Adaptive Team Formation for Shared Situation Awareness

Eelke van Foeken
TNO Defense and Security and UvA
Oude Waalsdorperweg 63
2597 AK The Hague
Email: eelke.vanfoeken@ortec.nl

Leon Kester
TNO Defense and Security and UvA
Oude Waalsdorperweg 63
2597 AK The Hague
Email: leon.kester@tno.nl

***Abstract*—To reach their goals in the most effective way a team of entities needs to have a high situation awareness. In addition, the awareness needs to be shared among the entities and preferably identical in order to coordinate their actions. Unfortunately, optimal identical shared awareness is often not possible due to limited communication capabilities, particularly when the system includes complex sensors that generate vast quantities of data. Here we consider the case of multiple ships that are threatened by hostile objects. We present a method in which each team member maximizes its performance by evaluating what information to share and which members of the team to share it with given the current information needs and communication constraints of the team.**

I. Introduction

The evaluation algorithm presented in this article is an extension on *Request and Constraint Based Evaluation*—RCBE [1] and is called *Adaptive Team Formation*—ATF. Both algorithms have the goal to improve the awareness that is shared and equal between a number of entities—*Identical Shared Awareness* (ISA). Where RCBE run-time determines whether each new collected information is more valuable than costly to share with the other entities in a team, ATF extends RCBE by dynamically determining which entities should be part of the team to create ISA. ATF calculates the contribution of an entity based on the quality of features such as the sensors, the communication capabilities and the utility of its data for the *Identical Shared Awareness* (ISA). Shortly, RCBE determines whether to share and ATF determines whom to share with.

What are other techniques that do simultaneous value and cost evaluation, what can we learn from them and which aspects do we need that are not present in these methods?

In [2], a sensor management approach is introduced that, in case of limited resources of processing or communication, performs a priority assignment per *object*. This assignment is done by evaluating the risk that each object brings to the completion of the current mission objective. Thereafter, a sensing function is assigned to an object, like a tracking function to do on a certain object in the detection range. Following a sensor is selected to perform the task. Then, in case multiple tasks are assigned to a sensor, the sensor performs a resource allocation based on the priority of the tasks.

In order to cope with limited processing and communication resources in networks of cameras that perform automatic localization and tracking of people, [3] apply dynamic sensor selection based on user-defined objectives. They use Partially Observable Markov Decision Processes (POMDPs) to decide which set of cameras is active in the next time-frame. POMDPs assign rewards to actions and its resulting states. The reward is based on the number of cameras used, which defines the costs of resource use, and the value in satisfying either the coverage objective or the uncertainty objective or both. Where the coverage objective is satisfied when the persons to be tracked are observed by at least one camera and the reward in terms of uncertainty increases with lesser uncertainty.

In achieving an optimal ISA of an object in case of limited communication, work has been offered that forms coalitions of sensors that share information about the object, based on the quality that the sensors bring and the communication limitations. [4] present such an approach, by forming a coalition that shares all measurements and occasionally updating non-members with tracklets (an aggregation of multiple sensor successive measurements), and thereby using less bandwidth. Through this method they achieve ISA—they call it the *identical tactical picture*—between members and a delayed ISA between non-members. This delay is caused by the time it takes to form tracklets. They suggest three methods to determine the coalition members. All three methods achieve dynamically changing teams. At the base of these methods stands a certain information theoretic measure to find the reduction of uncertainty by measurements to the track estimate. A non-member could then for example join if the reduction exceeds a certain threshold. The first method tracks the values of this reduction over time. When two entities are competing for membership and have the same reduction at some point in time, the one with an increasing contribution should be preferred over one with a decreasing contribution. The second method aims to achieve small coalitions with optimal—that is with many different viewing angles—positioning with respect to the object. The third method aims for coalitions with members with varying qualities.

These three methods present useful approaches to coalition forming. Compared to ATF there are a few differences and a few overlaps. One main difference is that they either share all measurements or share tracklets, while in case of coalition we also use RCBE as an evaluation method for single measurements and in case of no coalition we do not share anything. Sharing tracklets or some state once in a while might be a valuable addition to our ATF. Another difference is that their method is focussed on tracking alone, while ATF is more generic. What is promising of their first method is that they compare the contribution of entities over time and compare the track-records with each other. ATF also locally keeps track of the contribution of an entity and uses this to make decisions about excluding entities from the team. They use an information-theoretical measure to find the reduction of uncertainty. We claimed in [1] that such a measure is too indirect and implicit.

