Modern Principles of Identity Fusion

Thomas Kausch **Felix Opitz**
Defence & Communication Systems/Air and Naval Defence
EADS Deutschland GmbH
Wörthstr. 85, 89077 Ulm, Germany

thomas.kausch@eads.com felix.opitz@eads.com

ABSTRACT

This paper addresses the principles of identity fusion used in modern air defence systems. In the last years identity fusion gets more and more attention. The new proposed STANAG 4162 (Identification Data Combining Process) is exclusively devoted to identity fusion. The performance of the identity fusion process renders the opportunity to increase the overall system situation awareness. As a consequence the spatial range and timeliness of engagement capability can be optimised. Further it is of central importance with respect to interoperability and network feasibility of the current and future air defence systems.

Different sources of identity related data exist: The data may be measured by primary radar, e.g. high-range-resolution, engine modulation, cross-section fluctuations, polarizations, inverse scattering. Or they stem from other sensors, especially secondary radar (IFF) or ESM systems. The target trajectory itself gives a very relevant identity indication and there are also doctrine based aids. Further, modern air defence systems are required to be highly interoperable. There exist a multitude of different data link systems, which are commonly integrated within air defence. Therefore link information is another important source for identity fusion.

However, the identity related data may be contradicting, paradoxical and uncertain. To deal with these effects, several approaches to the identity fusion problem have been developed. The most popular techniques are the Bayesian method and the Dempster-Shafer Theory. Newer considerations are concerned with the benefit using the recently founded Dezert-Smarandache Theory. Other approaches are based on rule or voting based concepts, on expert systems or fuzzy logic.

The identity fusion must be understood as a component integrated in the data fusion concept of an air defence system, which is commonly described by the JDL model. Hence, identity fusion must be integrated and aligned with special overall data fusion architecture concepts. The common dependencies and impacts between the so called position fusion (i.e. data association and filtering) and identity fusion have to be analysed carefully. Identity fusion requirements must be in accordance with the threat evaluation and weapon assignment component of an air defence system. Finally, the identity fusion has also to take into account standardisation agreements - like STANAG 5516 for LINK 16 – to ensure interoperability and future network feasibility.

1.0 IDENTITY FUSION IN THE DATA FUSION CONTEXT

Identification and classification of objects in context with air surveillance operations is a task which can no longer be viewed as something which is operating independently or sequentially with other tasks of an air surveillance system, namely the target tracking, tactical situation assessment, threat evaluation and other mission depending tasks. As the amount of information available increases more and more a

Kausch, T.; Opitz, F. (2005) Modern Principles of Identity Fusion. In *Design Considerations and Technologies for Air Defence Systems* (pp. 11-1 – 11-14). Meeting Proceedings RTO-MP-SCI-143, Paper 11. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

common structure for data fusion was introduced to create a framework for building and understanding a data fusion system with all its inherent dependencies. In

Figure 1 the data fusion process model of the Joint Directors of Laboratories (JDL) is shown. The whole process is divided into six different levels [1]:

- Level 0: Source Pre-Processing
  Low level processing and filtering by the sensors
- Level 1: Object Refinement
  Fusion of data from several sensor to determine kinematics and identity of an object
- Level 2: Situation Refinement
  Refinement of the estimated situation by automated reasoning, determination of relationships between objects
- Level 3: Impact Refinement
  Create hypothesis of possible threats and future conditions
- Level 4: Process Refinement
  Monitoring of the data fusion process to improve processing results
- Level 5: Cognitive Refinement
  Improvement of the interpretation of the results by interaction of the data fusion system with the operator

