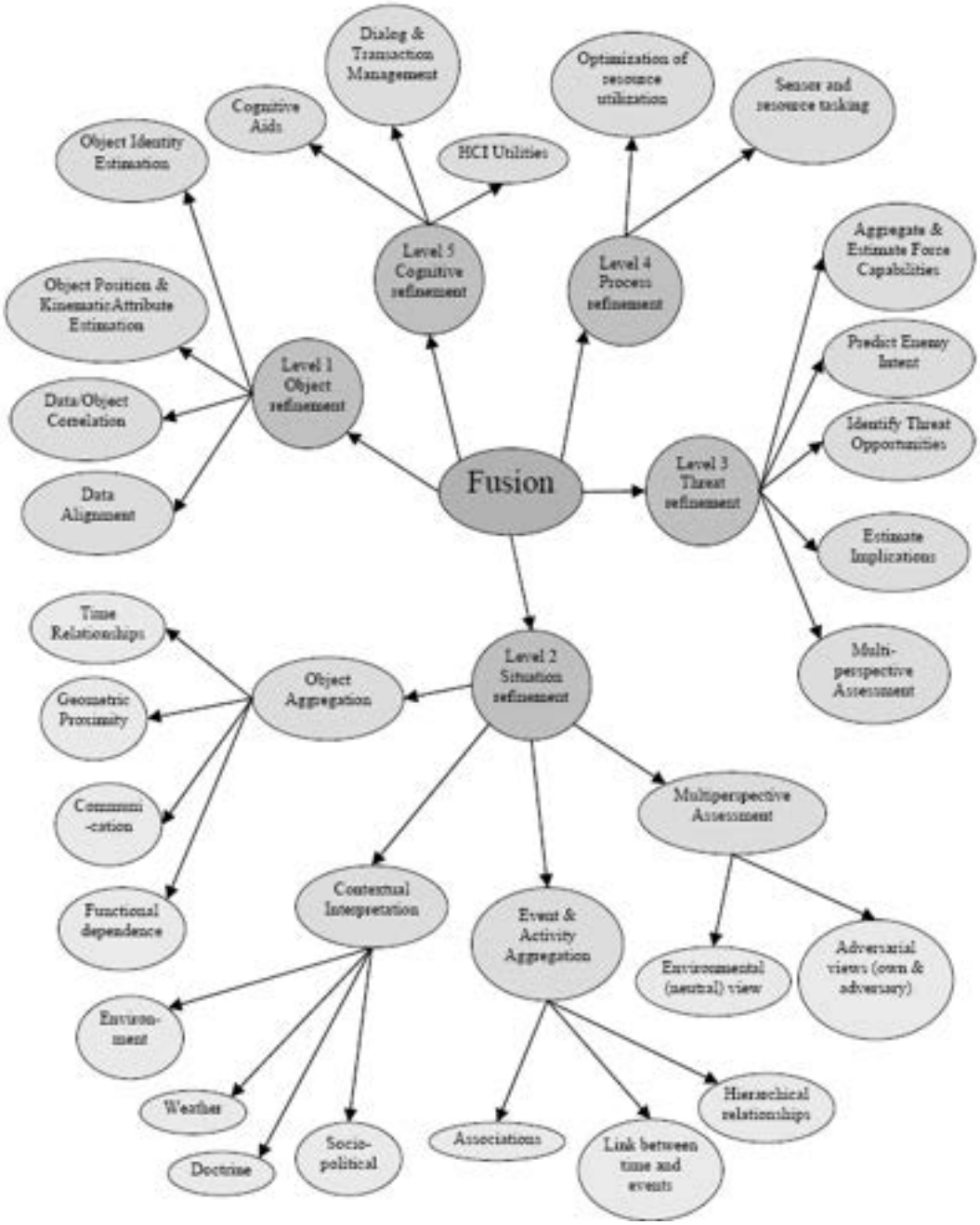

Figura 16. Tassonomia di Hall e McMullen

Blasch in un suo lavoro del 2006 [32] fornisce una ulteriore vista del JDL indicandolo come *User Fusion Model*, così come illustrato in Figura 17.

Ad i primi 6 livelli il *Data Fusion Information Group* (DFIG) ha affiancato infine un ulteriore livello, indicato come *Mission Management* [33], [34]. Questo livello (*Level 6*) nasce dall'esigenza di gestire le attività e gli obiettivi in maniera adattativa sia nel tempo che nello spazio al fine di supportare i team nel prendere delle decisioni sulla base di vincoli sociali, economici e politici. Il lavoro del DFIG è illustrato in Figura 18.

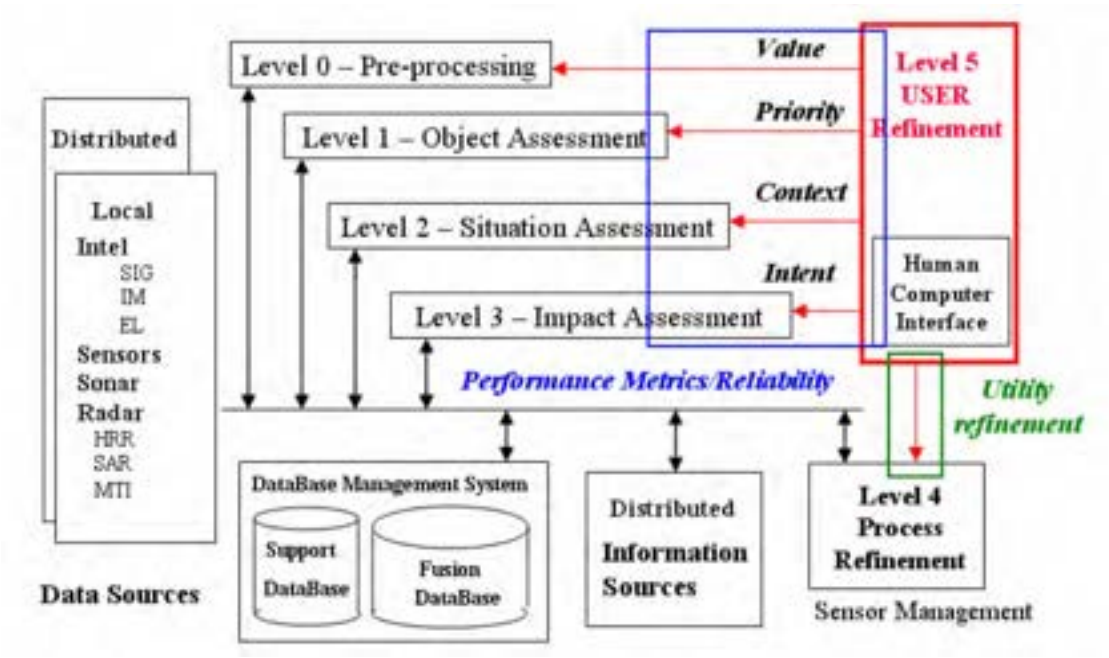

Figura 17: User Fusion Model

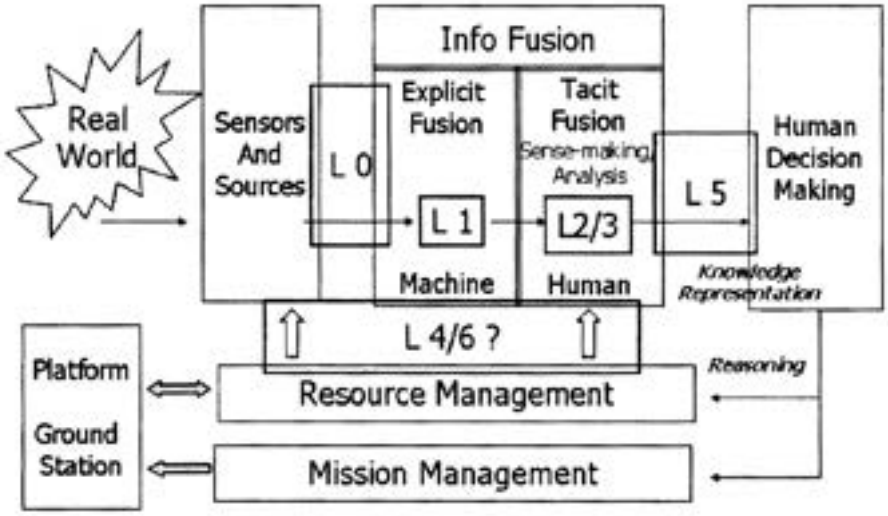

Figura 18. Modello DFIG

All'interno del modello DFIG è evidenziato anche il blocco che si occupa della gestione delle risorse all'interno del sistema di fusione. Un esempio di tali relazioni è illustrato in Figura 19 [27].

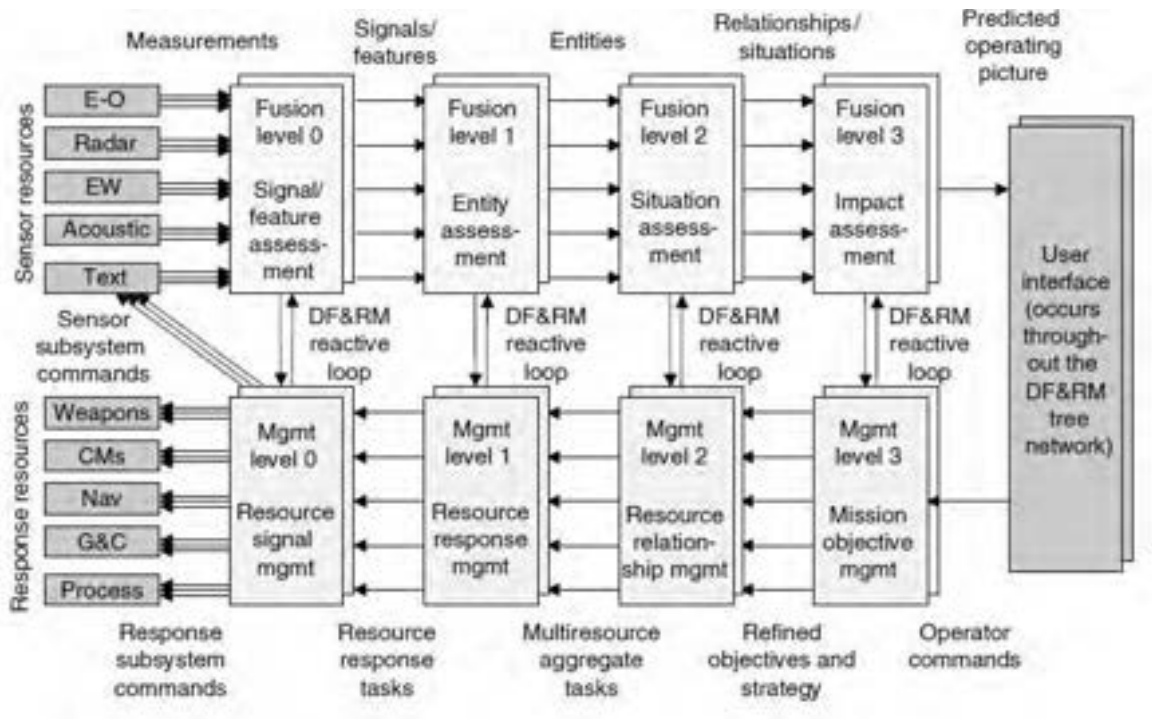

Figura 19. Relazioni esistenti tra sistema di Fusione e sistema di Gestione delle Risorse

In un panel report di marzo 2008 [35], Blasch, Salerno ed altri presentano un'ulteriore vista del JDL Model così come riportato in Figura 21.

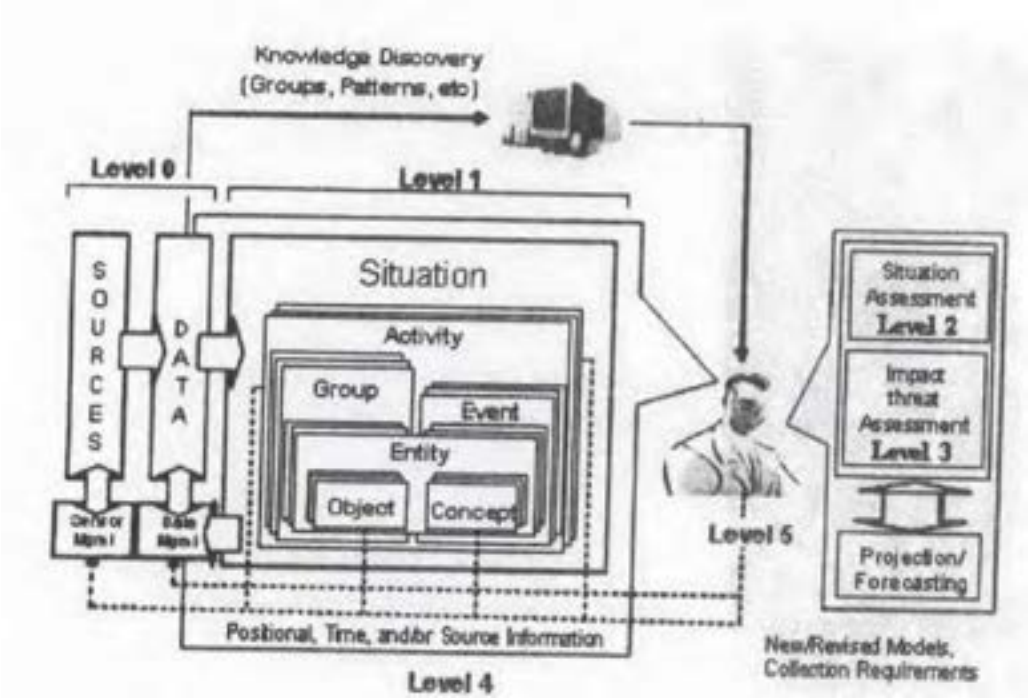

Figura 20. Il modello JDL secondo Blasch et al.

Le informazioni vengono scambiate tra i livelli all'interno del processo di elaborazione così come illustrato in Figura 21 [27].

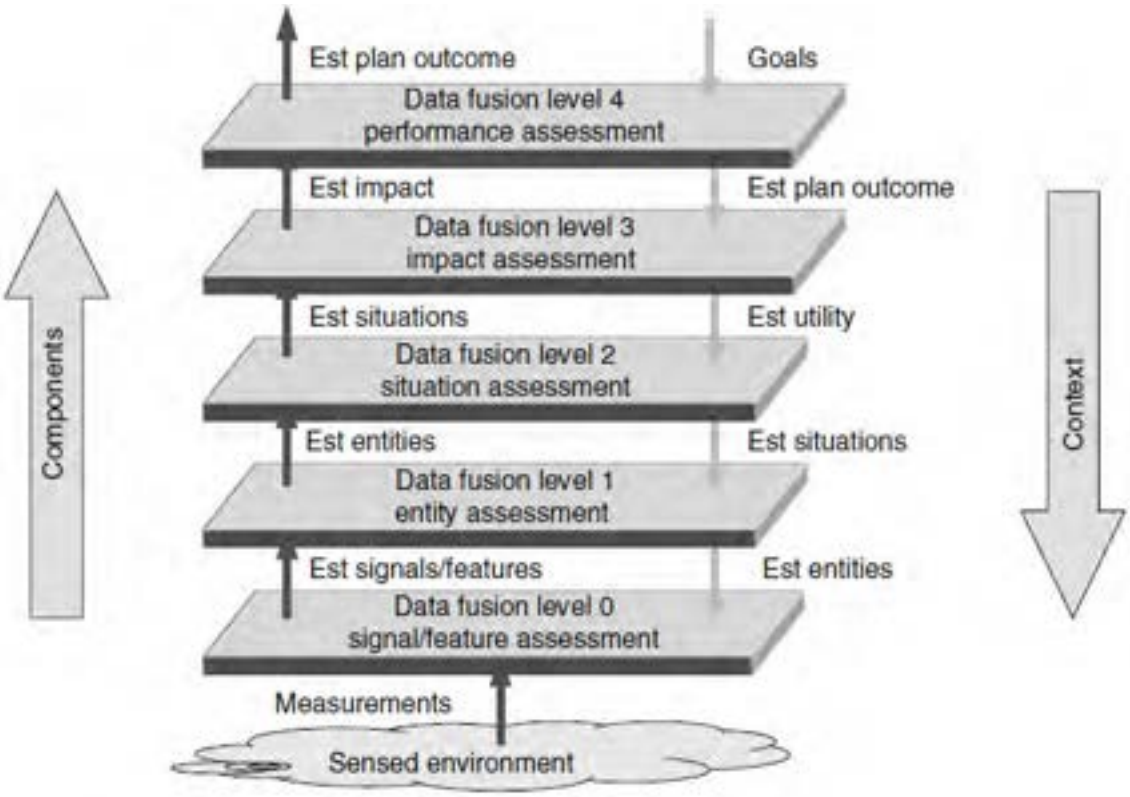

Figura 21. Flusso delle informazioni tra i vari livelli del modello JDL

3.3.1.5 DFD Model

Il modello *Data-Feature-Decision* (DFD) è strutturato su tre livelli principali di astrazione [36]. Secondo tale modello in ingresso si ha un insieme di dati, questi vengono poi trasformati in *feature* che a loro volta vengono fuse al fine di determinare il vero output del sistema di *Information Fusion*, cioè delle decisioni.

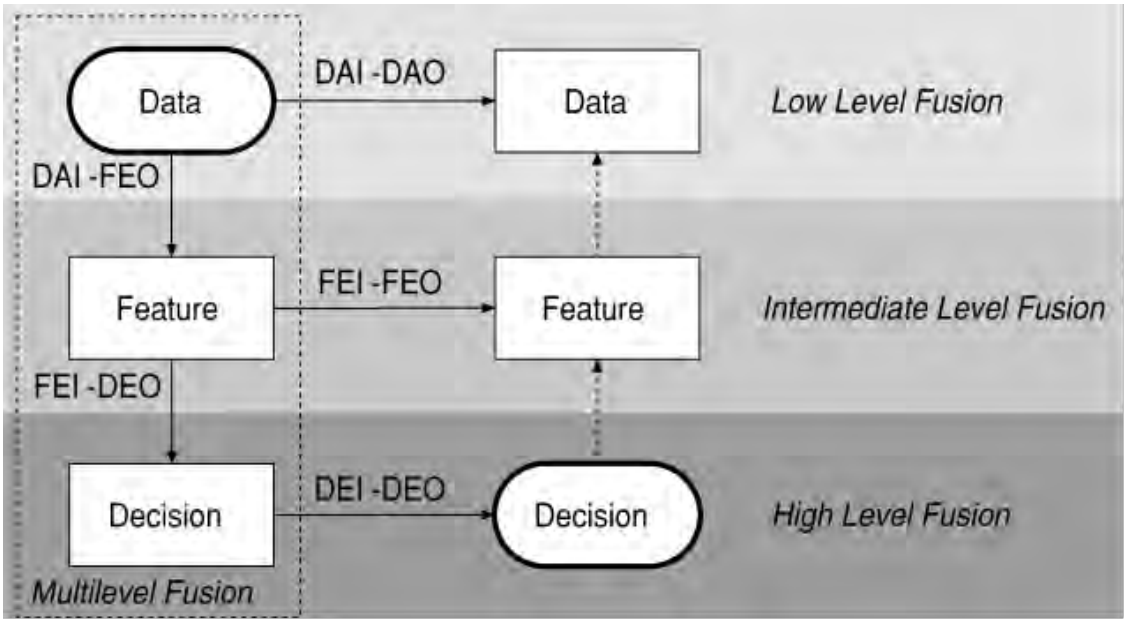

Figura 22. Modello DFD

Il modello DFD può anche essere visto come nella seguente versione schematizzata (Figura 23), che mappa il DFD con i primi due livelli del modello JDL [27].

Input \ Output features	Data	Objects
Data	Signal detection DAI/DAO	Feature extraction DAI/FEO	Gestalt-based object characterization DAI/DEO
Features	Model-based detection feature extraction FEI/DAO	Feature refinement FEI/FEO	(Feature-based object characterization) FEI/DEO
Objects	Model-based detection estimation DEI/DAO	Model-based feature extract DEI/FEO	Object refinement DEI/FEO

Level 0 — Level 1

Figura 23. Mapping del modello DFD con i primi due livelli del modello JDL

Nel 2001 Steinberg e Bowman modificarono la mappatura illustrata in Figura 23 al fine di includere anche i livelli 2, 3 e 4 del modello JDL (Figura 24) [37].