The method by [2] inspired some of the features of ATF. One of them being to evaluate the contribution of sensors to single objects instead of a group of objects, since one object may endanger the mission in a totally different way than the other. One can imagine that different features of information are important when dealing with a missile threatening an entity compared with a Naval mine. Where [2] selects a single sensor to do a task, we advocate the advantages of multiple sensors observing objects. Moreover, ATF determines at run-time the expected reward that each sensing entity can bring to the state estimation of certain features of the object. This way teams are dynamically organized per object.

Even though in [3] they focus on one specific application their method is highly generic. Generic in that it can be applied to other applications with different sensors, different tasks and different user-objectives. But also generic in how to define the costs of resource use. We believe we can install ATF in this generic method. However, our method is novel, firstly, because it particularly focusses on constructing ISA, and secondly, because it uses a new mechanism of determining utility and reward of information.

II. Adaptive Team Formation

The leading example of this article is a maritime scenario where the environment is observed by multiple ships and these ships are threatened by incoming objects e.g. missiles. Such situations can only be averted when the ships have a sufficiently accurate estimate of the location, speed, direction and acceleration of the missiles. In these typically time-critical situations, the ISAs of the objects need to have a certain quality. First, the contribution is dependent on whether a ship has the object in its *effector range* and in its *detection range*—for example, if a ship both has no effector range as well as no detection range on the object, it should not waste bandwidth to receive information about it. Second, there are constraints in the communication channel and the communication system that influence the delay of transmission and therefore delay the synchronization of the ISA. If one ship has a really bad connection with the other ships, it might be that the best sensory information loses its value and could better be used locally. Third, the quality of the sensory information itself is a measure of how it impacts the ISA and the effectiveness on the object. When the sensor itself has a low resolution, a low detection rate or a small range, it may not at all be able to deliver detections worthy of communication. Even if its connections are good, it has good sensory information, and it can affect the object, the information might not influence the currently important features, hence the the information brings low utility to the ISA.

To deal with either improving or degrading contributions of entities to the information needs for the object in the environment, Adaptive Team Formation—ATF—is needed, which periodically adapts the teams of agents for every object. The method evaluates the contribution, and determines which agents should remain, be excluded or be included in the team.

III. Method/Approach

Deciding to include or exclude agents involves the problem of recognizing to what degree it influences the expected effect on an object. Clearly, a disadvantage of exclusion is that an excluded agent does not have an updated ISA, which disables it to *optimally* coordinate actions with the other entities. It also stops delivering its collected information to the other agents, which can result in worse qualities of the ISA, such as a less accurate estimated location, or delayed detection of objects because they enter the smaller detection range of the smaller team later. Advantages of exclusion may be that the delay of communication among the members of the new team decreases so that the ISA is synchronized sooner and reaction time is shortened and that the costs of communication decrease. Moreover, if the excluded agent has the object within visual range, it logically has *no delay* in updating its local object awareness—LA, and can therefore affect the object *directly*.

Linking new agents in the team has the advantage that there is a greater area-coverage, robustness and that the ISA is fed with additional object-information so that, for example, accuracy increases. The additional entities on which the new agents reside also may increase the effectiveness on the objects, because first, they can work together instead of alone and, second, the joint effectors have exponentially more impact. The downside, however, is that the costs of communication will increase and the delay of communication, synchronization and reaction will lengthen.

In short, the decision of including or excluding is based on the difference between the *expected reward* of including an entity and the *expected reward* of excluding the entity. The expected reward being a trade-off between the utility of the awareness and the communication costs and delay.

The agents have knowledge about:

- the communication situation with the other entities,
- their sensors,
- their local tracking algorithm,
- the current information-requests,
- the tracks they have within their detection range.

The ATF utilizes all this knowledge to calculate the difference in expected reward that an agent brings to the ISA.