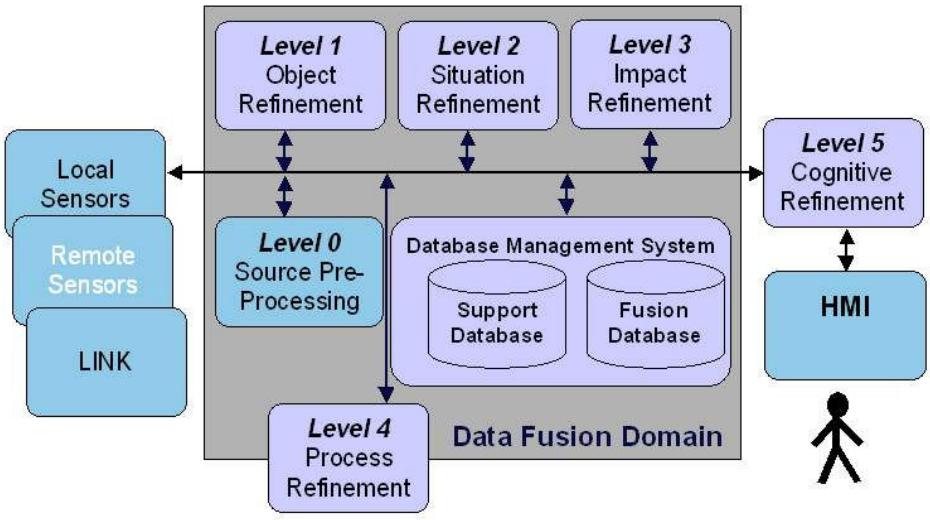

Figure 1 The JDL Fusion Model [1]

2.0 SOURCES OF IDENTITY

Information suitable for the identity finding process may came from sensor measurements, target dynamics and trajectory, or Link data. In the following paragraphs we will give an overview of this

information types

2.1 Sensor Information

Common to all sensor sources is the information about the position of a target and the time this measurement is made. This information is the most important part of sensor information since all possible identification has to be associated with a real object. This kind of information is linked to the JDL-level 0. Here we give a rough overview over the information one can expect from different kind of sensors.

2.1.1 Radar

The information of a radar can be divided into different parts. First, there is the high level information which result from a radar internal analysis of the received signals.

These information may be

- the position of the target either in 2D or 3D
- Doppler
- the tactical environment,
- some kind of classification, where the possible target type is given (e.g. jet aircraft, helicopter, missile,…),
- jamming,

and also the non-existence of a target response in a search volume can be used as important information.

Other so-called low-level information, which strongly depend on the capability of a sensor are:

- high-range-resolution or cross-section fluctuations.
- engine modulation caused by turbines or propellers of aircrafts or by hubs and blades of helicopters which can be used as a source of target classification.
- Other sources may be polarizations or inverse scattering.

These enumerations are far from being complete but give a rough hint on the huge amount of information that can be expected when using a modern radar system.

It has to be stressed that all this information is helpful in describing the real object behind these measurements but that every measurement has some inherent uncertainty incorporated.

2.1.2 IFF

For cooperative targets, the source of identity is the secondary radar or IFF, but even for those targets their answers may be corrupted through superposition, whenever they belong to the same segment. Sometimes the answers cannot be separated with respect to the targets involved, an effect known as garbling. Targets may also be queried by side lobes, which hamper the precise azimuth measurement. The azimuth measurement may be corrupted by delays in the different transponders. The own IFF interrogator may also receive answers caused by other IFF interrogators, an effect which is called fruit.

The altitude information depends on the correct setting of the barometric altimeter in the targets. Other targets or objects may reflect IFF, so that it is difficult to distinguish between a real and a mirror target. The IFF antenna of the target may be covered by the target itself. Hence, missed detections may be

experienced depending on the targets aspect angle during movement or manoeuvres.

2.1.3 ESM

The great advantage of measurements made by ESM is that they are done in a passive way. This means that this kind of source gives information about the Ids of targets in the operational area also in case of radar silence. An other factor that makes the ESM measurements of great value is the fact that those sensors are experts in analysing the incoming signal (waveform, frequency, PRF, …) in a way that it is possible to get information about the platform emitting this signal. The results of this evaluation may also be used to get a hint about the operational task of an unknown platform. E.g. when detecting a lock-on of a fire control radar will make one reconsider about the allegiance of the emitter.

This task is allocated to the JDL-level 1, but has also significant implications for JDL-level 2 (time relationships and functional dependence of objects, the time-event link) and 3 (prediction of enemy intent and to identify threat opportunities).

2.2 External Data

A huge amount of external data sources is available to feed the identification process. Some of them are

- the tactical data links like Link11, Link16, Link22, …, proprietary links (see e.g. STANAG 5511, 5516, 5522)
- intelligence data
- SAR pictures

This feature have implications for all JDL-levels.

2.3 Target dynamics, trajectory and behaviour

The trajectory and dynamics are derived from the position and Doppler measurements made by the local sensor and in addition from the information received from external sources by link. The trajectory contains important information about the target capabilities and the flown manoeuvre profiles giving a hint on the flight mission and target class. The following enumeration of useful features deducible from the target dynamics and trajectory shows how important it is to have an accurate kinematics estimation of the target state with correct association of different source of information to support the task of the identification fusion. This is part of the JDL-level 1.