Input \ Output	Data	Features	Objects	Relations	Impacts	Responses
Data	Signal detection DAI/DAO	Feature extraction DAI/FEO	Gestalt-based object extract DAI/DEO	Gestalt-based situation assessment DAI/RLO	Gestalt-based impact assessment DAI/IMO	Reflexive response DAI/RSO
Features	Model-based detection/ feature extraction FEI/DAO	Feature refinement FEI/FEO	Object characterization FEI/DEO	Feature-based situation assessment FEI/RLO	Feature-based impact assessment FEI/IMO	Feature-based response FEI/RSO
Objects	Model-based detection/ estimation DEI/DAO	Model-based feature extraction DEI/FEO	Object refinement DEI/DEO	Entity-relations situation assessment DEI/RLO	Entity-based impact assessment DEI/IMO	Entity-relation based response DEI/RSO
Relations	Context-sensitive detection/est RLI/DAO	Context-sensitive feature extraction RLI/FEO	Context-sensitive object refinement RLI/DEO	Micro-macro situation assessment RLI/RLO	Context-sensitive impact assessment RLI/IMO	Context-sensitive response RLI/RSO
Impacts	Cost-sensitive detection/est IMI/DAO	Cost-sensitive feature extraction IMI/FEO	Cost-sensitive object refinement IMI/DEO	Cost-sensitive situation assessment IMI/RLO	Cost-sensitive impact assessment IMI/RLO	Cost-sensitive response IMI/RSO
Responses	Reaction-sensitive detection/est DEI/DAO	Reaction-sensitive feature extraction RSI/FEO	Reaction-sensitive object refinement RSI/DEO	Reaction-sensitive sit assessment RSI/RLO	Reaction-sensitive impact assessment RSI/RLO	Reaction-sensitive response RSI/RSO

Level 0 — Level 1 — Level 2 — Level 3 — Level 4

Figura 24. Mapping del modello DFD con i primi cinque livello del modello JDL

3.3.2 Modelli basati sulle attività

Un ulteriore modo di classificare i sistemi di fusione dell'informazione consiste nell'individuare delle categorie in funzione delle attività che devono essere eseguite all'interno del processo di fusione [3]. Si possono individuare quindi i seguenti modelli che appartengono a questa famiglia:

- *Intelligent Cycle,*
- *Boyd Control Loop* (oppure OODA),
- *Waterfall Model,*
- *Omnibus Model.*

All'interno delle prossime sezioni vengono presentate le caratteristiche di tali modelli e delle attività che sono in essi incluse.
In particolare, i primi due modelli sono generali e possono essere impiegati in applicazioni appartenenti a qualsiasi dominio, di conseguenza non soddisfano ne specificano tutti gli aspetti relativi alla fusione, in particolare non specificano nei dettagli i vari task da eseguire.

L'ultimo modello, l'*Omnibus Model*, che potrebbe essere visto come un'integrazione dei modelli *Waterfall* ed OODA, si preoccupa di dettagliare i *task* di fusione che devono essere eseguiti ad ogni fase.

3.3.2.1 Intelligent Cycle model

L'*Intelligent Cycle* (IC) è rappresentato in Figura 25 [38], [39]. La prima fase, *Collection*, è di raccolta delle informazioni che poi vengono inviate alla fase *Collation* che le analizza, correla e compara. Il risultato di questa fase va poi in ingresso alla fase di *Evaluation* che fonde le informazioni ed il risultato di questo processo viene distribuito agli utenti nella fase di *Dissemination.*

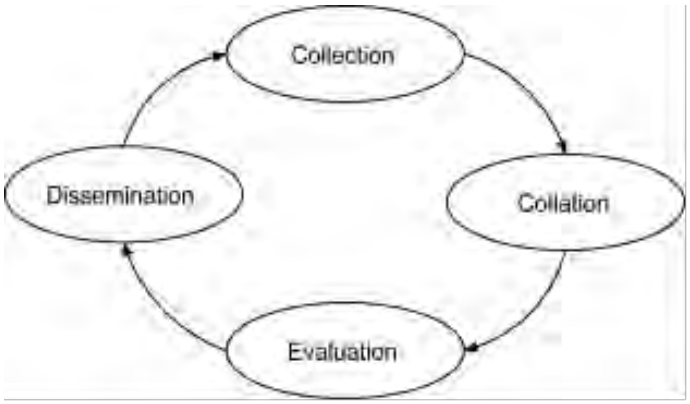

Figura 25. Intelligent Cycle Model

In Tabella 3 si riportano in maniera schematica le similitudini tra i livelli del JDL *model* con quelle che sono le attività caratteristiche dell'*Intelligent Cycle Model*.

Tabella 3. Corrispondenza tra i modelli IC e JDL

Attività del modello IC	Livello del modello JDL corrispondente
Collection	Livello 0
Collation	Livello 1
Evaluation	Livello 2, Livello 3

3.3.2.2 OODA model

Il *Boyd Control Loop* [40] viene anche indicato con l'acronimo OODA, il quale ha origine dalle quattro fasi che compongono il processo di fusione secondo questo modello: *Observe*, *Orient*, *Decide*, *Act*. Tale modello è di fatto molto simile all'*Intelligent Cycle Model*.
L'OODA *model* è un modello ciclico, in cui ciascuna fase accetta in ingresso il risultato della fase precedente ed invia alla fase successiva il prodotto delle azioni svolte al suo interno.
La prima fase è *Observe* che ha lo scopo di raccogliere le informazioni sullo scenario dalle sorgenti.
La seconda fase è *Orient* che interpreta le informazioni raccolte, fornendo alla fase di *Decide* gli strumenti per effettuare delle scelte sulle azioni da compiere.
L'ultima fase è *Act*, che realizza quanto deciso nella precedente fase.

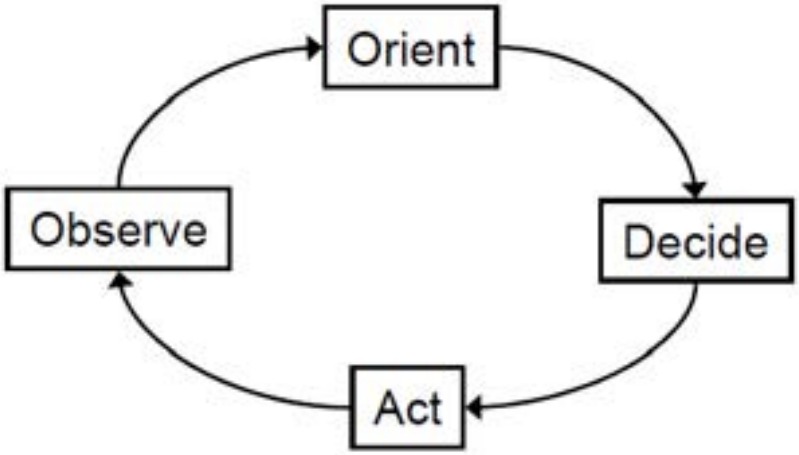

Figura 26. Body Control Loop OODA

Al fine di migliorare tale modello (Figura 27) Brehmer ha sviluppato il DOODA (*Dynamic*-OODA) Loop riportato in Figura 28, utilizzando modelli cibernetici per il comando e controllo [41].

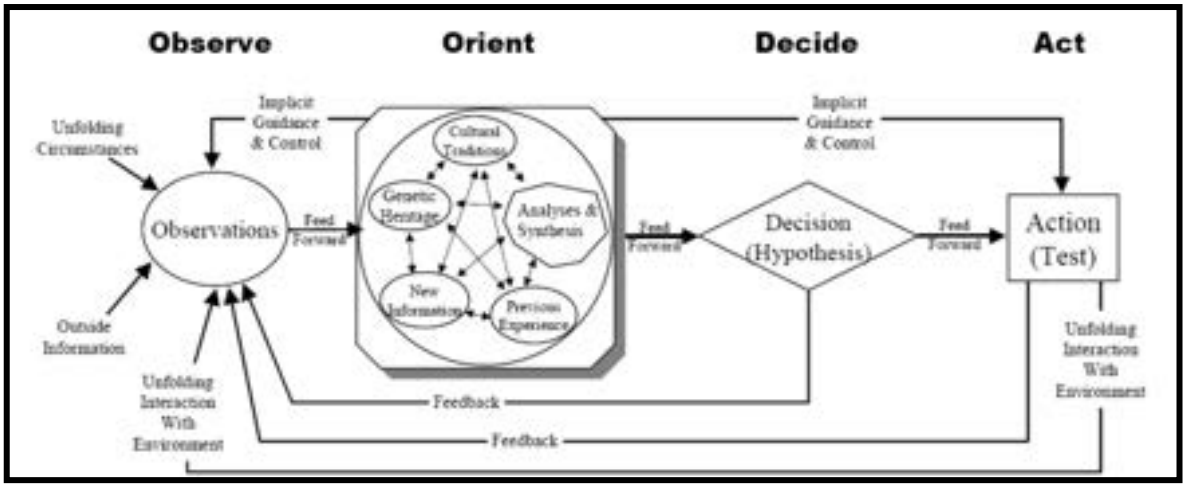

Figura 27. Modello OODA-Loop rivisto da Brehmer

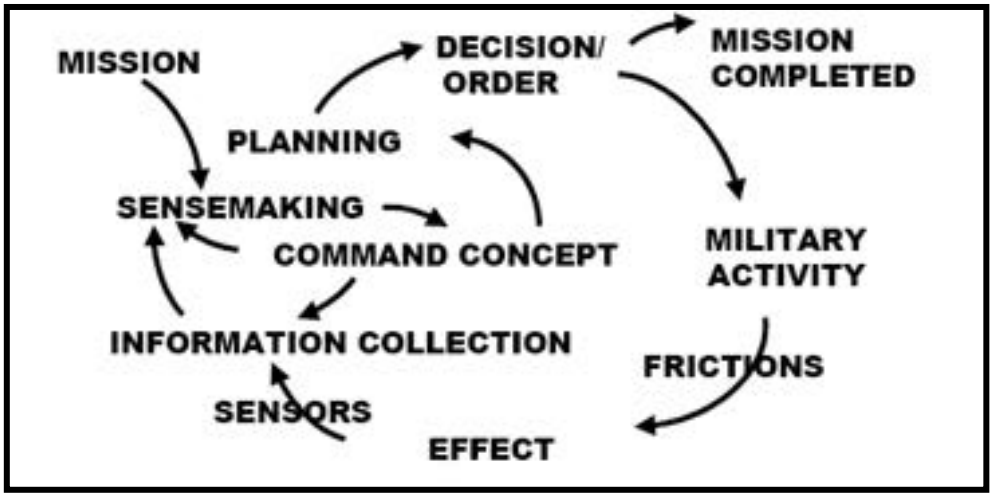

Figura 28. Modello DOODA-Loop di Brehmer

In Tabella 4**Errore. L'origine riferimento non è stata trovata.** si riportano le similitudini (descritte da Bedworth ed O'Brien in [38]) tra l'OODA Loop ed il modello JDL model e il modello *Intelligence Cycle*.

Tabella 4. Corrispondenze tra i modelli OODA, JDL e IC.

Attività del modello OODA	Parallelo con il JDL model e l'IC model
Observe	*Observe* è ampiamente comparabile sia al livello 0 del JDL model che a parte della fase di Collection dell'Intelligence Cycle.
Orient	Orient racchiude i livelli 1, 2, e 3 del JDL model ed include gli elementi strutturati delle fasi di *Collection* e *Collation* dell'Intelligence Cycle.
Decide	Decide include il livello 4 del JDL model e le attività della fase di *Dissemination* dell'Intelligence Cycle.
Act	*Act* non ha alcuna corrispondenza con i livelli o le fasi che caratterizzano i modelli in oggetto.

3.3.2.3 Architettura a cascata (waterfall)

L'architettura *a cascata* Figura 29, è un modello gerarchico in cui vi è una prima fase di rilevazione delle informazioni (*Sensing* e *Signal Processing*). Successivamente vi è una fase di estrazione delle caratteristiche ed elaborazione dei *pattern* che consentono alla fase successiva di *Situation Assessment* di valutare lo stato del sistema in modo da supportare il processo decisionale finale.

Figura 29. Modello architetturale a cascata

Un modello *a cascata* rivisitato (Figura 30) prevede che si porti in ingresso al sistema un feedback sulle decisioni prese nello stadio finale [42], [43].

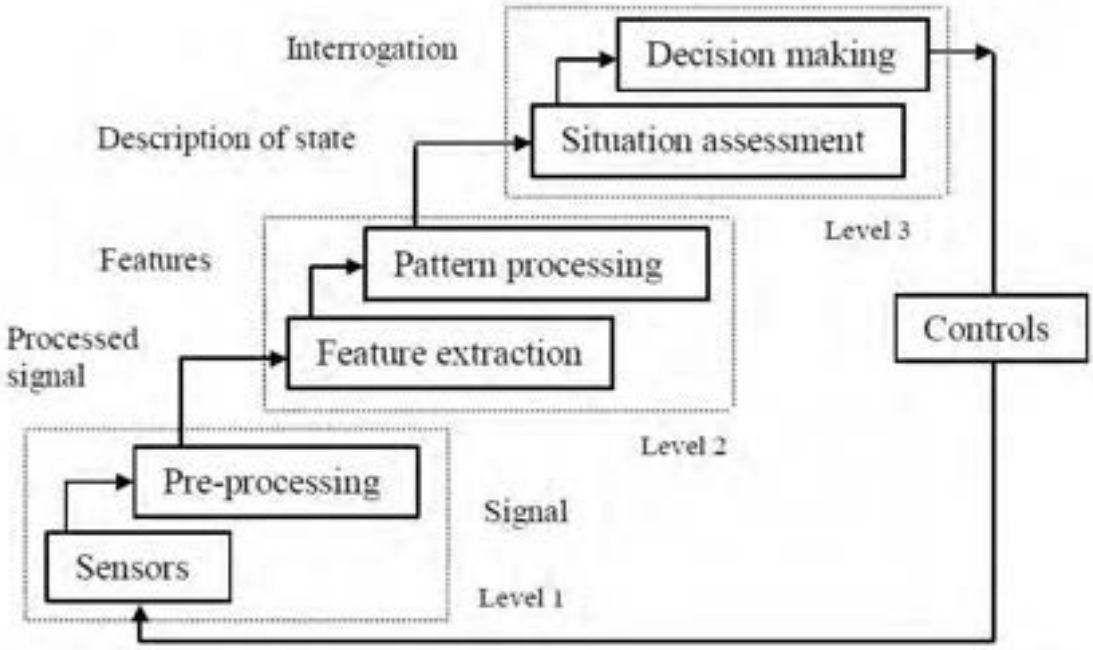

Figura 30. Modello di architettura a cascata con feedback

3.3.2.4 Modello Omnibus

Il modello Omnibus cerca di includere al suo interno i vantaggi degli altri modelli (OODA, *Waterfall*) e superarne allo stesso tempo le limitazioni [38].

Come gli altri modelli basati sulle attività, l'*Omnibus* si compone di diverse fasi:

- La fase di *Observe* include gli stadi di *Information Gathering* e *Information Preprocessing*, che hanno appunto l'obiettivo di raccogliere dati e preprocessarli, producendo così i *Raw Data Fusion.*
- La fase di *Feature Fusion* include la *Feature Extraction* e il *Pattern Processing* dando origine alla *Soft Decision Fusion*. Essa estrae dai dati grezzi, prodotti nella precedente attività, gli attributi salienti in funzione dei propri obiettivi.
- In *Decide* avviene il *Context Preprocessing* e la *Decision Making* ovvero vengono prese le decisioni da attuare a seconda del contesto in cui il sistema ipotizza di essere.
- Il *loop* viene chiuso con la fase *Act* che attua le decisioni prese nel precedente stadio e intraprende determinate azioni, il cui esito si riflette ovviamente in un'interazione con l'ambiente circostante i cui esiti vengono rilevati nella fase di *Observe.*

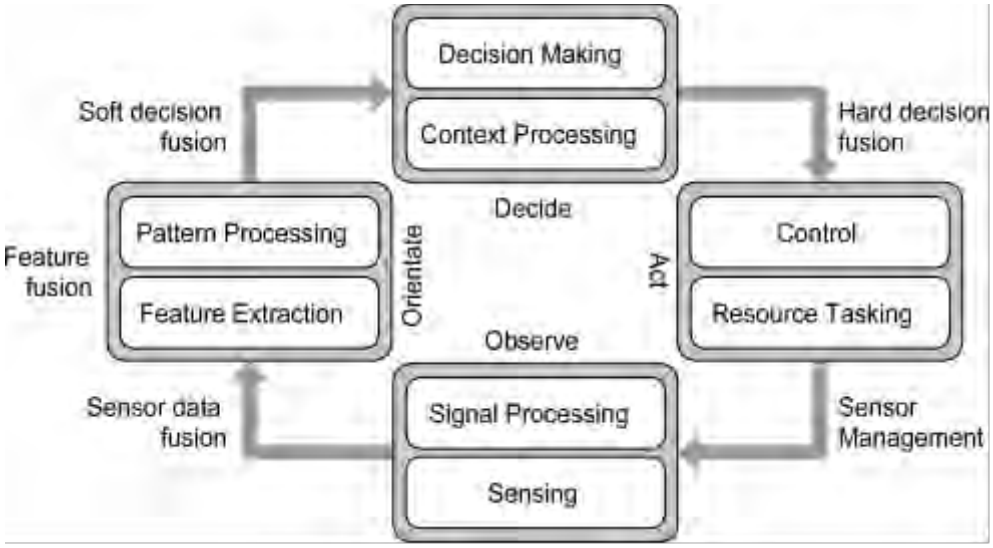

Figura 31. Il modello Omnibus

Come si può osservare dalla Figura 31 le informazioni trattate da tale modello sono unicamente quelle derivate dai sensori, ma con alcune piccole modifiche suggerite da Nakamura in [3] tale modello è stato generalizzato in maniera da risultare più appropriato per essere di riferimento nella progettazione di un IFS che tratta qualsiasi tipo di sorgente (inclusi i sensori).
In particolare:

- *Sensing* e *Signal Processing* sono stati rimpiazzati rispettivamente da *Information Gathering* ed *Information Preprocessing*;
- *Sensor Data Fusion* è stato rimpiazzato da *Raw Data Fusion*;
- *Sensor Management* è stato rimpiazzato da *Source Management.*

3.3.3 Modelli basati sulle regole

Un'ulteriore famiglia di modelli per la fusione delle informazioni comprende i modelli basate sulle regole. Questa famiglia consente appunto di classificare i modelli per l'*Information Fusion* in funzione non delle attività del processo di fusione, ma in funzione delle regole che legano le varie fasi, individuando così gli attori del processo di fusione ed il loro ruolo al suo interno.
A tale categoria appartengono i seguenti modelli:

- *Object Oriented Model*;
- Frankel-Bedworth *Architecture*.

Tali modelli consentono di comprendere il processo di fusione anche se non forniscono in maniera esplicita molti utili particolari che risultano invece evidenti nelle altre categorie di modelli.

3.3.3.1 Object Oriented Model

L'*Object Oriented Model* [44] si basa sulle relazioni tra quattro insiemi di regole ed ha una struttura ciclica, simile a quanto visto ad esempio nell'OODA *model*.
Gli attori del processo di fusione sono riassunti in Tabella 5.

Tabella 5. Attori e ruoli nel processo di fusione delle informazioni

Attore del processo di fusione	Ruolo
Actor	L'*Actor* interagisce con il mondo esterno sia effettuando misure sullo stato dell'ambiente che andando ad agire su di esso.
Perceiver	Il *Perceiver* valuta le informazioni raccolte dall'*Actor* e le organizza in maniera contestualizzata per fornirle al *Director.*
Director	Il *Director* realizza un piano d'azione che il *Manager* si occupa di mettere in opera.
Manager	Il *Manager* si occupa di supervisionare l'esecuzione del piano da parte dell'*Actor*.

In Figura 32 vengono illustrati gli attori del processo di fusione, le azioni che compiono e le loro relazioni.

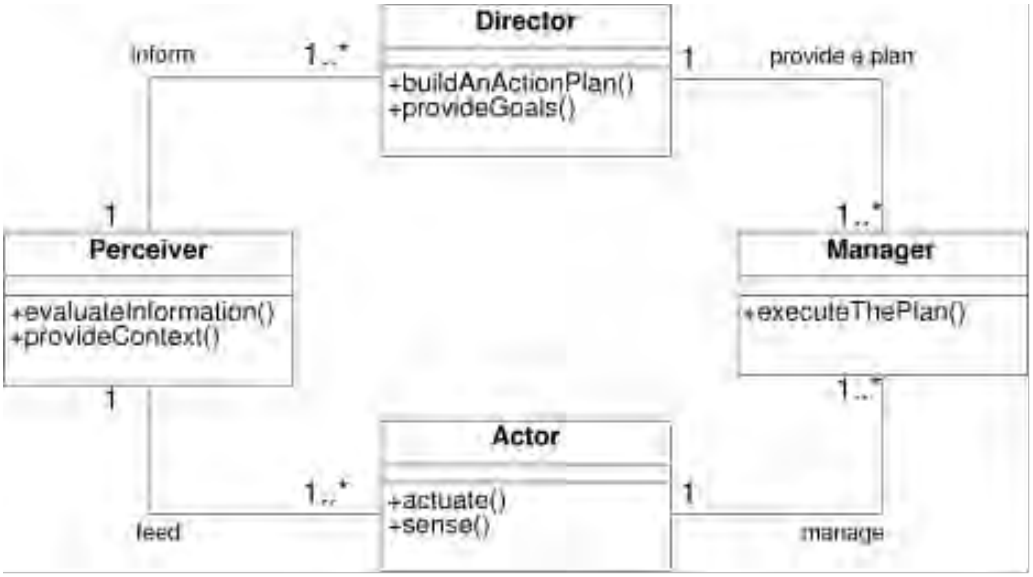

Figura 32. Modello orientato agli oggetti

Un'esplosione dell'*Object Oriented Model* di Figura 32 è mostrato in Figura 33([45]).