A. Set Up

The system under consideration consists of entities $\mathcal{J} = j_1, j_2, j_3, \ldots$. These entities have got radars that produce detections, $\mathcal{Z} = [z_1, z_2, z_3, \ldots]$ out of an uncertain environment $X \equiv \mathbb{R}$. The radars surveil the environment for visible objects $\mathcal{O} = [o_1, o_2, o_3, \ldots]$ Each entity j has multiple information abstraction levels as defined by [5], where the evaluation methods are acting on the *object assessment* level. Each entity can harbor multiple agents at this level; one for each object.

Each agent, a_{oj}, is cooperating in a team with other agents, $\mathcal{S}_o$, to observe and act upon a singular perceived object o—$o \in \mathcal{O}$. This cooperation consists of maintaining an *identical shared awareness* (ISA), $\hat{x}_o$, of the object. All ISAs of all visible objects are $\hat{X} = [\hat{x}_1, \hat{x}_2, \hat{x}_3, \ldots]$ Each agent uses a *core* functionality that is a domain-dependent algorithm ϕ fusing level 0 detections, $\mathcal{Z}_o = [z_1, z_2, z_3, \ldots]$, that associate to object o, into level 1 updated state estimates, $\hat{x}_o \mid \mathcal{Z}_o = \phi(\hat{x}_o, \mathcal{Z}_o)$.

Important to note is that an agent, a_{oj}, comes into existence when the entity j starts participating in the team of object o. In practice an agent is born when the entity finds a new track or when the entity contributes positively to a track maintained by other agents than itself and is allowed to partake in the team.

The agent contains two adaptive functions: Request and Constraint Based Evaluation (RCBE) and Adaptive Team Formation (ATF). Both functions determine some kind of expected reward. ATF is a function defined as $\Delta\hat{\mathcal{R}}\left(\mathcal{S}_o, \mathcal{S}_o^+\right)$ for determining whether $\mathcal{S}_o$ should be enlarged with a new entity to become $\mathcal{S}_o^+$ or as $\Delta\hat{\mathcal{R}}\left(\mathcal{S}_o, \mathcal{S}_o^-\right)$ for determining whether $\mathcal{S}_o$ should be smaller by excluding an entity to become $\mathcal{S}_o^-$. These decisions are based on the information $\mathcal{Z}_o$ that the new entity associates with the shared state estimate $\hat{x}_o$. RCBE is a function that determines whether single observations, such as z_o, should be shared with team $\mathcal{S}_o$ based on its contribution to $\hat{x}_o$ and therefore used to update $\hat{x}_o$: $\hat{x}_o \mid z_o = \phi(\hat{x}_o, z_o)$.

Adaptive Team Formation—ATF—is a method used for every object that an entity can observe, affect or both. It acts at run-time, and is triggered by different events. Its task is to determine for every perceived object whether its team remains constant, continues with a new member or continues with a member less. Depending on the type of evaluation that ATF needs to do it needs a subset of the following information:

- $\hat{x}_o$ — the current identical state estimate—ISA—of the perceived object o,
- $\hat{y}_o$ — the current local state estimate of the perceived object o,
- $\hat{x}_{ro}$ — the current reference state estimate of the perceived object o,
- $\mathcal{S}_o$ — the current object-team,
- $\hat{p}$ — the current Expected Delay Distributions (EDDs), which gives an estimation of the delay that communication will have,
- $\hat{\mathcal{C}}$ — the current Expected Cost of Communication (ECC), which gives an estimation of the amount of resources (*e.g.* bandwidth),
- U — the utility functions that represent the information-requests,
- pos — the locations of each entity, which is a reasonable expectation for allied forces.

The team, $\mathcal{S}_o$, shares an identical state estimate—that is an ISA, $\hat{x}_o$, of the perceived object. Such a team can be very transitory, whose lifetime is limited by several constraints, such as object-visibility duration or the duration of interest for the object. Next to keeping an ISA, each team-member with object-visibility maintains a local state estimate, $\hat{y}_o$, and a reference state estimate, $\hat{x}_{ro}$ of the object.