2.3.1 Dynamics

Different dynamic properties are useful to get constraints for the determination of the class to which a target belongs. Some of them are

- crossing of some acceleration thresholds which are typical for specific target classes
- height thresholds
- velocity thresholds.

Often the tracking component of the sensor data fusion can provide internal data from its filter process which are useful for the identification and classification. These may be the conditional probability of the used manoeuvre models, the parameters of some specialized manoeuvre models like the cross accelerations or the change in the vertical velocity. Also the association weight may be valuable to weight

the contribution of some target attributes in the identification process.

With these features also the search in the intelligence database can be refined to minimize the number of possible target types to get a platform classification. If the resultant platform classification is unique it can be used to get information about the allegiance of the target and therefore about the identity.

2.3.2 Trajectory

The target trajectory gives information about the path the target followed up to a certain time. Also from this data valid conclusions can be drawn. Some of these are

- the origin (under the assumption the take-off is observed), to deduce the identification by origin (IBDO)
- flight envelope.

With the flight profiles, different manoeuvres can be identified and compared with typical attack profiles, typical flight envelopes of missiles (weaving target or high diver profile,…) or mission profiles.

Comparisons of the trajectory data with civil or military flight plans or with intelligence data provide insight into the identity and possibly the mission.

2.3.3 ECM

Target exhibiting electronic counter measurements attempt to camouflage their own existence. Even when all other information is ambiguous these signal characteristics can be used to "tune" the own sensors and to deal with this targets properties or to use information other than that of the own disturbed sensors (JDL-level 4). The kind of ECM used by the target can be used for identification purposes.

3.0 IDENTITY FUSION ALGORITHMS

Many different proposals exist to describe the combination or fusion of different information to get an identification and classification of an object. The most common algorithms are based on Dempster-Shafer or the Bayes approach. The latter one is used in STANAG 4162 (Identification Data Combining Process) to find a standardized formulation for the identification process. Recently the Dempster Shafer theory was reformulated to avoid the known paradoxes found in this theory. This reformulation led to the foundation of the Dezert-Smarandache Theory. In the following sections we give a short introduction to these approaches and describe their pros and cons.

3.1 Bayes inference

3.1.1 Theory

The power of the very general Bayesian inference approach is the possibility to derive the probability $P(H_i|E_j)$ of a hypothesis H_i given some evidence E_j from the capacity of the source of information, i.e. the probability $P(E_j|H_i)$ of getting some evidence E given that H_i is true. This relation is called the Bayes theorem:

$$P(H_i|E_j) = \frac{P(E_j|H_i)P(H_i)}{\sum_i P(E_j|H_i)P(H_i)}$$

Here $P(H_i)$ is the a priori probability that the Hypothesis H_i is true. The set of all possible hypotheses

$\Theta=\{H_1, H_2,\ldots, H_n\}$ has to be mutually exclusive and exhaustive, which results in following relations:

$$\sum_i P(H_i) = 1$$

$$H_i \cap H_j = \varnothing \text{ for } i \neq j$$

With a set of evidences, which is the standard in an multisensor environment, the probability of hypothesis H_i given a set of evidences, i.e. observations, for fixed j is

$$P(H_i \mid E_j) \propto P(H_i) \prod_j P(E_j \mid H_i)$$

when the assumption of statistical independency of the evidences $E_1,\ldots, E_m$ can be made. The decision which hypothesis is preferred at the end of the processing can be made in different ways:

- selection of the maximum posteriori probability (MAP)
- estimation of the wrong decision risk

3.1.2 IDCP as an application of the Bayes' Theorem

In the STANG 4162 the framework is described for the identification data combining process which is based on the Bayes' Theorem.

At the beginning of each measurement evaluation from a source is the so called source probability matrix (SPM) which expresses the performance of the source. It contains the probability $P(E_j|H_i)$, representing the probability of the source to give evidence E_j given object H_i. Since the absence of a measurement, i.e. no response from the target, is also in important information, the set of all possible evidences or declarations has to be extend with the 'no response' evidence. Each evidence E_j results in a likelihood vector (LV) consisting of the $P(E_j|H_i)$.