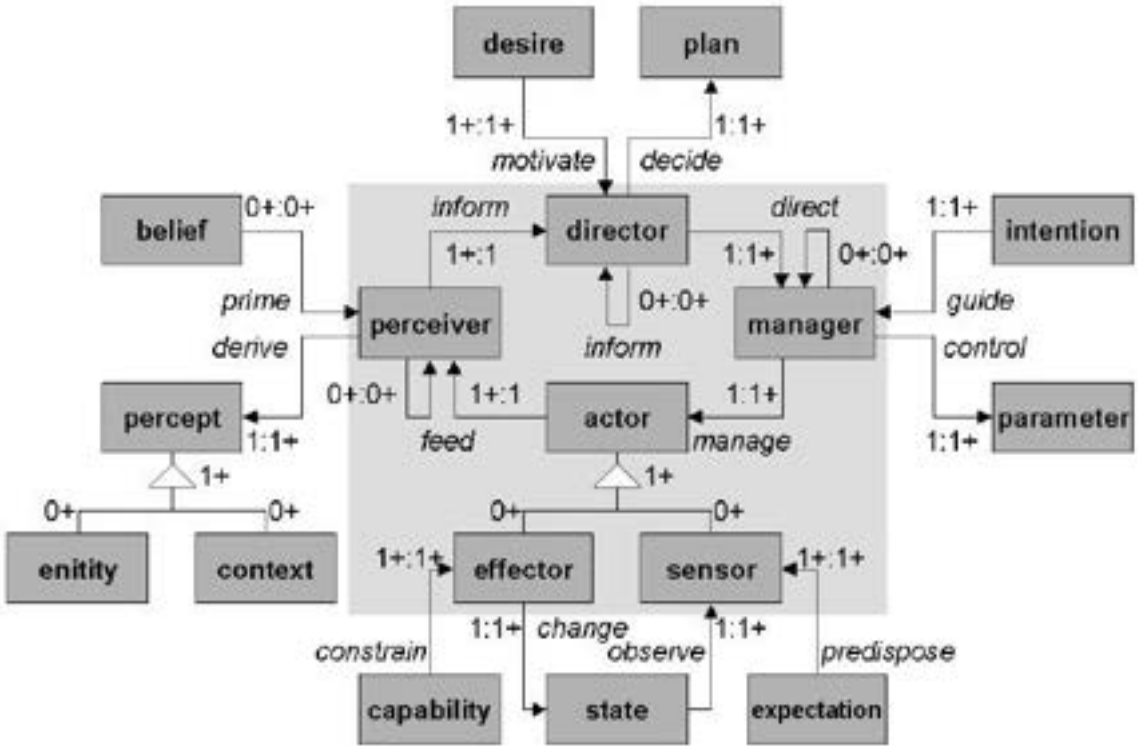

Figura 33. Versione in dettaglio del modello di fusione object-oriented

L'OOM è un modello intermedio che pone le basi per la Frankel-Bedworth *Architecture*.

3.3.3.2 Architettura Frankel-Bedworth

L'architettura Frankel-Bedworth [45] si basa sulle relazioni tra due processi che sono il *Local Process* e il *Global Process*.

Ciascun processo ha obiettivi differenti ed è composto da un insieme di regole differenti. Il *Local Process* persegue degli obiettivi nel rispetto di determinati standard, mentre il *Global Process* gestisce gli obiettivi del *Local Process* al fine di migliorare le performance dell'intero sistema.

Entrambi i processi contengono inoltre differenti *task*. Nel *Local Process* l'interazione avviene sia tra *task* interni che con quelli dell'altro processo; analogamente anche per il *Global Process* si possono individuare relazioni tra i *task* inter-processo e intra-processo.

L'analogia tra Frankel-Bedworth *Architecture* e *l'Object Oriented Model* sta nel fatto che il secondo includeva al proprio interno solo il *Local Process.*

In [3] è presente l'illustrazione della Frankel-Bedworth *Architecture* di Figura 34.

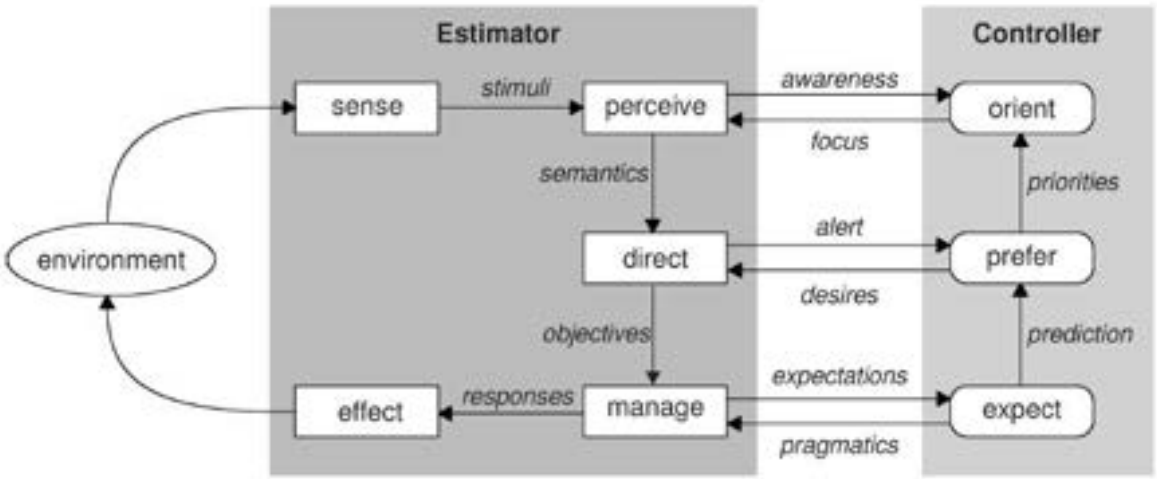

Figura 34. Architettura Frankel-Bedworth

Gli autori propongono inoltre un'altra rappresentazione del loro modello, così come mostrato in Figura 35 la cui spiegazione è ampiamente discussa in [45].

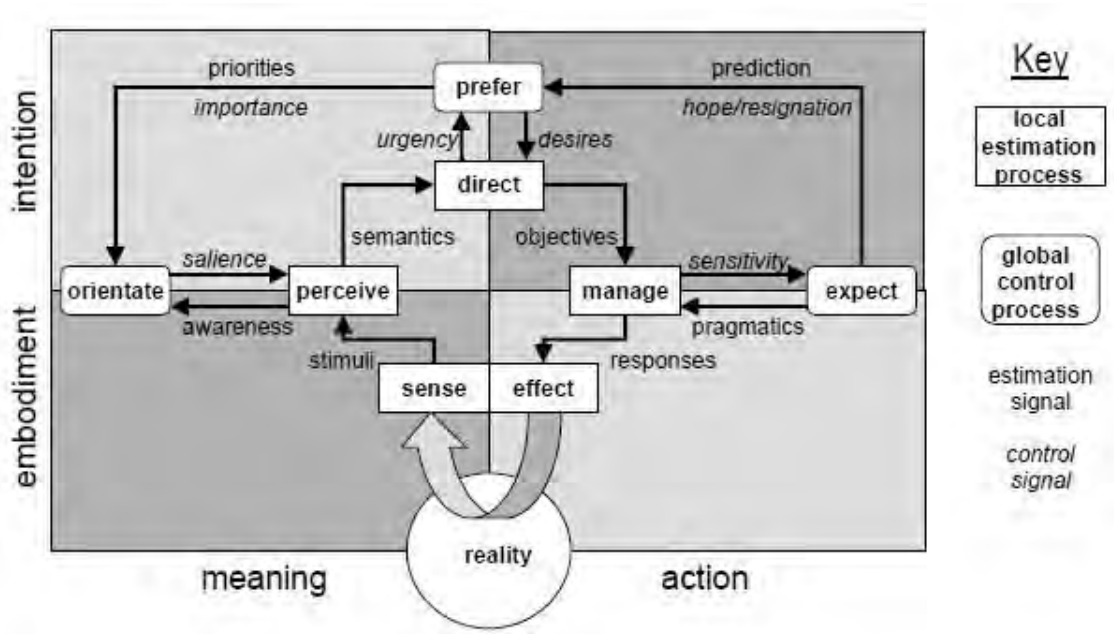

Figura 35. Rappresentazione alternativa dell'architettura di Frankel-Bedworth

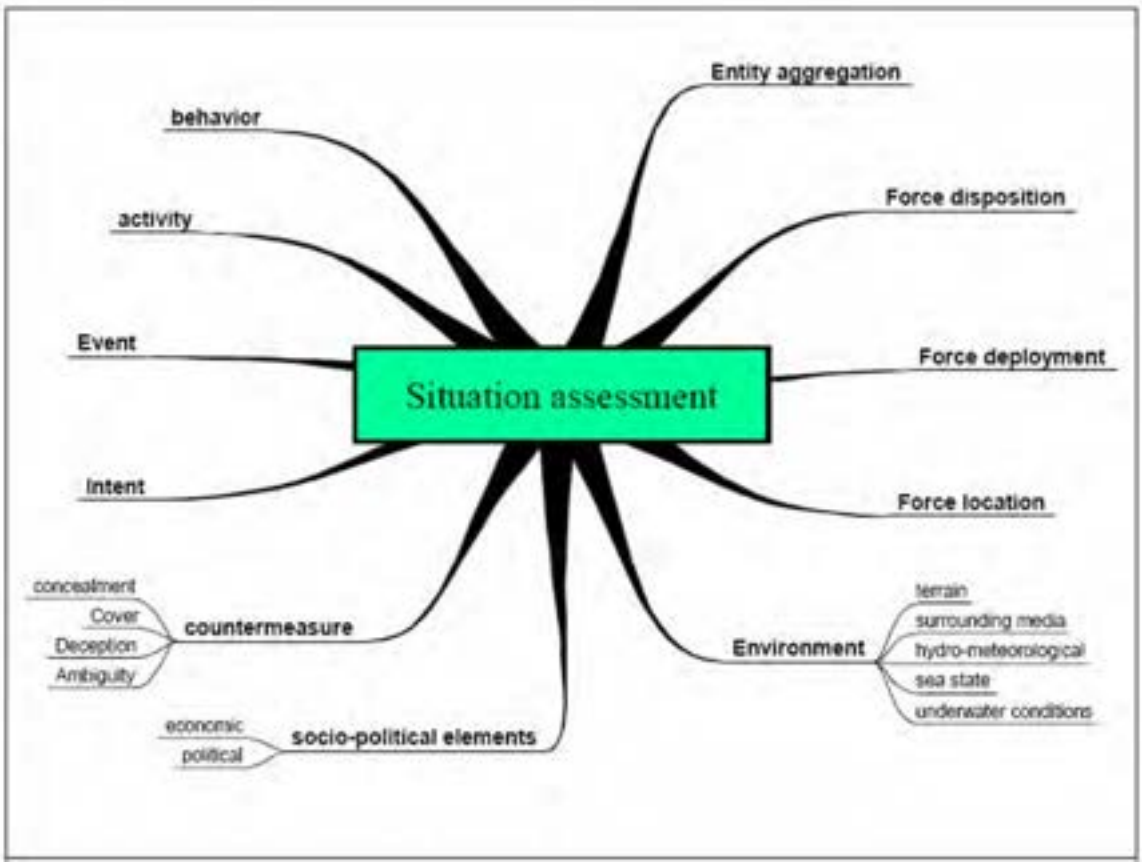

Figura 36. Elementi della situation assessment

3.3.4 Modelli basati sulla situational awareness

I modelli per i sistemi di *Information Fusion* possono essere classificati anche sulla base di come all'interno del processo di fusione si acquisisce conoscenza della situazione (*situational awareness*). Alcuni elementi che contribuiscono alla *situational awareness* [46] in ambito militare sono evidenziati in Figura 36.
All'interno di questa famiglia si individuano i seguenti modelli:

- *Endesley Situational Awareness Model*;
- *Recognition Primed Decision Making*;
- *Fusion Situation Awareness Model*;
- *SA Framework*;
- *Predictive Situation Awareness Systems.*

In generale, con il termine *situational assessment* si fa riferimento al concetto di come le persone diventano consapevoli di ciò che sta succedendo nell'ambiente in cui si trovano.

Nel settore dell'*Information Fusion* il *situational assessment* è una importante parte del processo di fusione per i seguenti motivi:

1. consente di sintetizzare le informazioni multiple;
2. fornisce un'interfaccia tra l'utente e l'automazione;
3. si occupa del reperimento e della gestione dei dati.

3.3.4.1 Endesley Situational Awareness Model

Lo *Human in the Loop* (HIL) di un sistema semi-automatico deve essere adeguato al *situation assessment*. Il *Situational Awareness Model* [47] consente di modellare il processo di acquisizione di conoscenza sulla situazione attuale da parte del sistema ed è composto da tre livelli:

- il *Livello 1* percepisce informazioni sull'ambiente;
- il *Livello 2* realizza la comprensione della situazione corrente sulla scorta di quanto percepito dall'ambiente;
- il *Livello 3* prova a proiettarsi nel futuro per immaginare quali saranno gli stati possibili.

Come illustrato in Figura 37 i tre livelli costituiscono così il processo di *Situation Awareness* (SAW). Un altro elemento molto importante del modello è il *workload* che ha influenza sia sulla fase di acquisizione della conoscenza che sulla fase di decisioni che il sistema dovrà intraprendere.

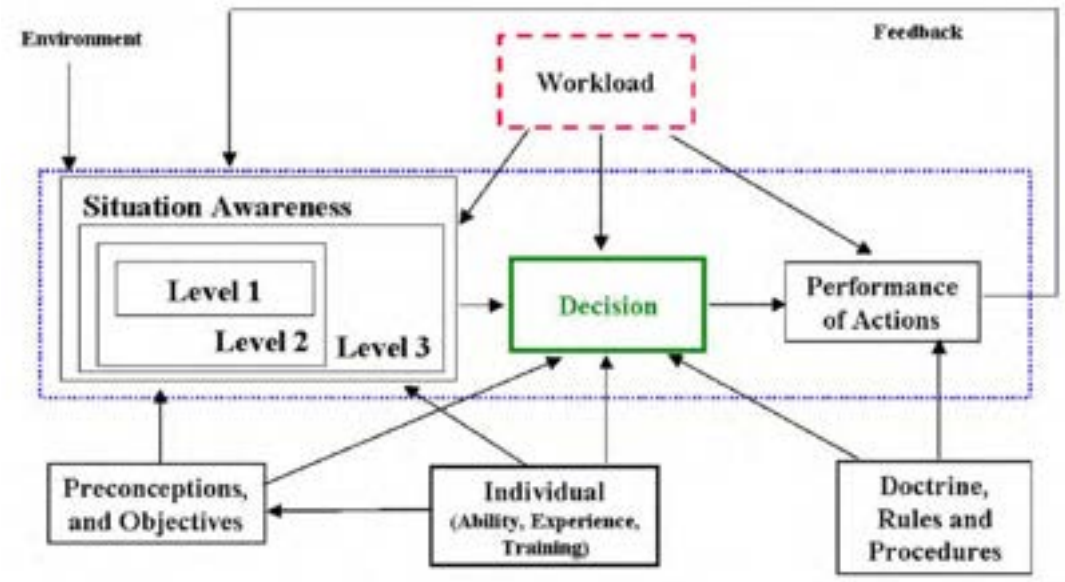

Figura 37. Modello di Situation Awareness di Endsley

3.3.4.2 Recognition Primed Decision Making

Il modello *Recognition Primed Decision Making* proposto da Klein in [48] consente di rappresentare la capacità dell'utente di prendere decisioni basate sulla conoscenza della situazione corrente e pregressa.
In Figura 38 viene rappresentato il processo di acquisizione della conoscenza secondo il modello *Recognition Primed Decision Making.*

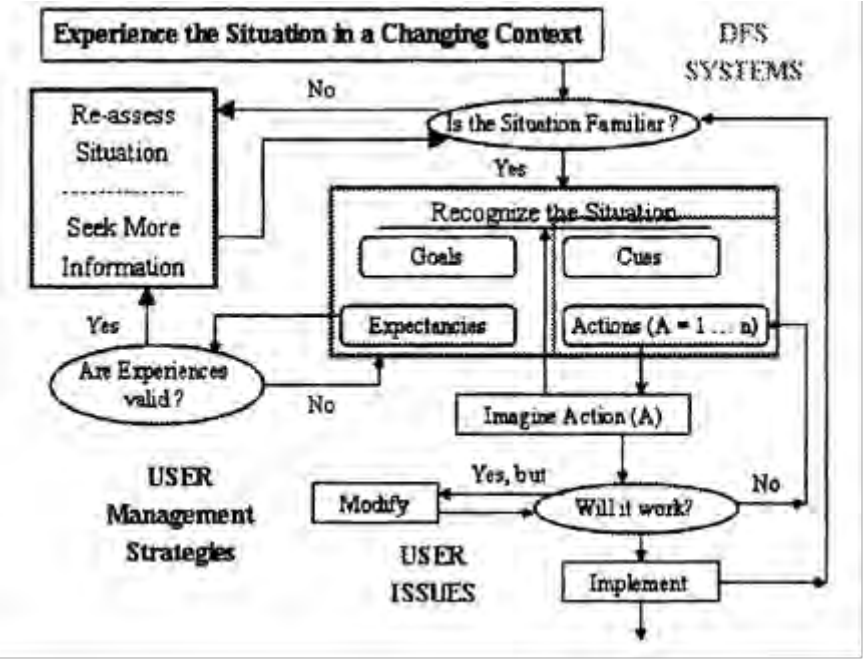

Figura 38. Modello di Situation Awareness RPD

3.3.4.3 Fusion Situation Awareness Model

Il *Fusion Situation Awareness Model* proposto da Roy in [49] organizza le informazioni delle quali l'utente ha bisogno, per poter ottenere un buona comprensione della situazione ed una buona capacità di proiezione nel futuro (SAW). Le tipologie di informazioni vengono raggruppate in diverse famiglie:

- *Situation Element Acquisition*
- *Element Association*
- *Situation Element Refinement*
- *Situation Element Contextual Analysis*
- *Situation Element Interpretation*
- *Situation Classification and Recognition*
- *Situation Assessment*
- *Situation Projection and Impact Assessment*
- *Situation Watch*
- *Process Refinement*

In Figura 39 vengono illustrate le diverse famiglie con i rispettivi elementi caratteristici.

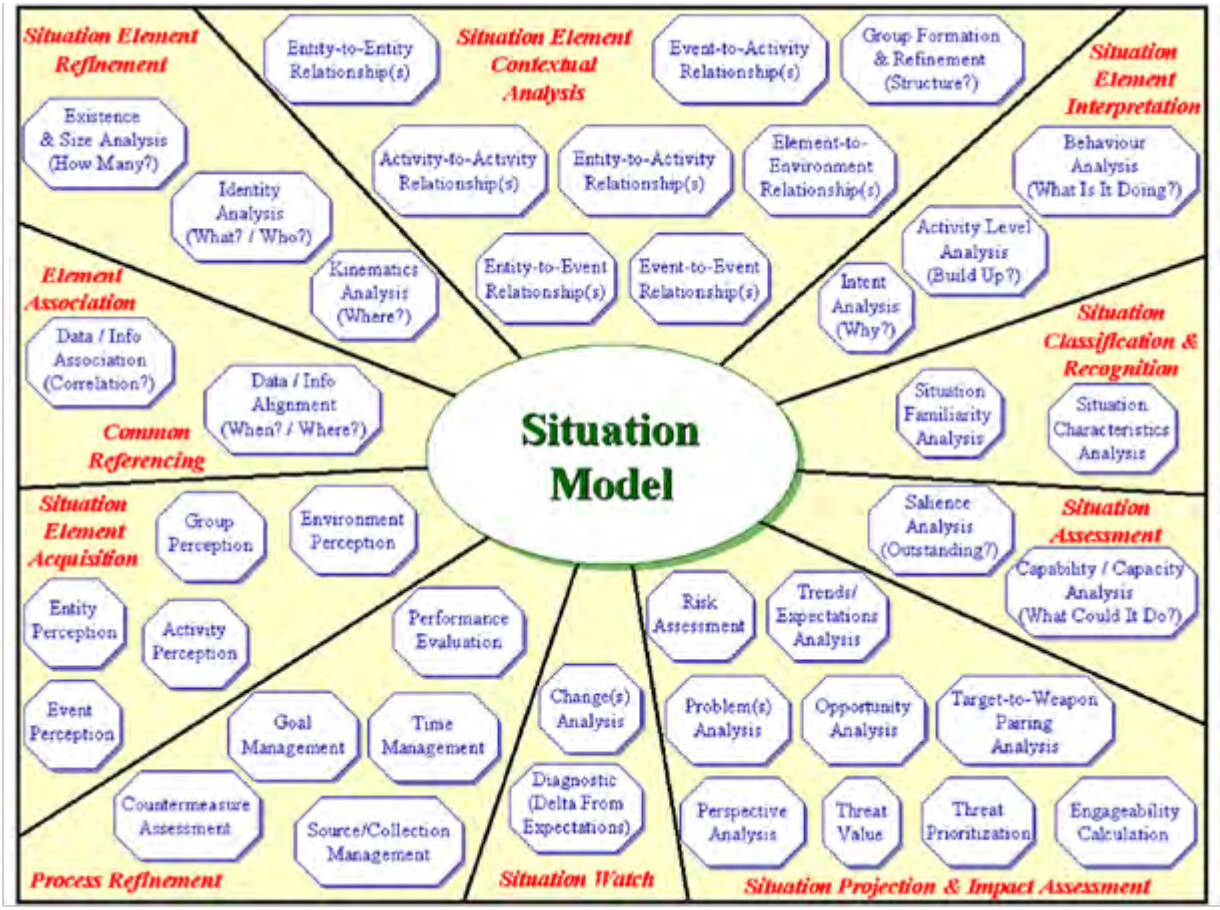

Figura 39. Modello di Situation Awareness di Fusione

3.3.4.4 Perceptual Reasoning Model

Il *Perceptual Reasoning Model* [50] si compone di due parti (così come illustrate in Figura 40 e in Figura 41):

- il *Generic Information Fusion Process Model System* (PMS);
- la *Perceptual Reasoning Machine* (PRM).