Fig. 1 shows a diagram that illustrates the process of ATF. ATF can be evoked when:

1) a new local track, $\hat{y}_o$, or an already shared track, $\hat{x}_o$, can be observed or can be acted upon by another entity.
2) a certain time-interval has passed after the last time the team has done an ATF.

If a new local track, $\hat{y}_o$, is found of object o by entity j, the object will be represented by a new agent, a_{oj}, and a new team, $\mathcal{S}_o$. When this object comes in either the effector, detection or both ranges of another entity, agent a_{oj} starts the ATF process for inclusion. If awareness of an object is already being shared, $\hat{x}_o$, and the object is in or enters either or both ranges of another agent, the leader of the team, $a_{oj} \in \mathcal{S}_o$, will start the ATF process for inclusion.

The moment that a single agent starts sharing its local object-awareness with another agent, a time-variable, m_o, is set to indicate when ATF has last been run. Another variable, n, determines the interval, $[0, n]$, after which ATF has to be run again. This interval is put such that it is significantly longer than the interval between two successive detections. m is reset when ATF has been run. At time $m_o + n$, each agent within $\mathcal{S}_o$ will perform ATF for exclusion.

IV. Utility

It is important what characteristics an *information-request* should minimally have. First of all, from [6] we learned that an *information-request* posed by a higher information abstraction level needs to have a clear relation with the type of information that the lower information abstraction level produces. The request should quantitatively indicate what and/or when information gathered by the lower abstraction level is relevant. We believe that the *information-request* should be a *utility function*:

> The utility function quantifies the utility as a function of the important features of information.

Formally: level 2 agents give *information requests*. An information-request takes the form of a utility function $U_{\mathcal{S}}(\Gamma) \in [0, 1]$. It is defined as $U : \mathbb{R}+ \rightarrow \mathbb{R}+$, and assumed to behave monotonically—either decreasing or increasing. It expresses the utility of any number of currently important evaluation features of information, where the features are expressed as a set of uncorrelated evaluation parameters $\Gamma = [\gamma_1, \gamma_2, \ldots]$. Examples of features that may be important are *region*, *accuracy* or *timeliness*.