The set of different object types observable by a source is called source discrimination object class (SDOC). This is for consideration of source specific discrimination of different objects, e.g. a Q&A sensor like IFF can give information whether a response is correct or not but cannot give information whether the object is friend, foe or neutral.

The likelihood vectors of equal source types undergo a pre- or post-combining process dependent whether they are combined before or after the application of the SPM. After the combining process, which is the component wise multiplication of the LV to be combined, the likelihood vectors are called combined LV (CLV).

Before combining this source type specific CLV with CLVs from other source types to the joint LV, the SDOC representation of the LV has to be mapped to the members B_i of the flexible output object class (FOOC) by using a mapping matrix $MM_{ki}=P(H_k|B_i)$:

$$P_{FOOC}(E_j \mid B_i) = \sum_k P_{SDOC}(E_j \mid H_k) P(H_k \mid B_i).$$

Care has to be taken when the CLVs are combined to the JLV since the CLVs might be contrary and represent conflicting likelihoods. This might result in to a JLV of length zero. The IDCP therefore suggests to add to components which have a value of zero a small ε. This is also the point where some conflict detection has to be made. A good proposal is here to use the distance $d=\sum|x_i-y_i|$, which is

comparable to the measure of conflict in the Dempster-Shafer theory. When d is greater than a certain threshold the operator should be informed to resolve the conflict.

Now the Bayes theorem is used to get $P(B_i|E_j)$, i.e. the probability that the target is of class B_i given the evidence E_j. In this step the a priori information $P(B_i)$, the so called force mixture ratio (FMR) is used. It defines the relative number of targets belonging to class B_i expected in the surveillance volume.

For the final decision process the FOOC-likelihood vector is mapped e.g. to the STANAG 1241 standard identities via a loss table which defines the risk when making a wrong decision. This loss table depends on the actual alert state and is influenced by doctrines.

For the application of the Bayes' Theorem the following has to be known

- the force mixture ratio, which is mostly based on intelligence information
- source probability matrix for each contributing information source, when contributes its information on source level and not on likelihood level

Figure 2: Principle workflow in the IDCP from source to decision

A major drawback of this approach is that often uniform prior distribution probabilities e.g. in the FMR or SPMs are used to represent complete ignorance but this competes with the case when known information suggests a uniform distribution. This may be a problem especially a time of crises or war when every piece of information has to be assumed as affected by miss-information.

The here described approach of the IDCP is applicable both for decisions about the identity (friend, hostile,…) as they are described in the STANAG 1241 and for the classification of objects (platform, platform type, …) as described in the STANAG 4420. For the application of the IDCP for classification purposes the FOOC has to be adapted to make the correct conversion. But since the possible classifications can be described in a hierarchical tree structure like proposed in the STANAG 4420 (see Figure 3) another approach can be made which utilizes this structure in an optimal way.

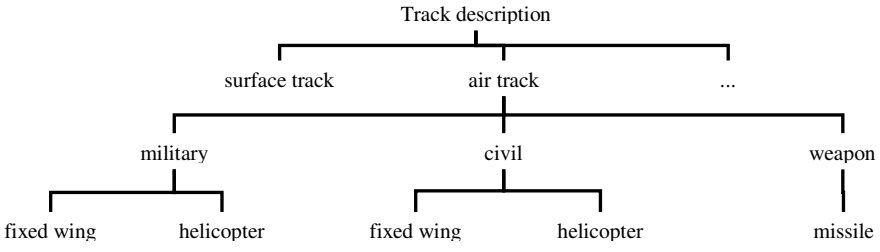

Figure 3: Proposed classification hierarchy by STANAG 4420

One approach we want to present is to use an evidence to update the specific hypothesis in the tree structure and to distribute the impact of this new information throughout the whole tree [2][3]. At the beginning of the processing the structure is initialized with the prior knowledge. For the calculations the

following reformulation of Bayes' Theorem is used

$$O(H \mid E) = \lambda_H O(H)$$

$$\lambda_H = \frac{P(E \mid H)}{P(E \mid \overline{H})} : \text{likelihood ratio}$$

$$O(H) = \frac{P(H)}{P(\overline{H})} : \text{prior odds}$$

$$O(H \mid E) = \frac{P(H \mid E)}{P(\overline{H} \mid E)} : \text{posterior odds}$$

With this nomenclature the update of a hypothesis can be written in the following recursive way:

$P(H \mid E_t, E_{t-1}, \ldots, E_1, E_0) = \alpha_H \lambda_H P(H \mid E_{t-1}, \ldots, E_1, E_0)$, where α_H is a normalization factor.