Il PMS costituisce in pratica un sistema generico di fusione dell'informazione mentre il PRM è in grado di dedurre e valutare ciò che sta avvenendo al fine di poter orchestrare l'acquisizione delle informazioni per aumentare le prestazioni dell'intero sistema. Inoltre, il PRM effettua un'analisi globale della situazione al fine di risolvere casi in cui le indicazioni ricevute siano in conflitto tra loro.

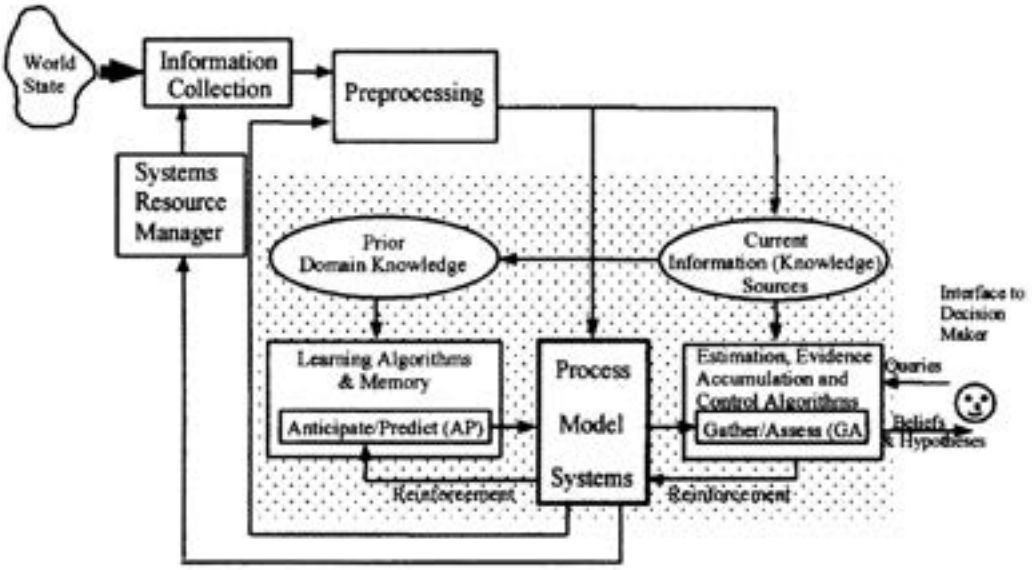

Figura 40. Modello di Fusione delle Informazioni Process Model System

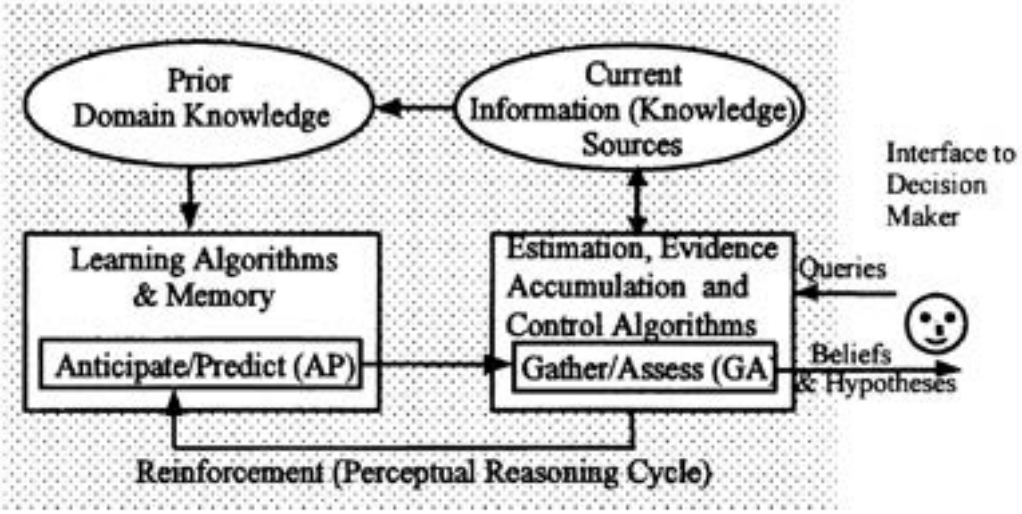

Figura 41. Modello Perceptual Reasoning Machine

3.3.4.5 SA Framework

Salerno et al. [51], [52], [53] hanno realizzato un framework in cui l'utente può interagire con il sistema al fine di realizzare un modello ex-novo o modificarne uno già esistente. Tale modello è mostrato in Figura 42 ed è stato applicato per la realizzazione di varie applicazioni strategiche, cibernetiche e tattiche.
Inoltre, sempre grazie a questa interfaccia con il sistema, si può attivare o meno modelli esistenti e ricevere informazioni su come i modelli sono correntemente usati o come sono stati impiegati nel corso del tempo.

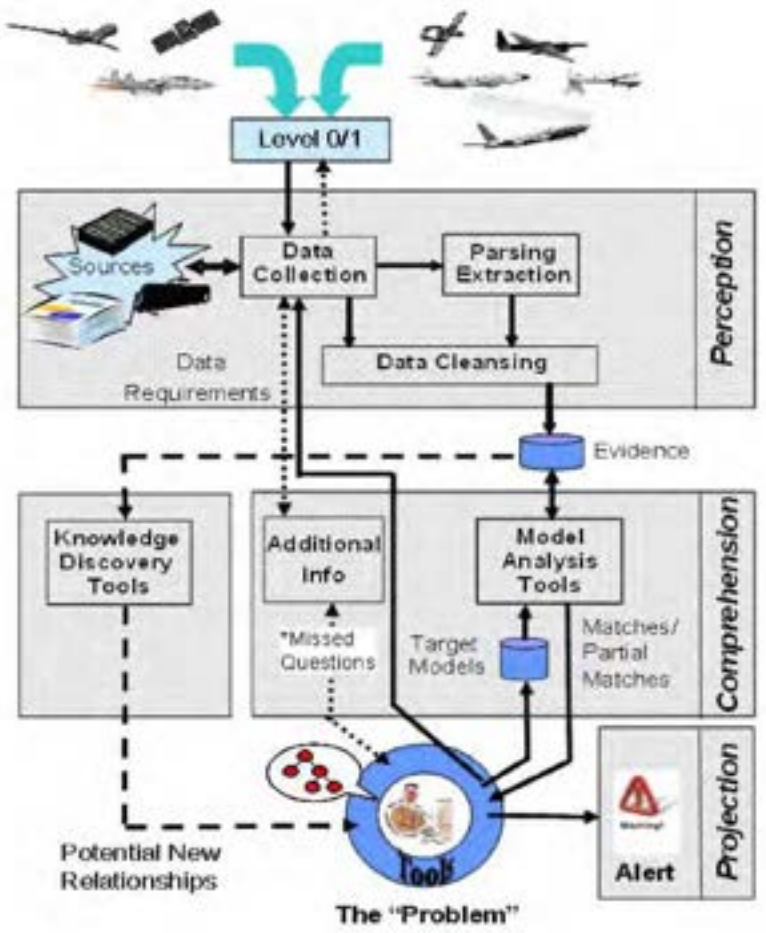

Figura 42. Modello di Salerno et al.

3.3.4.6 Predictive Situation Awareness Systems

I nuovi sistemi *net*-centrici hanno spostato il focus dei sistemi *Information Fusion* da piattaforme individuali alla rete, con un conseguente cambiamento di *focus* anche sulle tecnologie da impiegare (ad esempio SOA e *Semantically Aware Systems*). La maggior complessità dei sistemi che si intende affrontare oggi (ad esempio in ambito militare, di *Homeland Security o di Smart City*) necessita un approccio multidisciplinare, differente cioè rispetto al passato.

Un *Predictive Situation Awareness Systems* (PSAW) deve possedere un *framework* di rappresentazione in grado di catturare non solo le informazioni e le incertezze provenienti dalle varie fonti, ma anche le sottili e intricate relazioni che intercorrono tra entità di diverso tipo. Inoltre, i PASW distribuiti devono essere in grado di gestire grandi quantità di dati di diverse caratteristiche e natura, il che impone non solo algoritmi più efficienti, ma anche una migliore capacità di rappresentare e comunicazione semantica delle informazioni.

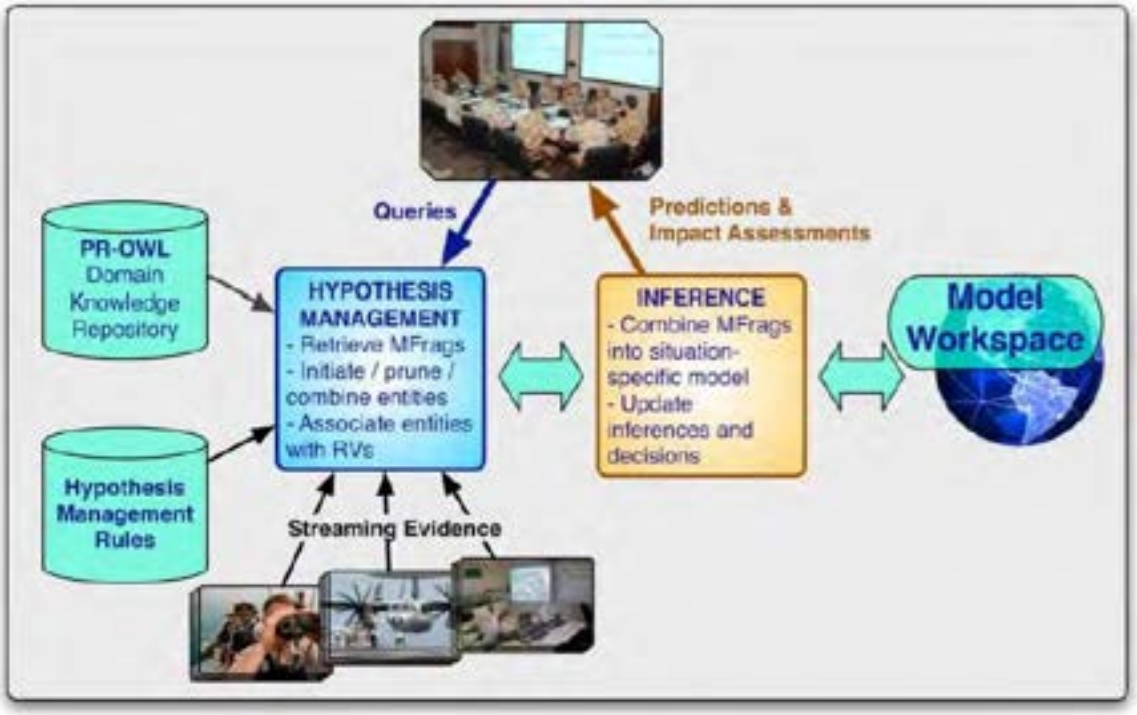

Figura 43. Predictive Situation Assessment and Impact [54]

Un moderno PSAW deve quindi [54]:

a) fornire metodologie interoperabili per la gestione dell'incertezza attraverso il processo di integrazione, al fine di fornire analisi predittive e di *impact assessment* per diversi comportamenti e in diversi ambienti;
b) disporre di una rigorosa base matematica e algoritmi efficienti per fornire previsioni affidabili;
c) includere tecniche automatiche per ridurre l'elaborazione delle informazioni da parte degli utenti;
d) operare in un ambiente distribuito.

L'obiettivo (a) può essere raggiunto mediante l'impiego di un *Probabilistic Ontology Language* come il PR-OWL. L'obiettivo (b) viene affrontato in [54] utilizzando come base rigorosa matematica il *framework* di rappresentazione delle funzioni MEBN/PR-OWL e come algoritmi si adottano lo *Spatio-Temporal Hypothesis Management* e l'*Efficient Hybrid Inference*.

3.4 Metodi per la Fusione delle Informazioni

La fusione delle informazioni è una materia multidisciplinare che affronta diverse problematiche, le quali appunto la rendono particolarmente stimolante come settore di ricerca. Le diverse problematiche vengono sintetizzate all'interno del paragrafo 3.4.1. Da una loro breve disamina risulta facilmente intuibile l'assenza in letteratura di un unico metodo per la fusione in grado di affrontare efficacemente tutte le problematiche. Successivamente, all'interno del capitolo **Errore. L'origine riferimento non è stata trovata.**, viene approfondito il rapporto tra metodi per la fusione, problematiche e prestazioni dei sistemi di fusione.

3.4.1 Criticità connesse al processo di fusione delle informazioni

Le problematiche principali derivano dalla natura dei dati oggetto del processo di fusione, dalle caratteristiche dei sensori impiegati e dalla particolare applicazione che si intende realizzare [2]. Tali problematiche sono in sintesi:

- **Imperfezione dei dati**: i dati in ingresso al sistema di fusione sono sempre in qualche modo affetti da qualche forma di imprecisione e incertezza.

- **Presenza di dati spuri**: le misure non solo sono affette da imprecisione e incertezza legate alle caratteristiche dei sensori, ma anche alla presenza di rumore all'interno dell'ambiente.
- **Dati in conflitto**: i dati da utilizzare nel processo di fusione potrebbero in qualche modo contraddirsi e portare, in prima analisi, ad affermazioni ambigue.
- **Modalità di acquisizione dei dati**: il sistema di fusione deve essere in grado di fondere sia rilevazioni effettuate da sensori aventi dello stesso tipo che informazioni provenienti da altri tipi di sensori.
- **Correlazione tra i dati**: questa problematica è particolarmente rilevante nei sistemi di fusione con topologia distribuita. Questo perché le misure effettuate in diversi nodi dell'architettura possono risultare polarizzate dalla presenza dello stesso rumore, inducendo così a considerazioni sbagliate per difetto o per eccesso di confidenza.
- **Allineamento/Registrazione dei dati**: i dati provenienti da sensori differenti devono essere allineati sulle stesse scale di riferimento (ad esempio bisogna adottare la stessa scala temporale) prima di poter essere fusi. Anche le modalità con le quali i dati vengono registrati sono importanti per le prestazioni del sistema di *Information Fusion*.
- **Associazione dei dati**: questa problematica è molto sentita in situazioni in cui le rilevazioni effettuate possono essere riconducibili a *target* differenti presenti all'interno dell'ambiente; ad esempio nel *multi-tracking* di aerei uno dei problemi che bisogna affrontare è come associare i dati misurati sul campo con gli aerei che effettivamente sono oggetto delle rilevazioni.
- **Framework di elaborazione**: l'efficienza del sistema di fusione dipende fortemente dall'architettura adottata; ad esempio nel caso di *wireless sensor networks* l'approccio decentralizzato è generalmente molto più performante di uno centralizzato.
- ***Timing* operativo**: problemi di temporizzazione sono riconducibili alle situazioni in cui i dati devono attraversare percorsi differenti prima di essere elaborati; ad esempio in un sistema di fusione decentralizzato può verificarsi che i dati arrivino non in ordine. Inoltre, problemi di *timing* possono verificarsi anche quando le misure vengono effettuate da sensori dello stesso tipo ma con frequenze diverse, imponendo quindi un loro riallineamento prima di essere trattati.
- **Fenomeni statici e dinamici**: i fenomeni sotto osservazione possono essere tempo invarianti o tempo varianti; in caso di fenomeni tempo varianti può essere utile disporre di informazioni storiche.
- **Dimensionalità dei dati**: in alcuni casi può essere rilevante la modalità con la quale si gestisce la dimensionalità dei dati al fine di ottimizzare le *performance* del sistema; ad esempio, in una architettura distribuita in cui la banda è una risorsa da gestire opportunamente, un approccio consigliabile prevede una preelaborazione nei nodi periferici dei dati al fine di inviare moli ridotte di informazioni all'interno della rete.

In Figura 44 Khaleghi e altri (2011) [2] presentano una tassonomia delle principali problematiche legate alla natura dei dati che un sistema di fusione delle informazioni deve affrontare.

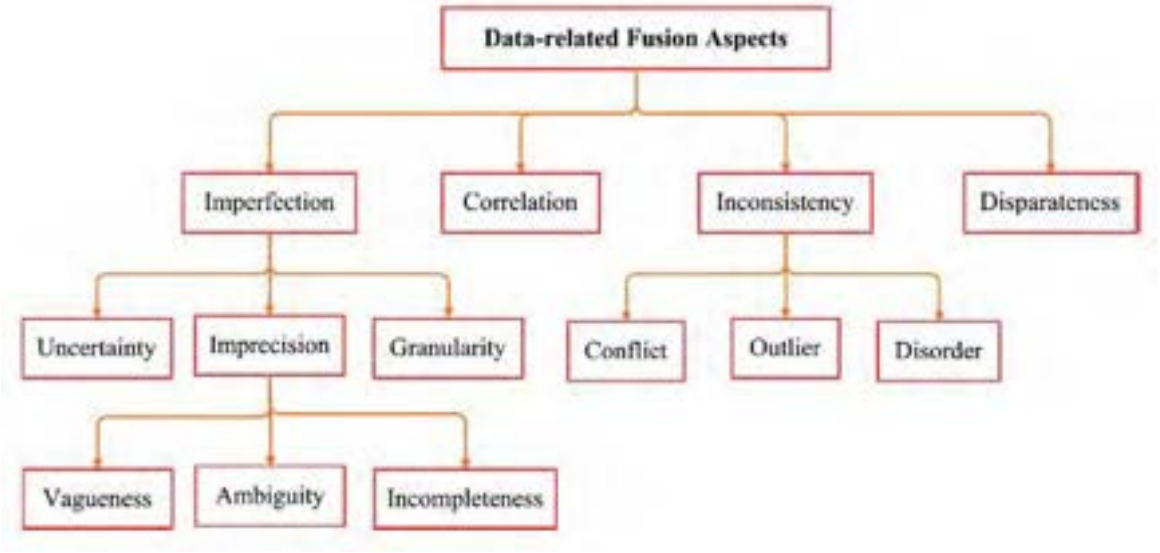

Figura 44. Tassonomia delle criticità connesse alla natura dei dati

3.4.2 Principali classi di algoritmi per l'Information Fusion

In letteratura esistono diversi criteri per classificare i metodi per la fusione delle informazioni. È possibile classificare infatti gli algoritmi in funzione di:

- livello di astrazione dei dati;
- scopo del processo di fusione;
- parametri da trattare;
- tipi di dati;
- fondamenti matematici.

In [3] Nakamura e altri descrivono la classificazione basata sullo scopo dei metodi, individuando le seguenti famiglie:

- *inference*;
- *estimation*;
- *feature maps*;
- *abstract sensors*;
- *aggregation*;
- *compression*.

In [2] sono state sintetizzate le problematiche relative alla natura dei dati. Da questa tassonomia si è tratto spunto per la realizzazione di una possibile classificazione dei metodi per la fusione delle informazioni in funzione delle problematiche associate alla natura dei dati. Tale classificazione categorizza appunto i metodi in funzione della specifica classe di problemi per la quale risultano essere più adatti. Le categorie di metodi individuate secondo questa modalità sono:

- metodi per la fusione di *dati imperfetti*;
- metodi per la fusione di *dati correlati*;
- metodi per la fusione di *dati inconsistenti*;
- metodi per la fusione di *dati disparati*.

L'analisi di questa classificazione viene ampiamente trattata all'interno del capitolo **Errore. L'origine riferimento non è stata trovata.**, con una particolare attenzione alle prestazioni dei metodi in risposta a determinate condizioni.

4 Aspetti prestazionali delle tecniche di Information Fusion

Come anticipato all'interno del paragrafo 3.4 le problematiche legate al processo di fusione delle informazioni sono tali da impedire l'individuazione di un metodo che risponda efficacemente a ciascuna di esse. In molte applicazioni si impone quindi la scelta di più metodologie, il cui utilizzo combinato è in grado di implementare il processo di fusione garantendo *performance* migliori.
All'interno di questo capitolo verranno discussi gli aspetti prestazionali degli algoritmi in relazione alla classe di problematiche per la quale risultano più adatti.

4.1 Metodi per la fusione di dati imperfetti

Uno dei problemi più diffusi legati alla natura dei dati è la presenza di imperfezioni all'interno delle rilevazioni effettuate. I dati possono considerarsi imperfetti sotto diversi punti di vista e caratterizzati da:

- incertezza;
- ambiguità;
- vaghezza;
- incompletezza;
- granularità.