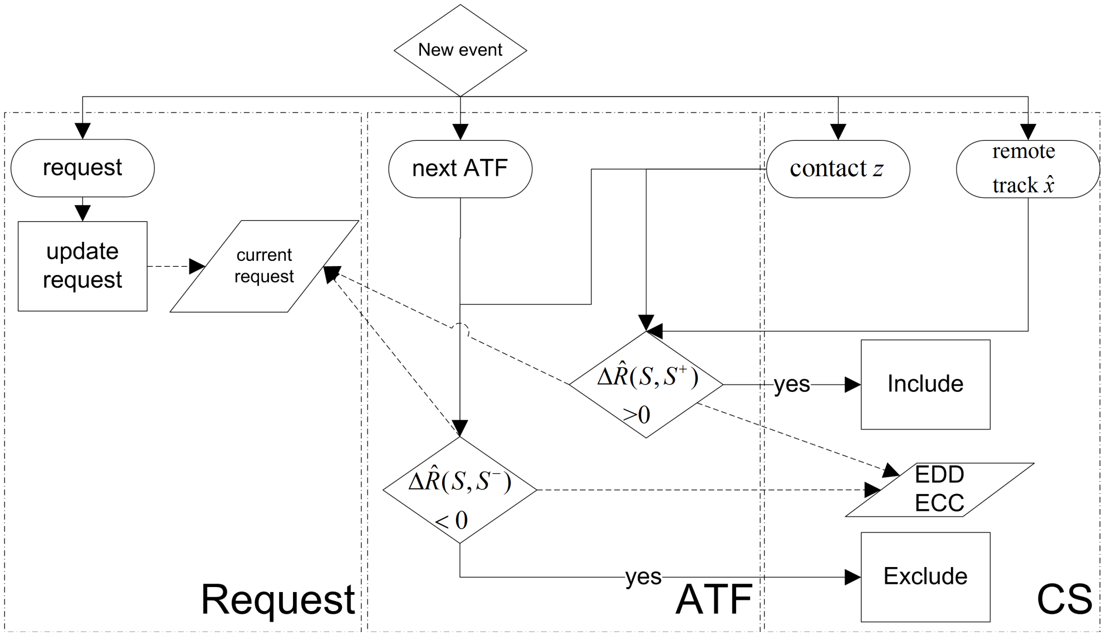

Fig. 1. Flowchart showing the steps for ATF. Event 1: request update. Event 2: n ms have passed; it is time for a new ATF. Event 3: entity receives new contact z. Event 4: entity receives a remote track $\hat{x}$. If the entity is up for inclusion after a new contact has been received it uses current EDD and ECC and request to determine if inclusion is more rewarding than exclusion. If it is the moment for possible exclusion it calculates if exclusion would be more rewarding than staying. If so, it is excluded.

Ultimately, the utility functions are derived from the wished actions in the environment. It seems logical that action effectiveness increases when entities cooperate instead of them acting individually. We would like to incorporate this by assuming that with the same state, $\hat{x}$, the action effectiveness on objects increases *exponentially* with more entities cooperating. This can be modeled by assuming that the sum of local utilities of a state is smaller than the team utility of the same state. Formally:

$$\forall\mathcal{F}|\cup\mathcal{F}=\mathcal{S} \text{ and } \cap\mathcal{F}=\emptyset:\left(\sum_{\mathcal{S}\in\mathcal{F}}\left[U_{\mathcal{S}}\left(\Gamma\right)\right]\leq U_{\mathcal{S}}\left(\Gamma\right)\right) \quad (1)$$

For all families of sets, $\mathcal{F}$, where the union of all subsets of that family is the same as $\mathcal{S}$ and the intersection of all members is an empty set, the summed utilities of all subsets in the family of features Γ, is smaller than the single utility of set $\mathcal{S}$. In other words, each team has a utility function that is a function of the individual agent utilities: $U_{\mathcal{S}}=f\left(U_1,...,U_{|\mathcal{S}|}\right)$. Such a function can be modeled in any way but in this article f is defined as:

$$f\left(U_1,...,U_{|\mathcal{S}|}\right)=|\mathcal{S}|^{\alpha}\sum_{a=1}^{|\mathcal{S}|}U_a. \quad (2)$$

where $(U_{\mathcal{S}}\mid(\alpha>0))>\sum_{a\in\mathcal{S}}U_a$. α is a variable exponent, where the higher its value the more the effectiveness of a larger team will increase. If α is zero there is no increased effect of coordination.

To calculate the utility of the state estimate $\vec{x}$ with respect to the reference state estimate $\hat{x}_r$ the utility function can be integrated over all state instances of *both* pdfs. Since it involves two pdfs a double integral needs to be performed. This results in the following 'difference measure', $\mathcal{U}$, between continuous pdf's $\hat{p}(\vec{x})$ and $\hat{p}_r(\vec{x}_r)$:

$$\mathcal{U}\left(\Gamma\left(\hat{p}_r,\hat{p}\right)\right)=\iint U\left(\Gamma\left(\vec{x}_r,\vec{x}\right)\right)\hat{p}\left(\vec{x}\right)\hat{p}_r\left(\vec{x}_r\right)\,\mathrm{d}\vec{x}\mathrm{d}\vec{x}_r. \quad (3)$$

The details of this function can be found in [1] This is the *integrated utility function*, and represents a direct and interpretable utility function. All the possible values of the state are directly measured for their utility and probability and lead to a scalar representing the average utility. It is defined as $\mathcal{U}\in[0,1]:\mathbb{R}+\to\mathbb{R}+$, and assumed to behave monotonically—either decreasing or increasing.

A. Expected Delay of Communication

For estimating the reward we also need to estimate the delay of communication. For this we use the model from [7]. Therein, the expected cumulative delay distribution (ECDD) is a function of the N transmission attempts to agents $\mathcal{S}$, $\hat{\mathcal{P}}\left(t,\mathcal{Z},\mathcal{S}\right)$:

$$\hat{\mathcal{P}}\left(t,\mathcal{Z},\mathcal{S}\right)=\sum_{n}^{N}p\left(\mathcal{S},n\right)\prod_{a=1}^{|\mathcal{S}|}\mathcal{G}\left(t-\Delta t_{\mathrm{det}}\left(a,n\right)\right), \quad (4)$$

where $\mathcal{G}$ is a Gamma Cumulative Distributed Function

To derive an EDD the differential equation from (4) is taken:

$$\hat{p}\left(t,\mathcal{Z},\mathcal{S}\right)=\frac{\mathrm{d}(\hat{\mathcal{P}}\left(t,\mathcal{Z},\mathcal{S}\right)}{\mathrm{d}t}. \quad (5)$$

The communication service delivers this probability distribution as EDD by request from the evaluation methods.