Now the children of the updated hypothesis are modified by the factor $\alpha_H \lambda_H$, the parent hypothesis F are updated via

$$P(F \mid E_t) = \alpha_H [P(F \mid E_{t-1}, \ldots, E_1, E_0) - P(H \mid E_{t-1}, \ldots, E_1, E_0)] + P(H \mid E_t, E_{t-1}, \ldots, E_1, E_0)$$

All other hypothesis are corrected by multiplying with factor α_H. This leads to a consistent structure where all hypothesis are updated correctly.

Some automatism can be introduced when the kinematics association, i.e. the association of the measurements of the source itself, is done in parallel with the identity fusion, i.e. this processes are integrated in one task. Then the conflict measure from the IDCP and also from the classification can be used to penalize the association itself and to get a combined association weight. This makes it possible to resolve dense target situation and to get in joint tracking and classification process.

3.2 Dempster-Shafer theory

3.2.1 Theory

As a generalisation of the Bayesian interference the Dempster-Shafer theory deals with proposition instead of hypothesis. This means that not only the set of hypothesis $\Theta=\{H_1, H_2, \ldots, H_n\}$ is tested but the power set $\wp(\Theta)=2^{\Theta}$. Obviously this includes also the proposition $\bigcup_i H_i$ which is identical to a general level of uncertainty. The number of possible propositions is $2^{|\Theta|}$ leading to a much higher numerical complexity compared to the Bayesian approach with n different hypothesises.

Probability masses $m(A) \in \Theta$ are assigned by an observer. The sum of all probability masses sums to one because of the exhaustive characteristics of Θ. The function m is also called a basic belief assignment (bba). With this definition one can define the so called belief ($Bel(A)$) into a proposition A which is the lower boundary for the probability that proposition A is true:

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

The upper boundary that proposition A is true is called the plausibility and is defined as

$$Pl(A) = \sum_{A \cap B \neq \varnothing} m(B).$$

An obvious advantage of the Dempster-Shafer theory is the very intuitive introduction of this measures plus the includence of the unknown. If the probability masses are defined that $m(A) \neq 0$, $A \in \Theta$ and zero for all possible disjunctions of $H_1, H_2, \ldots, H_n$ this approach is identical to the Bayesian interference approach and the following holds:

$$Bel(A) = Pl(A)$$
$$Bel(A \cup B) = Bel(A) + Bel(B)$$

The combination of evidences from different information sources is done by Dempster's rule of combination, the orthogonal sum

$$m(C) = [m_1 \oplus m_2](C) = (1-K)^{-1} \sum_{A \cap B = C} m_1(A) m_2(B)$$
$$K = \sum_{A \cap B = \varnothing} m_1(A) m_2(B)$$

where K defines the degree of conflict between the two evidences to be combined. In case of total contradiction of the two evidences K will become one and the orthogonal sum is not defined.

Especially when combining contradicting evidences ($K \approx 1$) gives problems in the interpretation of the result. A common example is the diagnosis of two doctors [4]. The set of hypothesis is $\Theta = \{m, t, c\}$ (m=meningitis, t=tumour, c=concussion). The mass probabilities given by the doctors are:

$$m_1(\mathrm{m}) = 0.99, m_1(\mathrm{t}) = 0.01 \text{ and } m_2(\mathrm{t}) = 0.01, m_2(\mathrm{c}) = 0.99.$$

Using Dempster's rule of combination results in $m(t)=1.0$, i.e. the patient suffers with certainty from tumour, which is intuitively in total contradiction to the small probability masses assigned to the diagnosis tumour by the doctors. This paradoxical situation led to a discussion about the application of Dempster's rule of combination and finally to a large set of alternative rules of combination [5] were developed.

Two general options exist to circumvent this paradoxical situation:

- to accept that the set of hypothesis Θ is not exhaustive, i.e. $m(\varnothing) > 0$. In the example this would mean that a mass probability of 0.9999 is given to the hypothesis that the patient suffers from any other illness than that in the hypothesis set
- not accepting that the set of hypothesis Θ is not exhaustive. Then e.g. the combinatorial rule of Yager [7] states, that a value of 0.9999 is assigned to the set of general uncertainty and only 0.0001 to $m(t)$, but $m(\varnothing)=0$.