Diverse teorie matematiche vengono impiegate per gestire dati di questo tipo, il cui comportamento è naturalmente diverso in relazione alle caratteristiche specifiche dei dati. In Figura 45 si evidenzia come alcuni *framework* matematici sono in grado di affrontare anche più classi di problemi legate alla natura imperfetta dei dati [2].
Le tecniche di fusione delle informazioni che meglio affrontano dati imperfetti sono:

- *Probability Theory*;
- *Soft Computing Techniques*;
- *Rough Set Theory*;
- *Possibility Theory*;
- *Dempster-Shafer Evidence Theory*;
- *Hybrid Framework*;
- *Random Set Theory.*

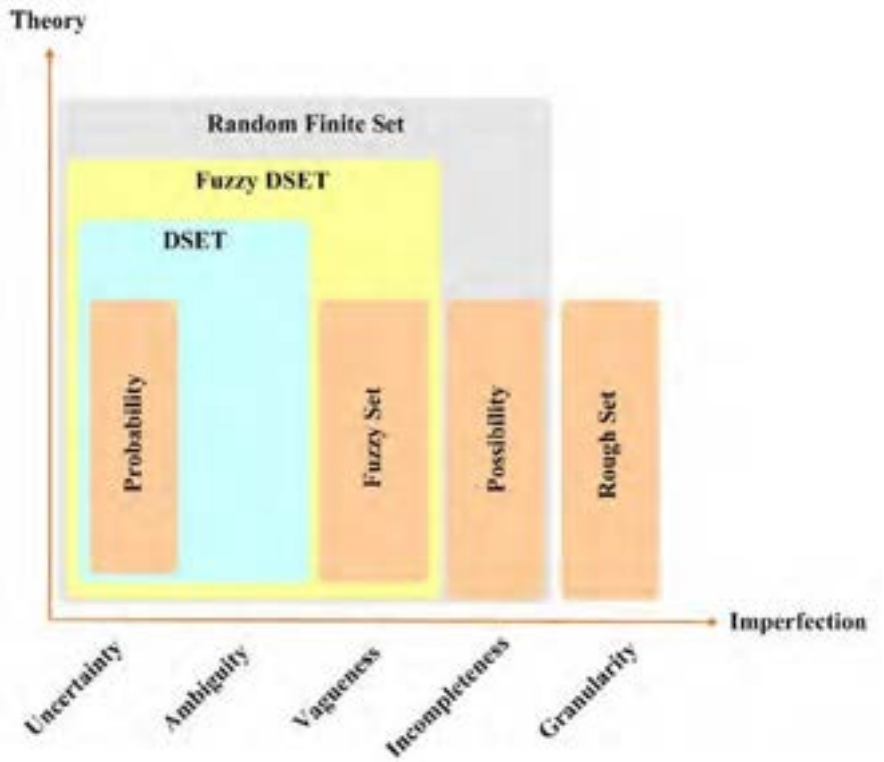

Figura 45. Relazioni tra imperfezione dei dati e framework matematici

4.1.1 Teoria della Probabilità

Molti metodi applicabili nel campo *dell'Information Fusion* trovano origine nella teoria della probabilità (*Bayesian Fusion*) [55]. Questa *framework* matematico è particolarmente adatto nell'affrontare problemi di incertezza dei dati. Tuttavia dimostra diverse debolezze nel caso di sistemi complessi e nei casi in cui non si ha una conoscenza approfondita dell'incertezza e dell'inconsistenza dei dati, nonché della precisione dei modelli.

La *Bayesian Fusion* si basa sul *Bayesian Estimator* ed è un approccio molto ben consolidato per la fusione di dati incerti, pur tuttavia mostrando incapacità nella gestione di altre tipologie di imperfezioni dei dati.

4.1.1.1 Estimatore Bayesiano

L'*Estimatore Bayesiano* consente la fusione dei dati attraverso il calcolo della probabilità a posteriori (regola di Bayes delle probabilità condizionate) [56], [57], [58], [59], [60].

Il *Bayes Estimator* è stato utilizzato in molte applicazioni, soprattutto nei precedenti decenni nei quali mancavano teorie matematiche nel campo dell'*Information Fusion* diverse dalla teoria della probabilità. Uno dei principali svantaggi è l'impossibilità in generale di valutare analiticamente determinati integrali nella regola di Bayes delle probabilità condizionate:

$$Pr(Y|X) = \frac{\Pr(X|Y)\,\Pr\,(Y)}{\Pr\,(X)}$$

In particolare tali problemi analitici si incontrano nella soluzione della probabilità a priori $\Pr(X|Y)$ e nel termine di normalizzazione $\Pr(X)$.

In letteratura si trovano molti casi di applicazioni di *Information Fusion* in cui si è utilizzata tale metodologia: Plan ed altri 1998 [56] hanno proposto tale metodologia per la realizzazione di un'architettura che loro definiscono NN-*Bayesian* (*Neural Network Bayesian*) per l'*Information Fusion*.

In [61] sono state utilizzate le *Dynamic Bayesian Network* nei livelli più alti di astrazione dei dati nell'architettura di *Information Fusion* (*High-Level Data Fusion*) per modellare attività terroristiche.

Il *Bayes Estimator* può essere classificato anche come un metodo di *inference* [3], in quanto consente di poter prendere una decisione sulla base della conoscenza dedotta dalla situazione presa in considerazione.

4.1.1.2 Metodo MAP (Maximum A Posteriori)

Il metodo *Maximum A Posteriori* (MAP) [62] consente di stimare lo stato *x* a partire dall'osservazione del suo valore in *k* osservazioni. Il valore stimato di *x* è il valore che massimizza la funzione *posterior distribution*:

$$\hat{x}(k) = \arg max_x \; p(x|\boldsymbol{z})$$

dove *x* è il parametro da stimare e ***z*** *= [z(1),…,z(k)]* è una sequenza di *k* osservazioni di *x* e *p(x)* è una pdf nota.
Il metodo MAP è utilizzato nell'*Information Fusion* come, ad esempio, nel caso della realizzazione di un sistema di riconoscimento biometrico basato sull'autenticazione della parola e del viso.
Il MAP è classificabile anche come un metodo di stima (*estimation*) [3].

4.1.1.3 Filtro Moving Avarage

Il *Moving Avarage Filter* [63], [64], [65] è un filtro di estrema semplicità realizzativa, particolarmente adatto alla fusione di dati aventi basso livello di astrazione (*Sensor Fusion*). Per queste sue caratteristiche trova largo utilizzo in campi applicativi analoghi al *Digital Signal Processing*. Consente di stimare il valore di un parametro all'istante k calcolando la media dei precedenti M campioni rilevati.
Il *Moving Avarage Filter* è molto utilizzato nella fusione delle informazioni all'interno di *Wireless Sensor Network* ed è inquadrato tra i metodi di stima (*estimation*) [3].

4.1.1.4 Metodo ML (Maximum Likelihood)

Il metodo *Maximum Likelihood* (ML) [66] [67] consente di stimare i parametri di interesse a partire da dei valori di training. Viene utilizzato quando i parametri sono quantità fissate ma sconosciute, a differenza dell'approccio bayesiano in cui i parametri sono variabili aleatorie aventi determinate probabilità a priori. La stima dei parametri di interesse viene effettuata andando a massimizzare la funzione di verosimiglianza ottenibile da modelli di sensori empirici o analitici.
Il *Maximum Likelihood Estimator* (MLE) è impiegato soprattutto nel dominio dell'*Information Fusion* legato alle *Wireless Sensor Network*. Il MLE viene inquadrato tra i metodi di stima (*estimation*) [3].

4.1.1.5 Metodi dei minimi quadrati

Il *Least-Square Estimator* [68], [69] è una tecnica di ottimizzazione matematica che consente di trovare una funzione che tenda il più possibile ad un'interpolazione di un insieme di dati. In particolare, viene ricercata la funzione che minimizzi la somma dei quadrati delle distanze dai punti dati (misure effettuate). Tale tecnica di ottimizzazione risente soprattutto di problemi legati alla presenza di rilevazioni spurie o corrotte.
Il *Least-Square Estimator* viene inquadrato tra i metodi di stima (*estimation*). La tecnica del *Least-Square Estimator* è molto utilizzata nell'*Information Fusion* nell'ambito delle *Wireless Sensor Network*.

4.1.1.6 Metodi basati sulla simulazione Monte Carlo (MC): MC Sequenziale e Markov Chain MC

I metodi basati sulla simulazione Monte Carlo sono tra i più flessibili per l'approssimazione di probabilità; infatti essi non si basano su alcuna assunzione sulla densità di probabilità da approssimare. In questa classe di metodi si individuano gli algoritmi *Sequential Monte Carlo* [70], [71], [72], come il *Particle Filter*, e i metodi *Markov Chain Monte Carlo* (di cui il *Metropolis-Hastings* vuole essere un'ottimizzazione).

4.1.1.7 Particle Filter

Il *Particle Filter* [73] [74] [75] [76] [77] [78] è un'implementazione ricorsiva dell'algoritmo *Sequential Monte Carlo*. Tale metodo, dati degli ingressi e delle uscite nel tempo, tenta di simulare la distribuzione di probabilità con la quale stimare lo stato del sistema.
In alcuni casi il *Particle Filter* viene utilizzato per stimare la probabilità a posteriori all'interno di un *framework* bayesiano mediante la generazione di valori casuali. Tali valori (*particles*) vengono ottenuti a partire dalla probabilità a priori. Il metodo *Particle Filter* risulta essere in alcuni casi oneroso da un punto di vista computazionale in quanto richiede numerosi valori casuali per poter stimare la distribuzione di probabilità.
Il *Particle Filter* risulta quindi di facile implementazione e adatto a trattare ogni tipo di non linearità e di non stazionarietà, nonché ogni tipo di distribuzione di probabilità. Viene anche indicato come algoritmo di *condensation* o di *sopravvivenza dei più appropriati*.
Il *Particle Filter* viene inquadrato tra i metodi di stima (*estimation*). Il *Particle Filter* è impiegato in numerosi campi dell'*Information Fusion* (*computer vision* [74], *tracking* [75], *location discovery* [76] in *Wireless Sensor Network*).

4.1.1.8 Algoritmi Markov Chain Monte Carlo

Gli algoritmi *Markov Chain Monte Carlo* (MCMC) [79] cercano di risolvere il problema che si può riscontrare nel *Particle Filter* in casi di dimensionalità elevata dello spazio degli stati. La soluzione proposta dagli algoritmi MCMC prevede non la generazione di valori casuali (*particles*) ad ogni iterazione, bensì una loro evoluzione nel tempo come in una catena markoviana. È stato dimostrato che, se il sistema è stato progettato correttamente, dopo un numero sufficiente di iterazioni (*burn-in period*) si tende ad un'unica densità stazionaria.

4.1.1.9 Metropolis-Hastings Algorithm

Diversi lavori svolti da Metropolis, Hastings e altri [80], [81] sono basati sui metodi MCMC e sono in grado di affrontare efficacemente il problema di dimensionalità elevata. Il loro metodo prevede la scelta di volta in volta di uno stato candidato a partire dallo stato corrente, mediante una apposita probabilità di transizione tra gli stati.
Il principale punto debole di questo metodo è una elevata sensibilità alla scelta dei campioni iniziali e delle probabilità di transizione che, se scelti male, potrebbero determinare un *burn-in period* eccessivamente lungo.

4.1.1.10 Kalman Filter

I *Kalman Filter* (KF) [82], [83], [84], [85] hanno come scopo la soluzione dei problemi analitici che si riscontrano nell'impiego dello stimatore bayesiano. In particolar modo vengono introdotte delle approssimazioni che consentono di ridurre la dinamica del sistema ad una forma lineare e gaussiana. Se queste forzature non si discostano eccessivamente dalla realtà, allora l'errore introdotto può considerarsi accettabile. Questi presupposti rendono il metodo KF facilmente implementabile e largamente impiegato in diversi campi applicativi (soprattutto utilizzati in problemi di localizzazione e *tracking*).
Tuttavia non sempre le approssimazioni possono essere effettuate in maniera coerente e reale con il sistema da rappresentare. In questi casi, come ad esempio quando il sistema è non lineare, si deve ricorrere a tecniche particolari di approssimazioni.
Inoltre i filtri di Kalman sono particolarmente sensibili alla presenza di dati corrotti e inadatti alle applicazioni in cui le caratteristiche di errore non sono facilmente parametrizzabili [2].

I metodi KF vengono largamente utilizzati nella fusione dei dati ai livelli più bassi del processo di fusione (*Low Data Fusion*). In conclusione i Kalman Filter sono algoritmi di stima dell'ottimo per sistemi dinamici lineari sui quali agiscono rumori gaussiani a media nulla.

4.1.1.11 Metodi EKF e UKF

I filtri di Kalman (KF) sono particolarmente adatti nel caso in cui il sistema sia lineare [2]. In casi di non linearità si deve ricorrere a delle particolari tecniche di approssimazione. Una di queste tecniche è il metodo *Extended KF* (EKF) [86]. L'*Extended KF* consente di affrontare tale problematica rappresentando il sistema mediante espansione in serie di Taylor approssimata al primo ordine. Qualora l'approssimazione del sistema non fosse ancora sufficientemente precisa è possibile ricorrere al metodo degli *Unscented KF* (UKF) [87] [88], dove l'approssimazione arriva al secondo ordine.

4.1.1.12 Metodi Grid-Based

I metodi *Grid-Based* [89] affrontano l'approssimazione della non-linearità del sistema in maniera differente rispetto a quanto fatto dagli EKF e UKF. Tuttavia al crescere della dimensionalità del sistema il problema computazionale cresce esponenzialmente fino a diventare irrisolvibile [2].

4.1.1.13 Hidden Markov Model (HMM)

Un *Hidden Markov Model* (HMM) [90], [91] è un modello stocastico per la valutazione della probabilità di una sequenza di eventi. Esso consente di determinare il percorso di transizione tra gli stati più probabile e stimare i parametri che rappresentano meglio tale percorso. I metodi legati all'HMM sono molto diffusi in applicazioni di vario tipo come ad esempio *di signal recognition*; in [92] vengono utilizzati in un'applicazione volta alla prevenzione di attacchi terroristici, in particolare sono stati impiegati per la fusione dei dati ai livelli più bassi per modellare reti terroristiche.

L'HMM consiste in due processi stocastici. Il processo *hidden* può essere osservato attraverso un altro processo stocastico che produce osservazioni O_s. Gli stati S rappresentano alcune condizioni non osservabili del sistema; in ogni stato c'è una data probabilità di produrre degli output osservabili del sistema O, contestualmente ad una probabilità che esprime i probabili prossimi stati. In [92] il caso di folla normale l'HMM è descritto dai seguenti parametri:

$$\lambda_N = (A, B, \pi, S, O_s)$$

dove A è la distribuzione di probabilità delle transizione tra gli stati, B è la distribuzione di probabilità delle osservazioni in ogni stato e π è la distribuzione di probabilità dello stato iniziale. I parametri A, B e π possono essere ottenuti effettuati un *training* di λ_N. Il comportamento della folla viene valutato calcolando il valore di verosimiglianza L_N per una sequenza di osservazioni O_s:

$$L_N = P(Os \mid \lambda_N)$$

Il valore di L_N viene confrontato poi con una soglia al fine di stabilire se si è in presenza di una situazione normale o meno.

4.1.2 Soft Computing Techniques

Le tecniche di *Soft Computing* consentono di trattare i casi in cui ottenere delle soluzioni presenta dei problemi computazionalmente intrattabili, offrendo delle soluzioni approssimate [93]. Le *Soft Computing Techniques* maggiormente adoperate all'interno dei processi di fusione delle informazioni sono legate alla logica fuzzy e alle reti neurali.

4.1.3 Fuzzy Set Theory

La *Fuzzy Set Theory* [94], [95], [96], [97], [98] consente di affrontare situazioni in cui i dati a disposizione risultano essere affetti da vaghezza. La logica fuzzy introduce la nozione di *appartenenza parziale* che consente di fondere dati vaghi o parziali messi a disposizione dai sensori.
Le regole di fusione *fuzzy* possono essere suddivise in:

- *congiuntive*;
- *disgiuntive*.

Le regole congiuntive risultano essere particolarmente adatte nei casi in cui i dati da fondere sono prodotti da sensori ugualmente disponibili ed omogenei. Le disgiuntive affrontano invece con maggior successo i casi in cui i dati sono profondamente in conflitto tra loro o quando le sorgenti non sono equamente disponibili.
Sono state inoltre sviluppate delle evoluzioni delle regole *fuzzy* in modo da poter fondere i pregi delle regole congiuntive e disgiuntive (*adaptive fuzzy fusion*).
La *Fuzzy Set Theory*, dato che affronta con efficacia il problema della vaghezze dei dati, è particolarmente adatta ai casi in cui si devono fondere dati vaghi come quelli prodotti da esseri umani (generalmente espresse quindi in linguaggio naturale). In genere viene utilizzata in combinazione con altre tecniche di fusione, come ad esempio con quelle probabilistiche o della teoria dell'evidenza [99], [100].

4.1.4 Fuzzy Cognitive Map

Una *Fuzzy Cognitive Map* consente di realizzare un modello di un dato scenario mentale utilizzando la logica *fuzzy* per rappresentare le relazioni tra gli elementi al suo interno. Le *Fuzzy Cognitive Map* vengono utilizzate in vari campi applicativi in cui è necessario modellare processi incerti e complessi.
In [101], [102], [103] le *Fuzzy Cognitive Map* vengono utilizzate all'interno di architetture di *Information Fusion*, anche in combinazione con altre tecniche di fusione, per avere costantemente la *situation awareness* del sistema.

4.1.5 Fuzzy Ontology

All'interno di [104] si pone l'attenzione su quali siano le peculiarità dei livelli più alti nell'architettura di un sistema di fusione dell'informazione e su come le ontologie possano essere impiegate nel supportare l'uomo nel prendere decisioni a valle del processo di fusione.
Uno scenario di esempio è quello illustrato in [105]. In tale scenario gli autori presentano un sistema per suggerire libri in funzione delle informazioni che si hanno sull'utente; in particolare si impiega la logica *fuzzy* e le ontologie per la descrizione di input quali ad esempio la fama, il prezzo e parole chiave per ottenere le raccomandazioni più interessanti per l'utente. L'ontologia è una relazione gerarchica tra concetti all'interno di un dato dominio che può essere vista come un grafo. La *fuzzy ontology* presentata consente di collezionare le relazioni *fuzzy* tra le diverse classi nell'ontologia.

4.1.6 Neural Network

Le *Neural Network* [106], [107] trovano larga diffusione nei casi in cui non si hanno sufficienti nozioni su come possano essere relazionati gli ingressi e le uscite di un sistema [2]. Tale caratteristica delle rete neurali è dovuta al fatto che esse apprendono le relazioni causa-effetto tra i dati attraverso l'apprendimento da esempi (*training data*). Le *Neural Network* sono quindi in grado di dare un significato ai dati visti in fase di *training* (categorizzare), per poi essere in grado di generalizzare dando risposte a ingressi mai affrontati in precedenza.
Tra i vantaggi delle reti neurali vi sono una buona immunità al rumore e tolleranza ai guasti. Uno svantaggio delle reti neurali è che generalmente non si può risalire al ragionamento che esse hanno sviluppato per fornire i risultati. Inoltre, è importante scegliere i dati di addestramento in maniera accurata e tenere in considerazione che il tempo di addestramento cresce sensibilmente con il numero dei nodi, limitando così la dimensione della rete.

Le *Neural Network* possono essere in diversi ambiti applicativi dell'IF, soprattutto nel campo dell'*Automatic Target Recognition* [108], [109] [110], [111]. Le *Neural Network* possono essere considerate un'alternativa alle teorie di Bayes e di Dempster-Shafer per la classificazione ed il riconoscimento dei task.

4.1.7 Rough Set Theory

La teoria dei *Rough Set* [112], [113] si basa matematicamente sulla *relazione di indiscernibilità* che induce una partizione dell'*Universo* in blocchi di oggetti indiscernibili. Tali blocchi vengono indicati come *insiemi elementari* o *granuli*. La teoria dei *Rough Set* consente di gestire la vaghezza e le incongruenze nelle informazioni. Ogni sottoinsieme *X* dell'*Universo* può essere espresso come un'unione di granuli e le informazioni relative ad un determinato problema possono essere rappresentate da regole decisionali. Questa metodologia consente di trattare sia dati qualitativi che quantitativi. Classicamente la *Rough Set Theory* non consente la possibilità di indicare delle preferenze nella fase di scelta della soluzione; tuttavia alcune sue estensioni, come la *Dominance-based Rough Set Approach* [114], superano questa limitazione.
Una volta che i dati sono stati suddivisi in granuli, allora possono essere fusi mediante la teoria classica degli insiemi. Affinché la fusione avvenga con successo i granuli non devono essere né troppo "grossi" ne troppo "fini". Il maggior vantaggio di questa teoria è che non si devono avere conoscenze preliminari o addizionali sulla distribuzione dei dati o su funzioni di appartenenza.
La *Rough Set Theory* è stata tuttavia poco usata nel campo degli *Information Fusion* in quanto relativamente nuova. Un esempio di utilizzo si trova in [115] per la gestione di robot.

4.1.8 Possibility Theory

La *Possibility Theory* [116], [117], [118], [119] è un approccio non probabilistico che intende affrontare il problema dell'incertezza dei dati. Come nella logica *fuzzy* (sulla quale si basa), le preferenze sono valori compresi tra 0 e 1 che vengono associate a degli eventi, esprimendone il livello di plausibilità con il quale l'evento si verifica.
A differenza della logica *fuzzy*, la *Possibility Theory* cerca di affrontare più il problema dei dati incompleti rispetto al problema di dati vaghi, trattando i dati imperfetti in maniera simile a quanto fatto dalla teoria dell'evidenza di Dempster-Shafer. Rispetto al classico approccio *fuzzy* la *possibility theory* è in grado di gestire scenari più complessi. Tuttavia non è molto impiegata dalla comunità dell'*Information Fusion*, anche se alcuni ricercatori hanno dimostrato che è del tutto competitiva nei confronti degli approcci basati sulla teoria probabilistica o sulla teoria dell'evidenza. Inoltre la fusione possibilista risulta essere più adatta in ambienti poveri di informazioni.
In un recente lavoro [120] è stata impiegata la fusione possibilista per la localizzazione di robot in un ambiente interno parzialmente noto.