3.2.2 Identity fusion via Dempster-Shafer

Identity fusion via Dempster-Shafer can be realized similar to the fusion described in the IDCP. But since we are now able to work on the power sets of object classes, a reformulation of the FOOC is possible to express the ignorance. This means that one is able give a certain probability mass to the hypothesis $H_1 \cup H_3$ when it is not possible to discriminate between H_i and H_j with an evidence. In the Bayes context H_i and H_j would be equal weighted to express the ignorance but then this might also be interpreted as a information that someone knows that they have to be equal weighted. An example is that some feature can

discriminate between military (m) and civilian (c) aircrafts but not whether they are friendly (f) or hostile (h). Then the a certain probability mass can be assign to $m(\mathrm{mf} \cup \mathrm{mh})$ and $m(\mathrm{cf} \cup \mathrm{ch})$ instead of assigning it to m(mf)=m(mh) and m(cf)=m(ch). The following example is given by Leung and Wu [8]:

Given are the hypothesis H_1=friendly commercial aircraft, H_2=friendly military aircraft, H_3=hostile commercial aircraft (or unexpected commercial airline), H_4=hostile military aircraft, and H_5=false targets. For the feature of manoeuvre detection the following probability matrices are used:

	H_1	H_2	H_3	H_4	H_5
manoeuvre	0.1	0.9	0.1	0.9	0.9
no manoeuvre	0.9	0.1	0.9	0.1	0.1

	$H_1 \cup H_3$	$H_2 \cup H_4 \cup H_5$	Θ
manoeuvre	0	0.9	0.1
no manoeuvre	0.9	0.0	0.1

Table 1: Example of the probability matrices for the feature "manoeuvre" or "no manoeuvre". Left table: for the Bayesian case, right table: for the Dempster-Shafer approach (Leung and Wu [8])

The hypothesis H_i are the elements of the FOOC defined in the IDCP. For each feature/sensor type such a matrix exists and give a basis belief assignment m_i. This basic belief assignment are now combined to get a fused basic belief assignment m.

$$m(A) = (1-K)^{-1} \sum_{A_i \cap B_j \cap C_k, \cap \ldots = A} m_1(A_i) m_2(B_j) m_3(C_k) \ldots$$

$$K = \sum_{A_i \cap B_j \cap C_k, \cap \ldots = \varnothing} m_1(A_i) m_2(B_j) m_3(C_k) \ldots$$

Using this bba the appropriate belief and plausibility can be deduced and used for the final decision process by mapping it e.g. to the STANAG 1241 standard identities via a loss table.

By the first view it seems that no a priori data are need in this approach. But that this is not true becomes obvious when going more into details. How is the probability mass calculated? Based on the example above $m(\Theta)$ is calculated as follows:

$$m(\Theta) = p(\text{manoeuvre} \mid E) \cdot 0.1 + p(\text{no manoeuvre} \mid E) \cdot 0.1$$

$$p(\text{manoeuvre} \mid E) = \frac{p(E \mid \text{manoeuvre}) \text{p(manoeuvre)}}{p(E)}, \; p(\text{no manoeuvre} \mid E) = 1 - p(\text{manoeuvre} \mid E)$$

This means that the a priori data is hidden in the determination of each feature probability and has to be determined for each feature separately or alternatively they are assumed to be unknown what means that a probability mass of one is assigned to the general level of uncertainty.

There exists a lot of critics about the computational load of the Dempster-Shafer approach, but very efficient implementations exist e.g. by using the Fast Möbius Transforms [9] but the load is still higher than that of the IDCP approach.

The target classification regarding the STANAG 4420 is also possible with the Dempster-Shafer approach. Since the frame of discernment Θ has normally a lot of different hypothesis when used in the target classification, since many platform types, platforms and platform specific types exist the computational load would be enormous. But the classification structure proposed in the STANAG 4420 is of hierarchical nature (see Figure 3) the proposal of Shafer and Logan can be used to work on a partition of Θ. They suggested that belief functions being combined should be carried by a partition $\mathcal{P}$ of Θ, which has fewer elements than Θ. The precondition for this is, that for every node $A \in \Theta$ only the elements for $A, \overline{A}$ and Θ

have non-zero masses. The processing in case of a new evidence received is four steps which will be described in the following. For a full description of the algorithm we refer to [2].

step 0: combine received evidence with existing belief value associated to this node and propagate this through the hierarchical true by using the following step 1 to 3

step 1: the belief functions attached to the terminal nodes are combined to find the degree of belief for and against their parents; the same is done also to the parents' parents throughout the whole tree.

step 2: obtain for each child A of Θ the value $(Bel_{\Theta}^{\downarrow})_{\ell_{\Theta}}$, where ℓ_{Θ} is the so called sib, which consists on the children of Θ and $Bel_A^{\downarrow}$ is the orthogonal sum of Bel_B of all nodes *B* strictly under *A*.

step 3: re-evaluate the belief of all nodes to take into account the influence of other nodes.

Finally for each node the belief and plausibility is calculated by $Bel(A) = Bel_{\Theta}^{\downarrow}(A)$ and $Pl(A) = 1 - Bel_{\Theta}^{\downarrow}(\overline{A})$.

The numerical complexity of this algorithm increases linear with number of nodes in the tree.

3.3 Dezert-Smarandache theory

A further enhancement of the so far described theories is the Dezert-Smarandache theory [6] realized in the following way [6]:

"…We extend here this notion and define now the hyper-powerset D^{Θ} as the set of all composite possibilities build from Θ with $\bigcup$ and $\bigcap$ operators such that $\forall A \in D^{\Theta}, B \in D^{\Theta}, (A \cup B) \in D^{\Theta}$ and $(A \cap B) \in D^{\Theta}$."

The cardinality of D^{Θ} is majored by $2^{2^{|\Theta|}}$ which rules the computational complexity. Belief and plausibility are defined equivalent to the Dempster-Shafer theory. The rule of combination is similar to that defined by Dempster but doesn't need any normalization factor since D^{Θ} is closed under $\bigcup$ and $\bigcap$ operators:

$$m(C) = [m_1 \oplus m_2](C) = \sum_{\substack{A,B \in D^{\Theta} \\ A \cap B = C}} m_1(A) m_2(B)$$

When we revisit the example given in the section about the Dempster-Shafer theory we now get following result

$$m(\mathrm{t}) = 0.0001, m(\mathrm{m} \cap \mathrm{c}) = 0.9801, m(\mathrm{m} \cap \mathrm{t}) = m(\mathrm{t} \cap \mathrm{c}) = 0.0099$$

which can be interpreted in that way that the patient suffers surly not from tumour but from both meningitis and concussion. This is in line with the input probabilities when the assumption holds that the set of hypothesis is exhaustive.

Beside the fact that this theory disposes the paradoxes known from Dempster-Shafer it also has the great advantage that it is possible to combine evidences from different sources which may have slightly different interpretations of the measured information, i.e. the frame of discernment. A consequence of this is, that the Dezert-Smarandache theory can be used in a modified IDCP framework like the Dempster-Shafer theory. But the usage in a hierarchical tree structure makes no sense since the hierarchy tree

structure is a only another representation of a disjunction: a parent node is the disjunction of its direct children. The conjunction which is the extension of the Dezert-Smarandache theory to the Dempster-Shafer theory finds support in the tree structure.

4.0 INTEROPERABILITY

When looking on interoperability aspects in the identity fusion process one has to look on the infrastructures and data exchange protocols available today and feasible solution in the near future.

Today the standard data exchange on force level is done be the standardized NATO tactical data links. These are mainly LINK11, LINK11B and LINK16 and in the near future LINK22. The bandwidths of these protocols are very limited, for LINK16 one can achieve up to 1Mbps when using the enhanced throughput concept. When looking at this limited available bandwidth it is obvious that a compromise has to be made which data are via the LINK for force identity fusion. In general three different levels can be identified, these the signal level, the standard identity level and the final ID level. This different levels are discussed in the following with respect to the IDCP framework.

4.1 Signal level data exchange

When the identity fusion of the ID authority should be based on the raw signals of all available sources in the force, all low level information of available sources has to be send via LINK every time a new signal is measured. In addition to this the source probability matrices for all available sources have to be available. Otherwise one has to assume a uniform prior distribution probability in order to take respect on the ignorance. Especially when a new entity enters the force this entity has to provide their knowledge about their own sensor capability and reliability. Also all entities in a force have to speak the same language with respect to the interpretations of the signals.