4.1.9 Dempster-Shafer Evidence Theory

La teoria dell'evidenza (nota anche come *Dempster-Shafer Belief Accumulation* o *Theory of Evidence* o *Dempster-Shafer Evidenzial Reasoning*) è una teoria matematica molto diffusa nella gestione di problemi in cui i dati sono incerti e imprecisi [2], [121], [122], [123] [124], [125], [126]. Può essere considerata come una generalizzazione della teoria bayesiana, che fornisce un formalismo per la rappresentazione della conoscenza incompleta. Questa teoria introduce l'assegnazione di opinioni e plausibilità per le ipotesi sulle misure possibili e regole per fonderle insieme.
A differenza della teoria bayesiana, la *Dempster-Shafer Theory* (DS-*Theory*) consente ad ogni sorgente di contribuire all'informazione a diversi livelli di dettaglio. La DS-*Theory* non assegna alcuna probabilità *a priori* a proposizioni sconosciute, mentre le probabilità vengono assegnate solo quando l'informazione a supporto è disponibile.
La teoria bayesiana si dimostra molto più accurata mentre la teoria dell'evidenza è molto più flessibile. LA DS-*Theory* si è dimostrata inoltre molto promettente e risulta essere quindi molto adottata nella comunità

dell'*Information Fusion* in diversi domini applicativi: dal campo medico al tracciamento, dalla localizzazione all'elaborazione dei segnali.
Tra gli svantaggi si riscontrano una crescita computazionale esponenziale e la possibilità di produrre dei risultati contraddicenti (soprattutto nel caso di dati in conflitto tra loro). Queste limitazioni tuttavia vengono affrontate e superate grazie ad alcune evoluzioni.

4.1.10 Framework ibridi

Alla base della realizzazione di *framework* ibridi [127], [128], [129] per la fusione dei dati vi è il presupposto che metodologie diverse (quali la teoria probabilistica, la teoria dell'evidenza e la logica *fuzzy*) affrontano il problema della fusione da diversi punti di vista e non in maniera concorrente tra loro [2]. All'interno di questo paragrafo vengono presentati alcuni dei metodi presenti in letteratura, ovvero la *Fuzzy Rough Set Theory* e la *Fuzzy Dempster-Shafer Theory*.

4.1.10.1 Fuzzy Rough Set Theory

La *Fuzzy Rough Set Theory* (FRST) è una soluzione ibrida che include la logica *fuzzy* ed è stata per la prima volta studiata da Dubois e Prade [130] per poi essere generalizzata da Yeung e altri [131] al fine di ovviare ad alcune sue limitazioni. Non vi sono molte applicazioni di questa teoria che, così come anche la *Rough Set Theory*, fa ancora fatica ad affermarsi nel campo dei sistemi di *Information Fusion*. Un esempio applicativo è riportato in [132].

4.1.10.2 Fuzzy-Dempster-Shafer Theory

In letteratura si possono incontrare diversi lavori in cui si combinano la *Fuzzy Set Theory* e la *Dempster-Shafer Theory* al fine di realizzare dei *framework* ibridi in grado di affrontare meglio il problema di dati imperfetti. Il lavoro più popolare è probabilmente quello svolto da Yen [133] che è stato inoltre ripreso da diversi autori, come nei lavori svolti da Zhu e Basir [100] [134] per la segmentazione di immagini.

4.1.11 Random Set Theory

Questa teoria matematica ha origine negli anni '70 e diversi ricercatori hanno dimostrato nel tempo come questa potesse essere utilizzata come base per un *framework* unificato per la fusione dei dati; soprattutto l'attività di ricerca di Mahler [135] ha valorizzato questa tesi. La maggior parte dei lavori presenti in letteratura è orientata a problematiche di *multitracking*. Analiticamente questi problemi possono essere molto complessi o irrisolvibili e diversi studi hanno fornito delle metodologie di approssimazione in grado di superare questo limite.
Altri lavori recenti hanno dimostrato la capacità di questa teoria di fondere informazioni in applicazioni di *target detection* [136], *tracking* [137], *identification* [138], *sensor management* [135] e *hard/soft data fusion* [139].
La *Random Set Theory* è una teoria relativamente nuova e non molto apprezzata nella comunità dell'*Information Fusion*, anche se potenzialmente potrebbe fornire un *framework* unificato per la fusione di dati imperfetti [2].

4.2 Metodi per l'eliminazione della correlazione tra i dati

La correlazione dei dati è particolarmente problematica nei sistemi di *data fusion* distribuiti. Si verifica soprattutto con la presenza di *data incest*, ovvero quando l'informazione raccolta da un sensore raggiunge il punto di fusione seguendo percorsi differenti. La correlazione tra i dati si può verificare anche quando un'uscita del sistema di fusione viene riportata in ingresso allo stesso.
Una soluzione alla correlazione tra i dati potrebbe essere il tentativo di eliminare la *data incest* prima del processo di fusione. Questo in genere impone una conoscenza a priori del problema ed una topologia e dei tempi di comunicazione delle informazioni ben definiti [140], [141], [142].

Un altro approccio prevede la ricostruzione dei dati prima del processo di fusione in modo da renderli decorrelati prima di portarli in ingresso al processo di fusione. Questa soluzione può consentire di trattare scenari di fusione più complicati [143], [144].

4.2.1 Metodi di Covariance Intersection

Il metodo della *Covariance Intersection*, a differenza dei metodi presentati nel paragrafo **Errore. L'origine riferimento non è stata trovata.**, cerca di tener conto della correlazione nel processo di fusione senza tentare di rimuoverla. Il concetto alla sua base è una sottostima del problema di correlazione, che in molti casi può causare delle valutazioni pessimistiche e quindi basse performance dell'algoritmo di fusione.
Per superare alcuni problemi del metodo di *Covariance Intersection* [145] è stato proposto dapprima il metodo *Largest Ellipsoid* [146] e successivamente un ulteriore algoritmo, noto come *Internal Ellipsoid Approximation* [147]. Tutti questi metodi tuttavia sono in grado di realizzare la fusione dei dati esclusivamente mediante l'impiego di tecniche basate sui filtri di Kalman; quindi non sempre possono essere raggiunte elevate performance.
Il metodo di fusione di Chernoff [148] è stato infine proposto per fondere un numero qualsiasi di PDF correlate.

4.3 Fusione di dati inconsistenti

All'interno di questo paragrafo vengono presentati i metodi che consentono la gestione di dati inconsistenti. L'inconsistenza dei dati può consistere nella presenza di dati spuri, dati disordinati o in conflitto tra loro.

4.3.1 Modellazione stocastica adattativa dei sensori per la riduzione degli outliers

In molti casi è possibile che le performance della fusione dei dati vengano ridotte sensibilmente dalla presenza di errori di misura come *spike* o condizioni di *failure* permanenti.
Una strada per affrontare il problema della presenza di dati spuri (*outliers*) è la possibilità di disporre di tecniche di validazione del sensore. Tuttavia tali tecniche richiedono conoscenze preliminari, nonché modelli delle condizioni di *failure*.
Per ovviare a questa limitazione è stato sviluppato un metodo di modellazione stocastico adattativo del sensore del tutto generale, che non richiede quindi conoscenze a priori particolari. Mediante questo metodo si è in grado di gestire opportunamente la presenza di dati spuri [149], [150].

4.3.2 OOSM/OOST Reducing Methods

Ai dati generati dalle sorgenti vengono associati i *timestamp* che rappresentano il momento in cui sono stati prodotti. Questi dati possono arrivare al punto di fusione anche in maniera non ordinata, ad esempio a causa di tempi di propagazione differenti. Questo problema è sempre più importante data la maggior diffusione di sistemi di fusione delle informazioni decentralizzati. Le tecniche più triviali possono essere o lo scarto del dato o una bufferizzazione dei dati. Scartare i dati tuttavia può portare, in alcuni casi, a notevoli perdite di informazioni, mentre bufferizzare i dati può essere causa di problemi computazionali.
In [151] e in altri lavori viene affrontata la realizzazione di algoritmi efficienti per la gestione di *Out Of Sequence Measurement* (OOSM) con ritardi arbitrari.
Un problema analogo all'arrivo di dati disordinati è costituito dagli *Out Of Sequence Track* (OOST). Questa problematica viene affrontata in [123] utilizzando misure equivalenti ottenute a partire dalle tracce registrate dai singoli sensori. Il problema degli OOST è molto meno affrontato in letteratura e solo di recente ha visto un maggiore interesse da parte della comunità scientifica.

4.3.3 Algoritmi *Transferable Belief Model*

La *Dempster-Shafer Theory* può condurre a risultati contraddicenti se applicata in casi di dati in conflitto. Un caso in cui i dati possono essere in conflitto si presenta ad esempio quando diversi esperti asseriscono osservazioni diverse in merito ad uno stesso fenomeno.
Mahler ha dimostrato che questo problema può essere risolto se si rispettano alcuni vincoli. Tuttavia tali vincoli possono essere troppo restrittivi e non fattibili in alcuni casi. Alcuni approcci, come la *Transferable Belief Model* (TBM) [152] e la *Dezert-Smarandache Theory* (DSmT) seguono le indicazioni di Mahler, meno che per il rispetto di alcuni vincoli.

4.3.4 Teoria di Dezert-Smarandache

La teoria *Dezert-Smarandache* (DSmT) del ragionamento possibile e paradossale è la naturale estensione della teoria classica Dempster-Shafer (DST). Tuttavia include delle fondamentali differenze rispetto a quest'ultima. La DSmT consente di combinare formalmente ogni tipo di sorgente di informazione indipendente rappresentata in termini di *belief function*, ma è principalmente orientata alla fusione di sorgenti di evidenza ad alto conflitto, incerte e imprecise dal punto di vista quantitativo e qualitativo.
La DSmT è in grado di risolvere problemi di fusione complessi, statici o dinamici superando i limiti che si possono incontrare con la classica DST. Questo soprattutto nel caso in cui i conflitti tra le sorgenti crescono e quando il raffinamento del problema in considerazione è difficile a causa della vaghezza e imprecisione intrinseca.

4.3.5 Algoritmi di Covariance Union

La fusione di dati in conflitto tra loro viene affrontata dal metodo di *Covariance Union* (CU) sviluppato all'interno di un *framework* bayesiano per essere complementare rispetto al metodo *Covariance Intersection*. L'algoritmo CU consente di trattare la fusione dei dati quando questi sono sia correlati che in conflitto [153].
Di recente è stato presentato un ulteriore approccio bayesiano che consente di trattare dati sia incerti che correlati e in conflitto [154].

4.4 Fusione di dati disparati

In questo paragrafo si presentano dei metodi particolarmente adatti a situazioni in cui i dati in ingresso al sistema di fusione delle informazioni sono particolarmente variegati. Infatti gli input possono spaziare da dati provenienti da sensori *fisici* (radar, sonde di temperatura, etc.) a sensori *umani* (testi tratti dal web, comandi di controllo, etc.).
Gli scenari più recenti legati all'*Information Fusion* vedono quindi, oltre alla fusione di dati provenienti da sensori elettronici classici (*hard data*), la fusione di dati prodotti dall'uomo (*soft data*) espressi in genere in un linguaggio naturale privo di vincoli. Un'applicazione tipica è la *Human Computer Interaction* (HCI) in cui l'uomo può essere visto come un sensore. Alcuni lavori presenti in letteratura sono [154], [155], [156], [157], [158]. Una delle sfide principali è affrontare l'ambiguità dell'interpretazione del linguaggio naturale.
Lo *Human Centered Data Fusion* è un paradigma recente che pone enfasi sull'uomo nel processo di fusione ed in particolare l'uomo non è più considerato come semplicemente un sensore di *soft data*, ma è visto anche come dei computer ibridi. Questa tipologia di paradigma è soprattutto legato allo sviluppo di nuove tecnologie di recente quali i *social network* e i mondi virtuali [159].

4.4.1 Semantic Information Fusion

Questa metodologia di fusione dell'informazione si articola generalmente in due fasi. La prima fase ha lo scopo di aggregare la conoscenza astratta in informazioni semantiche: tale fase si dice di *knowledge base construction*. La seconda conferisce ai risultati della prima fase un'interpretazione semantica. Per effettuare

tale interpretazione semantica, la fase di *pattern matching* lavora sui dati provenienti dalla fase di *knowledge base construction*, fondendo gli attributi che vengono considerati rilevanti.
Un esempio di questa metodologia di fusione viene descritto in [160].

4.4.2 Abductive Reasoning

La teoria dell'*abductive reasoning* [161] più che un metodo di fusione è un *reasoning pattern*. Si basa sul fatto che una volta osservato un evento l'abduzione deriva dalla migliore spiegazione per quel fatto.
Data la regola *a implica b*, l'abduzione prevede che una volta osservata la conseguenza *b* si deriva che la spiegazione di *b* è *a*.
Nel caso invece di ragionamento deduttivo il processo di ragionamento è che una volta osservato il fatto *a* allora si deriva *b* come sua conseguenza.
L'*abductive reasoning* può essere un valido strumento soprattutto a livelli medio-alti del processo di fusione.

4.4.3 Clustering Algorithms (CA)

In letteratura si incontrano diverse famiglie di algoritmi di clustering, particolarmente adatti alla gestione di dati disparati [2]. La prima famiglia è quella dei *Partional Clustering Algorithms* (come ad esempio gli algoritmi *k-means clustering*) mentre la seconda famiglia è quella degli *Hierarchical Clustering Algorithms* (come i *Divisive Hierarchical Clustering* e gli *Agglomerative Hierarchical Clustering*, dove i primi sembrano risultare superiori ai secondi) [162], [163].
Gli autori in [23] evidenziano come il numero di *match* inesatti può essere ridotto adottando un algoritmo di clustering; in tale clustering viene adottato una *Hypercube Distance Measure* e la misura risultante è successivamente confrontata con una *Fuzzy Hamming Distance Measure.*
Altre indicazioni sulle tecniche di *clustering* all'interno del processo di data fusion per il supporto alle decisioni vengono fornite in [164] dove l'applicazione specifica consiste nell'offrire ai geologi gli strumenti per prevedere i possibili *pattern* eruttivi dei vulcani.
La protezione da *cyber* attacchi presentata in [165] viene ottenuta mediante un algoritmo di clustering DHC.

5 Nuove Frontiere

Alcune delle nuove frontiere verso le quali il mondo dell'*Information Fusion* si espande vengono discusse all'interno di questo capitolo.
Un indice di come si stanno evolvendo i sistemi di fusione dell'informazione è il fatto che quello che una volta si considerava *noise* oggi è in realtà il *target*, o comunque una parte essenziale, del processo di fusione. Quanto asserito vale soprattutto nella tematica dei big data, delle smart city data network, wireless sensor network, delle Infrastrutture Critiche, della *Homeland Security*. Ad esempio nella lotta al terrorismo le informazioni vengono ricercate anche all'interno del web dove i siti Internet possono avere livelli di attendibilità bassi se non addirittura sconosciuti o dove comunque le informazioni possono essere facilmente falsificabili. Inoltre, nuovi stimoli alla ricerca nell'ambito dell'*Information Fusion* provengono dagli sviluppi del Web e dei *social network*.

5.1 La fusione nell'emergente net-centric *environment*

Una nuova frontiera per l'*information fusion* consiste nell'uscita da un ambito locale per pervenire ad un ambiente basato sulla rete (*net-centric environment*) [166].
L'obiettivo principale del processo di fusione è rendere le informazioni visibili, accessibili e comprensibili, ottenendo contemporaneamente diversi benefici: manodopera ridotta, migliore interoperabilità, facilità di integrazione, adattabilità/modularità, etc.
In un tale ambiente il problema di "come" trasferire le informazioni è di più facile soluzione rispetto al problema del "cosa" comunicare [166]. In un *net-centric environment* si richiedono l'impiego di applicazioni che lavorino in una *Service Oriented Architecture* (SOA) e di servizi che siano in grado di gestire il contenuto informativo. Ad esempio lo sviluppo di un sistema *enterprise* basato su rete si compone di tre moduli essenziali [166]:

- un'infrastruttura *enterprise*,
- un sistema *data strategy*,
- *mission applications*.

Nel campo di applicazioni militari, gli eserciti moderni stanno concentrando la propria attenzione sull'impiego di tecniche avanzate di analisi, sul *cloud computing* e sull'impiego di tecniche di analisi mediante grafi. Analoga è la situazioni in ambiti Finance e Commerce, così come nelle grandi reti di distribuzione di servizi come quelli informativi, di trasporto, energetici, etc.
Ad oggi l'architettura predominante di tipo *client-server tradizionale*, si sta spostando appunto verso architetture SOA e nello specifico cloud che producono ad esempio i seguenti benefici:

- riduzione delle connessioni in quanto tutte le applicazioni sono connesse al bus enterprise;
- riduzione dei *life-cycle time* dovuto alla drastica riduzione delle connessioni punto-punto;
- non si devono più progettare, realizzare e testare singole connessioni tra le applicazioni;
- il centro di gravità del sistema è costituito dai *Data Center* mentre nelle architetture *client-server* è costituito dai programmi, il che consente una gestione più razionale dei dati.

In [167] viene proposta un *framework* (INFORM Lab) per gestire sistemi distribuiti di *Information Fusion* e *Resource Management* che si basa su una SOA e su un'architettura *Multi-Agent*.
In [168] viene effettuato uno studio delle caratteristiche e delle performance degli algoritmi di fusione distribuiti in diverse condizioni. Tali algoritmi risultano essere una delle ultime frontiere nell'ambito dei sistemi di *Information Fusion* dato l'aumento di sistemi basati sulla collaborazione in rete tra diversi centri di fusione. Ogni centro di fusione infatti riceve dati dai sensori ad esso collegati direttamente ed effettua la fusione con altre informazioni provenienti da altri centri. Le performance dipendono dall'architettura della rete, dalla disponibilità dei link di comunicazione e dall'applicabilità di algoritmi distribuiti. In diversi ambiti

inoltre è necessario che gli algoritmi impiegati siano scalabili, adattandosi così al variare della topologia della rete.

5.2 Analisi dei comportamenti in rete

Tra gli scenari più recenti si trova il tentativo di modellare il comportamento degli utenti all'interno dei *social network*, il modo in cui questi interagiscono e lo stabilirsi di ruoli tra loro. Uno dei possibili campi applicativi in tal senso è la ricerca di comportamenti sospetti all'interno dei *social network*. In tale scenario le informazioni sono spesso frammentate, prive di precise strutture grammaticali o intenzionalmente alterate. In [169] gli autori propongono l'impiego di PCFGs (*Probabilistic Context Free Grammars*) per modellare le azioni degli utenti all'interno di un social network.
In [170] gli autori hanno realizzato uno studio su come scegliere opportunamente un algoritmo per effettuare la rilevazione della presenza di community (*community detection*) all'interno di una rete. La *community detection* può essere utilizzata come vincolo in problemi di inferenza dove la ricerca di attività sospette utilizza appunto la *community detection* per estrarre un sottoinsieme di persone sul quale effettuare osservazioni successive.

5.3 Approccio User-Centered nella gestione dei sensori

L'approccio tradizionale nella scelta di cosa, quando e dove produrre informazioni è seriamente compromesso dalla crescente mole di informazioni di cui si potrebbe disporre e dalla diversità delle sorgenti spesso in campo; un nuovo approccio *User-Centered* tenta di risolvere quindi i problemi posti dal tradizionale approccio *Data-Driven*.
Mentre le tecnologie per la condivisione di informazioni in ambiente distribuito (ad esempio le SOA) si sono sviluppate rapidamente negli ultimi anni, l'ottimizzazione di tali risorse è rimasto ancora poco sviluppato. Gli ultimi sviluppi nell'elaborazione di architetture *Market-Based* (paragrafo 3.2.5.6) hanno appunto posto l'attenzione su come gestire in maniera dinamica la produzione delle informazioni.

5.4 Trend recenti per la fusione negli High-Level

È stato illustrato all'interno del paragrafo 3.3.1 come i dati manipolati possono essere caratterizzati dal loro livello di astrazione; in particolare i livelli più bassi corrispondono ad esempio ai dati provenienti dai sensori, mentre nei livelli più alti del processo di fusione sono coinvolte astrazioni, entità e decisioni più vicine al processo cognitivo umano.
I livelli più alti nell'architettura di un sistema di fusione consentono in particolar modo di supportare gli operatori nel prendere decisioni cruciali in breve tempo e in situazioni di forte stress. Tali livelli tuttavia non sono molto trattati in letteratura e la comunità scientifica sembra rivolgere adesso maggiore attenzione nei loro confronti [171]. In tali livelli è molto forte inoltre la simulazione delle capacità cognitive dell'uomo nell'analizzare le informazioni, eventualmente già risultanti da un processo di fusione *low level*, al fine di prendere decisioni o predire le evoluzioni degli scenari presi in considerazione.
In [172] si evidenzia come il *trend* dello sviluppo di sistemi di fusione qualitativa delle informazioni negli ultimi anni cresca continuamente, soprattutto laddove vi siano più sorgenti informative e si debbano prendere decisioni in situazioni di incertezza intrinseca. Sempre in [172] vengono inoltre presentati gli approcci classici per la creazione di regole di combinazione sia qualitative che quantitative.
I livelli del JDL che vedono maggior lavoro dalla comunità negli ultimi anni sono quindi quelli posti più in alto nell'architettura, in particolar modo i livelli 2 (*situation assessment*) e 3 (*impact assessment*).
Gli scenari sempre più incentrati sulla rete (*net-centric environment*) richiedono un maggiore sforzo nel definire meccanismi di interoperabilità e di fusione più sofisticati. Questo è causato dalla crescente mole di sensori e processi umani che vengono coinvolti all'interno del processo di fusione. Sistemi di tale complessità non possono essere affrontati efficacemente con i soli livelli inferiori di *Information Fusion* in quanto non riescono a supportare lo scambio della conoscenza. In [173] si presenta quindi la necessità di una teoria matematica per i livelli più alti dell'*information fusion* allo scopo di:

- rappresentare concetti semantici e pragmatici;
- realizzare un *framework* matematico al supporto dei processi;
- consentire la scalabilità della soluzione.