This sort data exchange needs an enormous amount of bandwidth and up to now no message standard exists for this kind of information. A realization of this type of information exchange is known as cooperative engagement capability (CEC) which will be used by the US navy based on a proprietary data link.

4.2 Standard Identity categories

On this communication level the likelihood vectors with respect to a common standard identity categorisation are exchanged via link. The interpretation of this likelihood vectors is independent of the capabilities of the used data sources. The needed bandwidth is much smaller than on the signal level when using the data for the determination of the allegiance since this is a manageable number of attributes. When using this data for classification purposes the needed bandwidth is much higher.

Since all involved entities must use the same interpretation of the transmitted likelihood vectors the implementation of the of the identity fusion process should be the same on all entities.

Again no message format for this kind of data exchange is defined up to now for the known tactical data links.

4.3 Final ID exchange

This sort of ID exchange is what is defined in the actual STANAGS for the different data links in addition to the STANAG 1241. But here the ID authority has only the final identity without having information about the reliability or the basis this information is determined from. Also the chance of optimizing the

identity fusion process by using different and possibly complementary sources is not available.

5.0 CONCLUSSION

In this paper we have described several algorithms to fused evidences from different sources for identification and classification of targets. Each of the algorithms has its shortcomings.

The Bayes approach needs a good knowledge of the a priori data to give good results. In the case the prior is not known the ignorance is often modelled by a uniform prior distribution probability. When this is done the algorithm cannot identify whether this is done because of ignorance of the prior or whether the prior is known to be uniform distributed. On the other side this algorithm can be implemented very efficiently in the frame of the IDCP or the hierarchical structure propose by the STANAG 4420 for the target classification.

In the Dempster-Shafer approach ignorance can be modelled by assigning some probability mass to the disjunction of elements of the from of discernment. But the effect of the normalization constant used in Dempster's rule of combination may give paradoxical results. Also the numerical complexity is very high. The values of belief and plausibility can be interpreted very intuitively by a decision maker.

Which approach is the better one is not easy to decide but there is evidence, that the approach based on the Dempster-Shafer theory is more robust against disturbances due to imprecise prior knowledge [8].

Further investigation has to made into better integration of kinematical and identity fusion especially with focus on improved automatic conflict resolution methods which are a big issue, since conflicts can arise from wrong associations in the kinematical fusion part or, when the kinematical fusion is correct, from the identity fusion part. The problem is that the conflict resolution must work without knowing the ground truth as well of the association as of the identification and classification.

With respect to the data exchange to improve the identity fusion process further investigations have to be made to find a standard to exchange low level or raw sensor information via a data link. Also the available bandwidth for this communication has to be increased. In this two points we see the most challenging problems for an improvement for a distributed identity fusion process since the mathematics of the identity fusion can be based on is well defined and more or less well understood.

[1] D. Hall, S. McMullen: Mathematical Techniques in Multisensor Data Fusion, 2nd edition, Artech House, 2004

[2] des Groseilliers, L. Bossé, E. Jean, R. Jean: Fusion of hierarchical identity declarations for naval command and control, 1996

[3] J. Pearl: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgen Kaufmann Publishers, San Mateo, CA, 552 p.

[4] L.A. Zadeh: On the Validity of Dempster's rule of Combination of Evidence, Memo M79/24, Univ. of California, Berkeley, 1979)

[5] K. Sentz, S. Ferson: Combination Rules in Dempster-Shafer Theory, 6th World Multi-conference on Systemics, Cybernetics and Informatics, July, 2002

[6] J. Dezert: An introduction to the theory of plausible and paradoxical reasoning, Proceedings of NM &A 02, International Conference on Numerical methods and Applications, Springer Verlag Ed.,

Borovetz, Bulgaria, August 20-24, 2002

[7] R. Yager: On the Dempster-Shafer Framework and New Combination Rules."Information Sciences 41: 93-137, 1987

[8] H. Leung, J. Wu: Bayesian and Dempster-Shafer Target Identification for Radar Surveillance, IEEE Transactions on Aerospace and electronic systems VOL. 36, 2, 2000

[9] R. Kennes, Ph. Smets: Computational aspects of the Möbius transform, Uncertainty in Artificial Intelligence 6, P. P. Bonissone, M. Henrion, L. N. Kanal, and J. F. Lemmer, Eds. 1991, pp. 401–416, North Holland, Amsterdam.