Un *Common Operational Picture* (COP) è uno strumento che ha lo scopo di fornire informazioni tempestive e accurate in ambito militare, consentendo la condivisione della conoscenza sulla situazione tra più comandi. I COP del futuro faranno sempre maggior uso di ontologie che necessitano appunto di una nuova e rigorosa formulazione matematica. Si evidenzia inoltre come il numero di ipotesi in sistemi così complessi è talmente elevato da rendere necessario un sistema di *Hypothesis Management* (HM) e tecniche di inferenza ibrida [173]. Anche all'interno di [104] si pone l'attenzione su quali sono le peculiarità dei livelli più alti nell'architettura di un sistema di fusione dell'informazione e su come le ontologie possono essere impiegate nel supportare l'uomo nel prendere decisioni a valle del processo di fusione. Alcuni linguaggi ontologici (ad esempio LOOM e OCML) consentono una completa rappresentazione ontologica e hanno il vantaggio di essere UML-*like*.
Altri *framework* per la rappresentazione della conoscenza si basano sullo standard JSR-94 per l'implementazione in Java di architetture *Object Oriented* (ad esempio DROOLS, ILOG, JRules, Jess and Hammurabi). Tuttavia tale standard non definisce una esatta sintassi [174]. Il framework presentato in [174] lavora con regole che esprimono *first-Order Logic, Fuzzy Logic* e *Temporal Logic.*
Nei livelli più alti del processo di fusione sorgono difficoltà particolari nella rappresentazione di eventi incerti. Queste situazioni possono essere affrontate mediante l'applicazione dell'integrale di Choquet, che può essere utile per definire la probabilità di un evento incerto [175]. L'impiego dell'integrale di Choquet porta diversi benefici al sistema:

- maggiore stabilità e manutenibilità;
- garantisce a nuove informazioni un grado di conoscenza legate al contesto;
- aiuta ad evitare errori nell'integrazione di informazioni attraverso approccio top-down;
- aumenta la confidenza nei processi di *decision-making* mediante calcoli di verosimiglianza;
- consente decisioni accurate e attendibili su eventi incerti;
- limita i risultati attesi all'interno di margini di confidenza.

Sempre in [175] si propone di rappresentare la conoscenza mediante dei *conceptual spaces.*
La maggior parte dei sistemi di fusione presenti in letteratura sono di tipo tradizionale e lavorano con sorgenti che forniscono dati osservati e conoscenze a priori. Altri sistemi di *Information Fusion*, molto meno trattati rispetto ai tradizionali, utilizzano la conoscenza acquisita induttivamente e informazioni contestuali. In campi applicativi quali l'*Homeland Security* i sistemi di *Information Fusion* tradizionali sono molto meno adatti data la complessità dei contesti e delle forze in campo [176].
In [177] viene presentato un sistema che fornisce servizi di apprendimento *context-aware* basandosi su una modellizzazione ontologica del contesto. Lo scenario applicativo vede la realizzazione di un sistema di *ubiquitous learning* che differisce dalle tipiche piattaforme di *learning* basate su architetture *client-server* e sfrutta invece la possibilità di interagire ovunque e in ogni momento con il sistema (Figura 46).

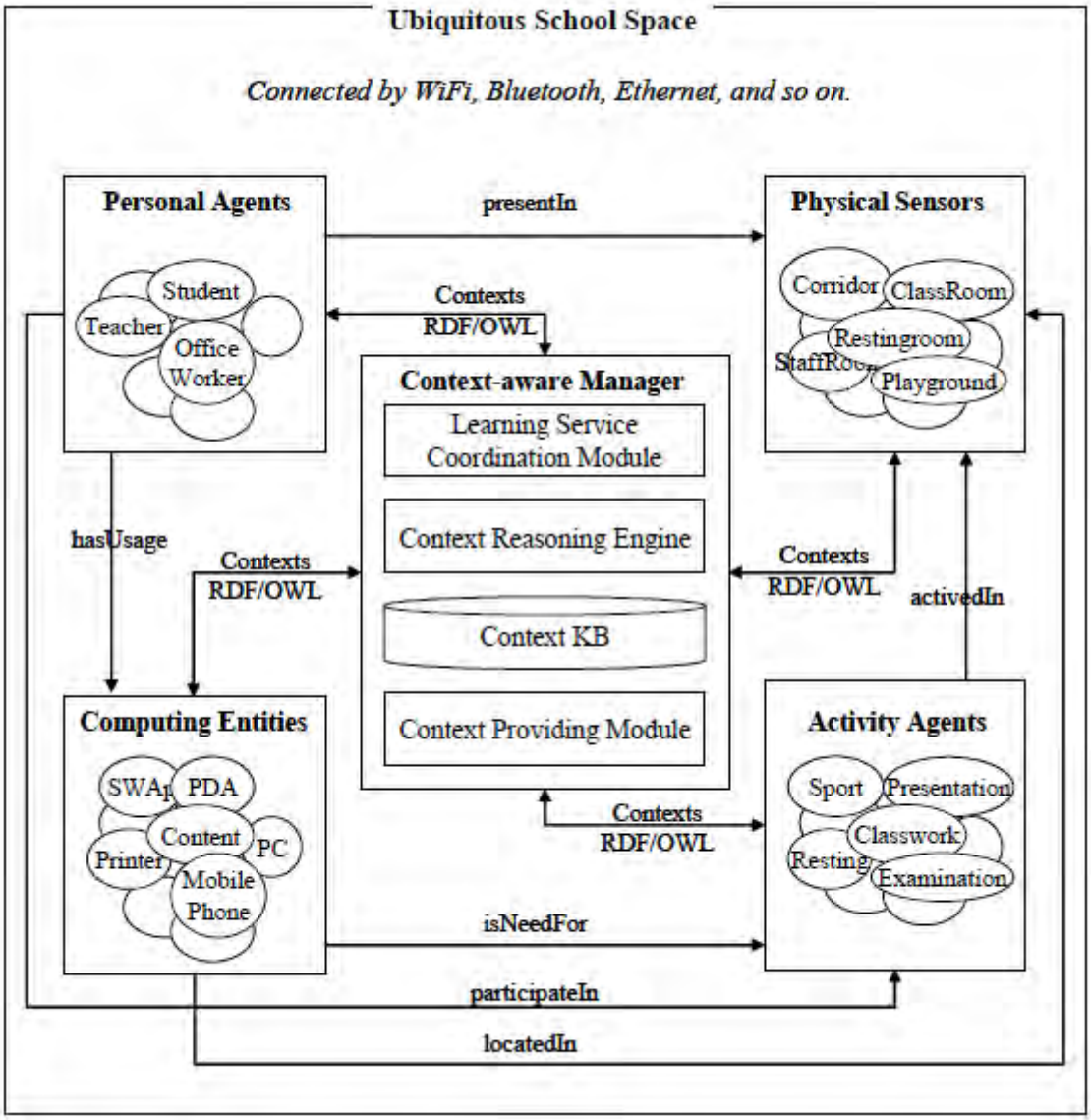

Figura 46. Architettura per il supporto al contesto

6 Conclusioni

E' stato condotto uno studio sui sistemi di *Information Fusion* presentandone i modelli e le tecniche proposte in letteratura utili ad un impiego nel contesto Smart City e Sistema Elettrico.
I modelli sono stati classificati in funzione dell'elemento base del processo di fusione. La progettazione del sistema di fusione può quindi essere guidata dalla scelta del modello che rappresenta in maniera più efficace le entità e le associazioni all'interno del dominio applicativo preso in considerazione. In particolare vi sono alcuni modelli, come il JDL, che sono in grado di rappresentare il processo di fusione in funzione dei livelli di astrazione dei dati manipolati. Altri modelli invece spostano il *focus* sulle attività che saranno svolte all'interno del processo di fusione, mentre una terza famiglia di modelli è in grado di rappresentare il dominio applicativo concentrandosi sulle regole che legano gli attori del processo di fusione. Infine è stata presentata una categoria di modelli in grado di rappresentare efficacemente il processo di acquisizione della consapevolezza della situazione.
È stato inoltre evidenziato come la topologia dell'architettura utilizzata per la realizzazione del sistema di fusione delle informazioni incida sulle prestazioni complessive. Le architetture particolarmente promettenti presenti in letteratura, specie in *net-centric environment*, sono l'architettura olonica o decentralizzata e l'architettura *market-based*.
Infine sono state discusse le varie problematiche legate alla natura dei dati, che i sistemi di fusione devono trattare, analizzando al contempo le metodologie che meglio rispondono a ciascuna problematica. I vari metodi sono stati classificati in funzione della natura dei dati ed è stato evidenziato come non esista un metodo in grado di affrontare contemporaneamente, in maniera efficace, ogni problematica legata ai dati da fondere. In conclusione, la scelta del metodo influisce notevolmente sulle prestazioni del sistema di fusione e in molti casi è possibile l'impiego integrato di più metodi.
Sono state infine discusse alcune delle nuove frontiere della ricerca nel campo dei sistemi di fusione delle informazioni, con particolare riferimento al Web ed al fenomeno del *social networking*.

7 Riferimenti bibliografici

[1] D. Hall, «http://ist.psu.edu/research/centers-and-labs/nc2if/nc2if-page,» [Online].

[2] B. Khaleghi, A. Khamis e F. O. Karray, «Multisensor data fusion: A review of the state-of-the-art,» *Information Fusion,* 2011.

[3] E. F. Nakamura, A. A. F. Loureiro e A. C. Frery, «Information Fusion for Wireless Sensor Networks: Methods, Models, and Classifications,» *ACM Comput. Surv.,* vol. 3, p. XXXIX, 2007.

[4] U.S. Department of Defense, Data Fusion Subpanel of the Joint Directors, «Data fusion lexicon,» 1991.

[5] Klein, L. A. Sensor and data fusion concepts and applications, Tutorial texts, SPIE Optical Engineering Press, 1993.

[6] L. Wald, «Some terms of reference in data fusion,» *IEEE Transactions on Geosciences and Remote Sensing,* pp. 1190-1193, 1999.

[7] V. Dasarathy, «Information Fusion - what, where, why, when, and how?,» *Information Fusion,* vol. II, n. 2, pp. 75-76, 2001.

[8] A. K. Hyder, E. Shahbazian e E. Waltz, «Multisensor Fusion,» 2002. [Online]. Available: http://books.google.it/books?id=7OQsbIX18cEC&printsec=frontcover#v=onepage&q&f=false.

[9] J. Yu, «Information Fusion in Multimedia Information Retrieval: An Overview; Department of Computer Science,» [Online]. Available: http://www.cs.utsa.edu/~qitian/seminar/Spring07/02-16-07/02-16-07.pdf.

[10] D. Whyte, «Sensor models and multisensor integration,» *Inter. J. Robotics Res.,* p. 97–113, December 1988.

[11] W.Elmenreich, «Sensor fusion in time-triggered systems,» *Ph.D. thesis, Vienna University of Technology, Austria,* 2002.

[12] C. Pohl e J. L. Van Genderen, «Multisensor image fusion in remote sensing: Concepts, methods and applications,» *Int. J. Remote Sens.,* p. 823–854, March 1998.

[13] S. S. Iyengar, K. Chakrabarty e H. Qi, «Introduction to special issue on distributed sensor networks for real-time systems with adaptive configuration,» *J. Franklin Inst,* p. 651–653, September 2001.

[14] Nakamura, Figueiredo e Loureiro, «Using information fusion to assist data dissemination in wireless sensor networks. Telecomm,» *Telecommunication Systems,* p. 237–254, November 2005.

[15] R. C. Luo, C.-C. Yih e K. L. Su, «Multisensor fusion and integration: Approaches, applications, and future research directions,» *IEEE Sensors J.,* pp. 107-119, April 2002.

[16] M. E. Q. Hamp1, B. S. Lee3, T. Becker3, D.Wiebeck4 e L. Reindl1, «Information Fusion based on graph analysis during Urban Search and Rescue,» *Information Fusion (FUSION), 2010 13th Conference,* pp. 1 - 7, July 2010.

[17] B. V. Dasarathy, «Sensor Fusion Potential Exploitation—Innovative Architectures and Illustrative Applications,» *Proceedings of the IEEE,* vol. LXXXV, n. 1, pp. 24-38, 1997.

[18] P. Lytrivis, G. Thomaidis e A. Amditis, Sensor and Data Fusion, Vienna: Dr. ir. Nada Milisavljević, 2009.

[19] A. Makarenko, A. Brooks, T. Kaupp, H. Durrant-Whyte e F. Dellaert, «Decentralised Data Fusion: A Graphical Model Approach,» in *12th International Conference on Information Fusion*, Seattle, 2009.

[20] V. Avasarala, T. Mullen e D. Hall, «A Market-based Approach to Sensor Management,» *Journal of Advances in Information Fusion,* vol. IV, n. 1, June 2009.

[21] J. Llinas, R. Nagi, D. Hall e J. Lavery, «A Multi-Disciplinary University Research Initiative in Hard and Soft information fusion: Overview, research strategies and initial results,» in *Information Fusion (FUSION), 2010 13th Conference on*, Edinburgh, 2010.

[22] «Semantic Knowledge Representations,» in *Efficient Decision Support Systems - Practice and Challenges From Current to Future*, Chiang Jao, 2011.

[23] K. Sambhoos, R. Nagi, M. Sudit e T. Rickard, «Hierarchical Higher Level Data Fusion using Fuzzy Hamming and Hypercube Clustering,» *JOURNAL OF ADVANCES IN INFORMATION FUSION,* vol. III, n. 2, December 2008.

[24] B. Forbes e R. Boudjemaa, «Parameter estimation methods for data,» Technical report, National Physical Laboratory, Report CMSC, 2004.

[25] S. Das, High-Level Data Fusion, Norwood, MA: Artech House, Inc., 2008, pp. 5-6.

[26] E. Bosse, A. Guitouni e P. Valin, «An Essay to Characterise Information Fusion Systems,» in *Information Fusion, 2006 9th International Conference on*, 2006.

[27] M. E. Liggins, D. L. Hall e J. Llinas, Handbook of Multisensor Data Fusion. Theory and Practice 2nd Edition, CRC Press, 2009.

[28] P. Smart, «Knowledge-Intensive Fusion for Situational Awareness,» 2005.

[29] M. J. Hall e S. A. Tate, «Removing the HCI Bottleneck: How the Human Computer Interaction (HCI) affects the performance of Data Fusion System,» in *MSS National Sysmposium on Sensor and Data Fusion*, 2000.

[30] E. Blasch e S. Plano, «JDL Level 5 fusion model: user refiinement issues and applications in group tracking,» *SPIE Vol 4729, Aerosense,* pp. 270-279, 2002.

[31] D. L. Hall e S. A. H. McMullen, Mathematical Techniques in Multisensor Data Fusion (2nd ed), Boston, MA: Artech House, 2004.

[32] E. Blasch, «Level 5 (User Refinement) issues supporting Information Fusion Management,» in *9th International Conference on Information Fusion*, 2006.

[33] E. Blasch e S. Plano, «DFIG Level 5 (User Refinement) issues supporting Situational Assessment Reasoning,» in *8th International Conference on Information Fusion*, 2005.

[34] E. Blasch, «Sensor, User, Mission (SUM) Resource Management and Their Interaction with Level 2/3 Fusion,» in *Proceedings of the International Conference of Information Fusion*, Florence, 2006.

[35] E. Blasch, J. Salerno, I. Kadar, J. Biermann, C. Chong e S. Das, «Panel Report - Resource Management Coordination with Level 2/3 Fusion Issues and Challenges,» *Aerospace and Electronic System Magazine,* vol. XXIII, n. 3, pp. 32-46, 2008.

[36] B. V. Dasarathy, «Sensor Fusion Potential Exploitation—Innovative Architectures and Illustrative Applications,» *Proceedings of the IEEE,* vol. LXXXV, n. 1, pp. 24-38, 1997.

[37] A. N. Steinberg e C. L. Bowman, «Revisions to the JDL Data Fusion Model,» in *Handbook of Multisensor Data Fusion (2nd edition)*, CRC Press, 2009, pp. 59-60.

[38] M. Bedworth e J. O'Brien, «The Omnibus Model: A New Model of Data Fusion?,» in *In Proceedings of the 2nd International Conference on Information Fusion (FUSION'99)*, Sunnyvale, 1999.

[39] U.S. Army Intelligence Center, Directorate of Combat Development, Fusion: An Operational Assessmen, Ft. Huachuca, Ariz, July 6, 2004.

[40] J. R. Boyd, «A discourse on winning and losing,» in *Unpublished set of briefing slides available at Air University Library*, Maxwell AFB, Alabama, 1987.

[41] B. Brehmer, «The dynamic OODA loop: amalgamating Boyd's OODA loop and the Cybernetic approach to command and control,» in *Proc. of 10th international command and control research and technology symposium*, 2005.

[42] C. J. Harris e A. D. Bailey, «Multi-sensor data fusion in defence and aerospace,» *The Aeronautical Journal,* pp. 229-244, 1998.

[43] J. Esteban, A. Starr, R. Willetts, P. Hannah e P. Bryanston-Cross, «A review of data fusion models

and architectures: towards engineering guidelines,» *Neural Comput. Appl,* pp. 273-281, 2005.

[44] M. M. Kokar, M. D. Bedworth e C. B. Frankel, «A reference model for data fusion systems,» *Sensor Fusion: Architectures, Algorithms and Applications IV. SPIE,* pp. 191-202, 2000.

[45] C. B. Frankel e M. D. Bedworth, «Control, Estimation and Abstraction in Fusion Architectures: Lessons from Human Information Processing,» in *Proceedings of the Third International Conference on Information Fusion*, 2000.

[46] H. Alizadeh, «Semantic information fusion system based on JDL,» 2011. [Online]. Available: http://www.google.it/url?sa=t&rct=j&q=semantic%20information%20fusion%20system%20based%20on%20jdl&source=web&cd=1&ved=0CC4QFjAA&url=http%3A%2F%2Fwtlab.um.ac.ir%2Fparameters%2Fwtlab%2Ffilemanager%2Fpresentations%2FSemanticinformationfusionsystembasedonJDL..

[47] M. R. Endsley, «Design and evaluation for situation awareness enhancement,» in *In Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA, 1988.

[48] G. A. Klein, «Recognition-primed decisions,» *In W. B. Rouse (Ed.), Advances In man-machine systems research,* vol. V, pp. 47-92, 1989.

[49] J. Roy, «From data fusion to situation analysis,» in *In Proceedings of the 4th International Conference on Information Fusion (Fusion 2001)*, Montreal, 2001.

[50] Kadar, «Knowledge representation issues in perceptual reasoning managed situation assessment,» in *Information Fusion, 2005 8th International Conference on*, 2005.

[51] J. Salerno, D. Boulware e R. Cardillo, «Knowledge representation requirements for situation awareness,» in *Information Fusion, 8th International Conference on*, 2005.

[52] J. J. Salerno, M. Hinman e D. Boulware, «Building A Framework for Situation Awareness,» in *In Proc of the International Society on Information Fusion (ISIF) Conference,*, Stockholm, Sweden, 2004.

[53] G. Tadda, J. J. Salerno, D. Boulware, M. Hinman e S. Gorton, «Realizing situation awareness within a cyber environment,» *In Proceedings of SPIE,* vol. 6242, 2006.

[54] P. Costa, K.-C. Chang, K. Laskey e R. Carvalho, «A Multi-Disciplinary Approach to High Level Fusion in Predictive Situational Awareness,» in *Information Fusion, 2009. FUSION '09. 12th International Conference on* , Seattle, 2009.

[55] H. Durrant-Whyte e T. Henderson, Multisensor data fusion, in: B. Siciliano, O. Khatib (Eds.), Handbook of Robotics, Springer, 2008, pp. 586-610.

[56] H. Pan, Z.-P. Anastasio e T. Andhuang, «A hybrid NN-Bayesian architecture for information fusion,» in *In Proceedings of the International Conference on Image Processing (ICIP'98)*,

Chicago,IL, 1998.

[57] D. Sam, C. Nwankpa e D. Niebur, «Decision fusion of voltage stability indicators for small sized power systems,» in *In IEEE Power Engineering Society Summer Meeting.*, Vancouver, British Columbia, Canada, 2001.

[58] C. Couè, T. Fraichard, P. Bessiere e E. And Mazer, «Multi-sensor data fusion using Bayesian programming: An automotive application,» in *In IEEE/RSJ International Conference on Intelligent Robots and System.*, Lausanne, Switzerland, 2002.

[59] B. Moshiri, M. R. Asharif e R. Hoseinnezhad, «Pseudo information measure: A new concept for extension of Bayesian fusion in robotic map building,» *Inform. Fusion 3,* pp. 51-68, March 2002.

[60] A. Tsymbal, S. Puuronen e D. Patterson, «Ensemble feature selection with the simple Bayesian classification,» *Inform. Fusion 4,* pp. 87-100, June 2003.

[61] J. Allanach, H. Tu, S. singh, P. Willett e K. Pattipati, «Detecting, Tracking and Counteracting Terrorist Networks via Hidden Markov Models,» in *IEEE Aerospace Conference, Big Sky MT*, 2004.

[62] S. Marcel e J. d. R. Millàn, «Person Authentication Using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation,» *IEEE Transactions on Pattern Analysis and Machine Intelligence,* pp. 743-752, April 2007.

[63] S. W. SMITH, The Scientist and Engineer's Guide to Digital Signal Processing. 2nd ed, San Diego, CA: California Technical Publishing, 1999.

[64] C.-L. B. S. A. C. W. J. Yang, «Location tracking with directional antennas in wireless sensor networks,» in *In 2005 IEEE MTT-S International Microwave Symposium Digest.*, Long Beach, CA, 2005.

[65] S. Rangwala, R. Gummadi, R. Govindan e K. Psounis, «Interference-aware fair rate control in wireless sensor networks,» in *In Proceedings of the 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM'06)*, Pisa, Italy, 2006.

[66] R. Niu e P. K. Varshney, «Target location estimation in sensor networks with quantized data,» *IEEE Trans. Sig. Proc.,* pp. 4519-4528, December 2006.

[67] X. Sheng e Y.-H. Hu, «Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks,» *IEEE Trans. Sig. Proc. 53,* pp. 44-53, January 2005.

[68] J. Liu, Y. Zhang e F. Zhao, «Robust distributed node localization with error management,» in *In Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'06)*, Florence, Italy, 2006.

[69] L. Xiao, S. Boyd e S. Lall, «A space-time diffusion scheme for peer-to-peer least-squares

estimation.,» in *In Proceedings of the 5th International Conference on Information Processing in Sensor Networks (IPSN'06),*, Nashville, TN, 2006.

[70] A. Doucet, J. F. G. d. Freitas e N. J. Gordon, An introdution to sequential Monte Carlo methods, Springer-Verlag, 2001.

[71] A. Doucet, S. Godsill e C. Andrieu, On sequential monte carlo sampling methods for bayesian filtering, Statist. Comput., 2001.

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

[73] W. Gilks, S. Richardson e D. Spie, Markov Chain Monte Carlo in Practice, London, UK: Chapman & Hall/CRC, 1996.

[74] M. Isard e A. Blake, «Contour tracking by stochastic propagation of conditional density,» in *In 4th European Conference on Computer Vision (ECCV'96), B. F. Buxton and R. Cipolla, Eds. Lecture Notes in Computer Science, vol. 1064.*, Cambridge, UK, Springer, 1996, pp. 343-356.

[75] P. Perez, J. Vermaak e A. Blake, «Data Fusion for Visual Tracking With Particles,» *Proceedings of the IEEE,* vol. 92, n. 3, 2004.

[76] A. Doucet, B.-N. Vo, C. Andrieu e M. Davy, «Particle filtering for multi-target tracking and sensor management.,» in *In Proceedings of the 5th International Conference on Information Fusion (Fusion 2002)*, Annapolis, Maryland, 2002.

[77] F. Gunnarsson e F. Gustafsson, «Positioning using time-difference of arrival measurements,» in *In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, 2003.

[78] M. Pitt e N. Shephard, «Filtering via simulation: auxiliary particle filters,» *Journal of the American Statistical Association,* pp. 590-599, 1999.

[79] B. Berg, Markov Chain Monte Carlo Simulations and Their Statistical, Singapore: World Scientific, 2004.

[80] S. U. N. Metropolis, «The Monte Carlo method,» *Journal of American Statistics Association,* pp. 335-341, 1949.

[81] W. Hastings, «Monte Carlo sampling methods using markov chains and their applications,» *Biometrika,* pp. 97-109, 1970.

[82] Kalman, «A new approach to linear filtering and prediction problems,» *Trans. ASME J. Basic Engin.,* pp. 35-45, 1960.

[83] G. Bishop e G. Welch, An Introduction to the Kalman Filter, University of North Carolina at Chapel

Hill, Department of Computer Science, 2006.

[84] G. Powell, D. Marshall, R. Milliken e K. Markham, «A Data Fusion System for Object Recognition based on Transferable Belief Models and Kalman Filters,» in *Proceedings of the Seventh International Conference on Information Fusion*, 2004.

[85] C. Mascolo e M. Musolesi, «SCAR: Context-aware adaptive routing in delay tolerant mobile sensor networks.,» in *In Proceeding of the 2006 International Conference on Communications and Mobile Computing (IWCMC'06)*, Vancouver, Canada, 2006.

[86] G. Welch e G. Bishop, «An introduction to the Kalman filter,» in *In SIGGRAPH 2001 Course Notes*, Los Angeles, CA, 2001.

[87] S. J. Julier e J. K. Uhlmann, «New extension of the Kalman filter to nonlinear systems,» vol. VI, n. 3068, p. 182–193, 1997.

[88] S. Ci e H. Sharif, «Performance comparison of kalman filter based approaches for energy efficiency in wireless sensor networks,» in *In Proceedings of the 3rd ACS/IEEE International Conference on Computer Systems and Applications (AICCSA'05). IEEE*, Cairo, Egypt, 2005.

[89] L. Stone, T. Corwin e C. Barlow, Bayesian Multiple Target Tracking, Norwood, MA: Artech, 1999.

[90] L. Rabiner e B. Juang, «An introduction to hidden Markov models,» *IEEE ASSP Magazine,* pp. 4-16, 1986.

[91] M. Andersson, J. Rydell e J. Ahlberg, «Estimation of crowd behaviour using sensor networks and sensor fusion,» in *12th International Conference on Information*, Seattle, WA, USA, 2009.

[92] J. Allanach, H. Tu, F. Tu e S. Singh, «Detecting, Tracking, and Counteracting Terrorist Networks via Hidden Markov Models,» in *Proceedings of the IEEE Aerospace Conference*, 2004.

[93] B. Khaleghi, S. N. Razavi, A. Khamis, F. O. Karray e M. Kamel, «Multisensor Data Fusion: Antecedents and Directions,» in *International Conference on Signals, Circuits and Systems*, 2009.

[94] L. Zadeh, «Fuzzy sets,» *Information and Control,* pp. 338-353, 1965.

[95] L. Zadeh, Fuzzy logic and approximate reasoning, 1975.

[96] W. C. Yet e U. Qidwai, «Intelligent sensor network for obstacle avoidance strategy.,» in *In Proceedings of the 4th IEEE Conference on Sensors.*, Irvine, 2005.

[97] I. Gupta, D. iordan e S. Sampalli, «Cluster-head election using fuzzy logic for wireless sensornetworks.,» in *In Proceedings of the 3rd Annual Communication Networks and Services Research Conference (CNSR'05).*, Halifax, Canada, 2005.

[98] M. N. Halgamuge, S. M. Guru e A. Jennings, «Energy efficient cluster formation in wireless sensor networks,» in *In Proceedings of the 10th International Conference on Telecommunications*

(ICT'03)., Papeete, French Polynesia, 2003.

[99] P. Escamilla-Ambrosio e N. Mort, «Hybrid Kalman filter-fuzzy logic adaptive,» *Proc. of the IEEE Conference on Decision and Control,* pp. 5215-5220, 2003.

[100] O. B. A. Zhu, «A novel fuzzy evidential reasoning paradigm for data fusion with applications in image processing,» *Soft Computing Journal,* vol. X, n. 12, pp. 1169-1180, 2006.

[101] S. Chandana, H. Leung, E. Bosse e P. Valin, «Fuzzy Cognitive Map based Situation Assessment for Coastal Surveillance,» in *Information Fusion, 2008 11th International Conference on* , Cologne, 2008.

[102] A. Siraj, R. B. Vaughn e S. M. Bridges, «Intrusion Sensor Data Fusion in an Intelligent Intrusion Detection System Architecture,» in *Proceedings of the 37th Hawaii International Conference on System Sciences*, 2004.

[103] H. Yue, G. Yue e G. Yi, «Application Study in Decision Support with Fuzzy Cognitive Map,» *INTERNATIONAL JOURNAL OF COMPUTERS,* vol. I, n. 3, 2007.

[104] E. P. Blasch e S. Plano, «Ontological Issues in Higher Levels of Information Fusion: User Refinement of the Fusion Process,» in *Proceedings of the 6th International Conference on International Fusion*, Cairns, Queensland, 2003.

[105] J. N. Liu, «Fuzzy Ontology Based System for Product Management and Recommendation,» *INTERNATIONAL JOURNAL OF COMPUTERS,* vol. I, n. 3, 2007.

[106] F. Rosenblatt, Two theorems of statistical separability in the perceptron. In Mechanization of Thought Processes, London, UK: National Physical Laboratory, 1956.

[107] B. Widrow e M. E. Hoff, Adaptive switching circuits, 1960 IRE Western Electric Show and Convention Record, 1960.

[108] R. H. Baran, «A collective computation approach to automatic target recognition,» in *In Proceedings of the International Joint Conference on Neural Networks*, Washington, D.C., 1989.

[109] A. Filippidis, L. C. Jain e N. Martin, «Fusion of intelligent agents for the detection of aircraft in SAR images,» *IEEE Trans. Patt. Anal. Mach. Intell.,* vol. XXII, n. 4, 2000.

[110] T. W. Lewis e D. M. W. Powers, «Audio-visual speech recognition using red exclusion and neural networks,» in *In Proceedings of the 25th Australasian Conference on Computer Science. Australian Computer Society, Inc.*, Melbourne, Victoria, Australia, 2002.

[111] L. Yiyao, Y. V. Venkatesh e C. C. Ko, «A knowledge-based neural network for fusing edge maps of multi-sensor images,» *Inform. Fusion,* vol. II, n. 2, pp. 121-133, 2001.

[112] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Norwell, MA: Academic

Publishers, 1992.

[113] L. Paolotti, «Pianificazione multiobiettivinelle aziende agrarie: l'approccio dei RoughSetsbasato sulla relazione di dominanza,» luglio 2008. [Online]. Available: http://ilo.unimol.it/sidea/images/upload/SE/dottori/Paolotti.pdf.

[114] S. Greco, B. Matarazzo e R. Słowiński, «Rough sets theory for multicriteriadecision analysis,» *European Journal of Operational Research,* vol. 129, n. 1, pp. 1-47, 2001.

[115] W. Haijun e 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.

[116] L. Zadeh, «Fuzzy sets as a basis for a theory of possibility,» *Fuzzy Sets and Systems,* 1978.

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

[118] D. Dubois e H. Prade, «Possibility theory in information fusion,» in *Proc. of the*, 2000.

[119] H. Borotschnig, L. Paletta, M. Prantl e A. Pinz, «Comparison of probabilistic, possibilistic and evidence theoretic fusion schemes for active object recognition,» *Computing,* pp. 293-319, 1999.

[120] S. Benferhat e C. Sossai, «Reasoning with multiple-source information in a possibilistic logic framework,» *Information Fusion,* pp. 80-96, 2006.

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

[122] A. Dempster, «A generalization of Bayesian inference,» *J. Royal Stat. Soc., Series B 30,* 1968.

[123] S. Challa e D. Koks, «Bayesian and Dempster–Shafer fusion,» *Sadhana,* pp. 145-176, April 2004.

[124] B. Bracio, W. Horn e D. Moller, «Sensor fusion in biomedical systems,» *Proc. of Annual International Conference of the IEEE Engineering in Medicine,* pp. 1387-1390, 1997.

[125] A. Benavoli, B. Ristic, A. Farina, M. Oxenham e L. Chisci, «An approach to threat assessment based on evidential networks,» *Proc. of the International,* pp. 1-8, 2007.

[126] P. Shenoy e G. Shafer, «Axioms for probability and belief-function,» in *Classic Works of the Dempster–*, Berlin, Springer, 2008.

[127] R. Yager, «Generalized probabilities of fuzzy events from fuzzy belief structures,» *Information Sciences,* pp. 45-62, 1982.

[128] F. K. Z. H. O. Basir, «Connectionist-based Dempster–Shafer evidential reasoning for data fusion,» *IEEE Transactions on Neural Networks,* pp. 1513-1530, 2005.

[129] D. S. Yeung, C. Degang, E. Tsang, J. Lee e X. Wang, «On the generalization of fuzzy rough sets,» *IEEE Transactions on Fuzzy Systems,* pp. 343-361, 2005.

[130] D. Dubois e H. Prade, «Rough fuzzy sets and fuzzy rough sets,» *International Journal of General Systems,* pp. 191-209, 1990.

[131] D. Yeung, C. Degan, E. C. C. Tsang, J. W. T. Lee e X. Wang, «On the generalization of fuzzy rough sets,» *IEEE Transactions on Fuzzy Systems,* vol. VI, n. 6, pp. 1513-1530, 2005.

[132] T. Guan e B. Feng, «Rough fuzzy integrals for information fusion and classification,» *Lecture Notes in Computer Science,* n. 3066, pp. 362-367, 2004.

[133] J. Yen, «Generalizing the Dempster–Shafer theory to fuzzy sets,» *IEEE Transactions on SMC,* pp. 559-570, 1990.

[134] O. Basir, F. Karray e Z. Hongwei, «Connectionist-based Dempster-Shafer evidential reasoning for data fusion,» *IEEE Transactions on Neural Networks ,* vol. VI, n. 6, pp. 1513-1530, 2005.

[135] R. Mahler, Statistical Multisource-Multitarget Information Fusion, Boston, MA: Artech House, 2007.

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

[137] M. Maehlisch, R. Schweiger, W. Ritter e K. Dietmayer, «Multisensor vehicle tracking with the probability hypothesis density filter,» in *Proc. of the International Conference on Information Fusion*, 2006.

[138] M. Florea, A. Jousselme e 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., 2007.

[139] B. Khaleghi, A. Khamis e 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.

[140] S. McLaughlin, R. Evans e V. Krishnamurthy, «Data incest removal in a survivable estimation fusion architecture,» *Proc. of the International Conference on Information Fusion,* 2003.

[141] S. McLaughlin, R. Evans e V. Krishnamurthy, «A graph theoretic approach to data incest management in network centric warfare,» *Proc. of the International Conference on Information Fusion,* pp. 1162-1169, 2005.

[142] T. Brehard e V. Krishnamurthy, «Optimal data incest removal in Bayesian decentralized estimation over a sensor network,» *IEEE International Conference on Acoustics, Speech and Signal*

Processing,, 2007.

[143] L. Pao e M. Kalandros, «Algorithms for a class of distributed architecture tracking,» *Proc. of the American Control Conference,* pp. 1434-1438, 1997.

[144] L. P. W. Khawsuk, «Decorrelated state estimation for distributed tracking of interacting targets in cluttered environments,» in *Proc. of the American Control Conference*, 2002.

[145] J. U. S.J. Julier, «A non-divergent algorithm in the presence of unknown correlation,» *Proc. of the American Control Conference,* p. 2369–2373, 1997.

[146] A. Benaskeur, «Consistent fusion of correlated data sources,» *Proc. of IEEE Annual Conference of the Industrial Electronics Society,* p. 2652–2656, 2002.

[147] Y. Zhou e J. Li, «Robust decentralized data fusion based on internal ellipsoid approximation,» *Proc. of World Congress of the International Federation of Automatic Control,* p. 9964–9969, 2008.

[148] W. Farrell e C. Ganesh, «Generalized chernoff fusion approximation for practical distributed data fusion,» *Proc. of the International Conference on Information Fusion,* p. 555–562, 2009.

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

[150] M. Kumar, D. Garg e R. Zachery, «A method for judicious fusion of inconsistent multiple sensor data,» *IEEE Sensors Journal,* p. 723–733, 2007.

[151] U. Orguner e F. Gustafsson, «Storage efficient particle filters for the out of sequence measurement problem,» *Proc. of the International Conference on Information Fusion,* pp. 1-8, 2008.

[152] P. Smets, «The combination of evidence in the transferable belief model,» *IEEE Transactions on Pattern Analysis and Machince Intelligence,* p. 447–458, 1990.

[153] J. K. Uhlman, «Covariance consistency methods for fault-tolerant distributed data fusion,» *Information Fusion,* vol. III, n. 3, pp. 201-215, 2003.

[154] S. Maskell, «A Bayesian approach to fusing uncertain, imprecise and conflicting information,» *Information Fusion,* p. 259–277, 2008.

[155] D. Hall, M. McNeese, J. Llinas e T. Mullen, «A framework for dynamic hard/soft fusion,» *Proc. of the International Conference on Information Fusion,* pp. 1-8, 2008.

[156] M. Pravia, O. Babko-Malaya, M. Schneider, J. White e C. Chong, «Lessons learned in the creation of a data set for hard/soft information fusion,» *Proc. of the International Conference on Information Fusion,* pp. 2114-2121, 2009.

[157] M. Pravia, R. Prasanth, P. Arambel, C. Sidner e C. Chong, «Generation of a fundamental data set for hard/soft information fusion,» *Proc. of the International Conference on Information Fusion,* pp. 1-8, 2008.

[158] A. Auger e J. Roy, «Expression of uncertainty in linguistic data,» *Proc. of the International Conference on Information Fusion,* pp. 1-8, 2008.

[159] D. Hall, M. McNeese, D. Hellar, B. Panulla e W. Shumaker, «A cyber infrastructure for evaluating the performance of human centered fusion,» *Proc. of the International Conference on Information Fusion,* pp. 1257-1264, 2009.

[160] D. S. Friedlander e S. Phoha, «Semantic information fusion for coordinated signal processing in mobile sensor networks,» *Int. J. High Perf. Comput.,* pp. 235-241, Fall 2002.

[161] C. S. Peirce, «Abduction and induction,» in *Philosophical Writings of Peirce*, New York, Eds. Dover, 1955, pp. 150-156.

[162] A. K. Jain, M. N. Murty e P. J. Flynn, «Data clustering,» *ACM Computing Surveys (CSUR),* vol. XXXI, n. 3, pp. 264-323, 1999.

[163] M. E. J. Newman e M. Girvan, «Finding and evaluating community structure in networks,» *Phisical Review E,* vol. LXIX, 2004.

[164] G. L. Rogova, M. I. Bursik e S. Hanson-Hedgecock, «Intelligent System for Interpreting the Pattern of Volcanic Eruptions,» *JOURNAL OF ADVANCES IN INFORMATION FUSION,* vol. III, n. 2, December 2008.

[165] C. T. Murphy e S. J. Yang, «Clustering of Multistage Cyber Attacks using Significant Services,» in *Information Fusion (FUSION), 2010 13th Conference on*, Edinburgh, 2010.

[166] F. E. White, «Fusion Update: What Will It Take To Make It Work Introduction to Data and Information Fusion and Current Status,» in *NDIA Information Dominance Conference*, October 2008.

[167] A. Guitouni, P. Valin, H. Wehn e J. Happe, «High-level Closed-loop Fusion and Decision Making with INFORM lab,» in *Presented at the 16th International Command and Control Research and Technology Symposium*, Quebec City, 2011.

[168] K. Chang, C.-Y. Chong e S. Mori, «Analytical and Computational Evaluation of Scalable Distributed Fusion Algorithms,» vol. XLVI, n. 4, October 2010.

[169] S. C. Geyik, J. Xie e B. K. Szymanski, «Behavior Modeling with Probabilistic Context Free Grammars,» in *Information Fusion (FUSION), 2010 13th Conference on* , Edinburgh, 2010.

[170] L. Peel, «Estimating network parameters for selecting community detection algorithms,» in *Information Fusion (FUSION), 2010 13th Conference on* , Edinburgh, 2010.

[171] M. Nilsson e T. Ziemke, «Information Fusion: a Decision Support Perspective,» in *Information Fusion, 2007 10th International Conference on* , 2007.

[172] A. Martin, C. Osswald, J. Dezert e F. Smarandache, «General Combination Rules for Qualitative and Quantitative Beliefs,» *JOURNAL OF ADVANCES IN INFORMATION FUSION,* vol. III, n. 2, December 2008.

[173] P. C. G. Costa, K.-C. Chang, K. Laskey, T. Levitt e W. Sun, «High-Level Fusion: Issues in Developing a Formal Theory,» in *Information Fusion (FUSION), 2010 13th Conference on*, Edinburgh, 2010.

[174] J. G. d. C. Jr., P. Rangel e N. F. Ebeken, «A multi-logic framework for multi-level fusion in real time,» in *Information Fusion (FUSION), 2010 13th Conference on*, Edinburgh, 2010.

[175] T. Schuck e E. Blash, «Description of the Choquet Integral for Tactical Knowledge Representation,» in *Information Fusion (FUSION), 2010 13th Conference on* , Edinburgh, 2010.

[176] J. Gómez-Romero, J. García, M. Kandefer, J. Llinas, J. Molina, M. Patricio, M. Prentice e S. Shapiro, «Strategies and techniques for use and exploitation of Contextual Information in high-level fusion architectures,» in *Information Fusion (FUSION), 2010 13th Conference on* , Edinburgh, 2010.

[177] M.-w. Hong e D.-j. Cho, «Ontology Context Model for Context-Aware Learning Service in Ubiquitous Learning Environments,» *INTERNATIONAL JOURNAL OF COMPUTERS,* vol. II, n. 3, 2008.