B. Expected Cost of Communication

To estimate the reward the costs of communication need to be estimated. The ECC used in this article can be found in [7] as well and is defined as the amount of resources used over time, with respect to the total amount of resources available. Here, time is related to the required number of transmission frames, F, times the number of expected transmissions required, N. The expected cost of the nth transmission attempt is then defined as

$$\hat{c}(n, \mathcal{Z}, \mathcal{S}) = Fn \frac{t^F B \sum_{a=1}^{|\mathcal{S}|} P(a)}{t^F_{\text{tot}} P_{\text{tot}} B_{\text{tot}}}, \tag{6}$$

and the expected cost of communication (ECC) is the sum of all attempts:

$$\text{ECC} = \hat{\mathcal{C}}(t, \mathcal{Z}, \mathcal{S}) = \sum_{n=1}^{N} \hat{c}(n, \mathcal{Z}, \mathcal{S}), \tag{7}$$

It is defined as $\hat{\mathcal{C}} : \mathbb{R}+ \rightarrow \mathbb{R}+$, has a lower bound of zero and increases with the number of frames, F, and number of attempts, n. Here, t^F_{tot}, P_{tot} and B_{tot} represent the total frame length, power and bandwidth. A communication network has a certain total bandwidth B_{tot} available. Assuming that an agent has a certain bandwidth B allocated for transmitting a message to agents $\mathcal{S}$, a larger B claims more of the total bandwidth B_{tot}, hence is more costly. The same applies to the frame length fraction T and the transmission power $P(a)$. Depending on the communication system properties and settings, certain resources, such as bandwidth, can be shared among multiple agents.

V. Inclusion

The first time an object o is detected by one of the entities $j \in \mathcal{J}$ the object-assessment agent on the entity will start a *new track*, $\hat{y}_o$. Later, another entity could be able to observe or affect the object. At such a moment the question is whether object-participation affects the object more than when they would work individually. If so, team $\mathcal{S}_o$ is born. Also, when the object has been tracked by multiple agents and enters the detection or effector range of a new entity, the same choice has to be made: should that new entity be included in the current team $\mathcal{S}_o$ to become $\mathcal{S}_o^+$?

There are three possible situations where an entity j is up for inclusion, which all require different calculations:

- The object only enters the effector range
- The object only enters the detection range
- The object enters both effector and detection range

For simplicity we only consider the last case; we assume that the effector and detection range overlap. In this case entity j has visibility of the object. The problem is that the entity has no history of detections associating with the track and is therefore unable to determine what its contribution would have been in the previous time-window. It can only determine what the impact of future detections will be. First it needs the track and therefore receives this from an agent in team $a_{oj} \in \mathcal{S}_o$. At reception the entity temporarily becomes part of the team so that it receives detections from the team and can send its own detections to the team. This to make the evaluation as accurate as possible. Remember that RCBE determines whether single detections are worth sharing.

We define a parameter g that determines the number of detections the entity will be using for evaluation. The higher g, the more accurate the evaluation. Surely, a single detection is little information to base inclusion on, and therefore this parameter is used. The formula is:

$$\begin{aligned} \Delta\hat{\mathcal{R}}\left(\mathcal{S}_o, \mathcal{S}_o^+\right) \quad &= \quad \sum_{z=z_1}^{z_g} \hat{\mathcal{R}}_{\text{inc}}\left(\hat{x}_o, z, \mathcal{S}_o^+\right) \\ &\quad + \hat{\mathcal{R}}_{\text{exc}}\left(\hat{x}_o, \emptyset, \mathcal{S}_o\right) - \hat{\mathcal{R}}_{\text{exc}}\left(\hat{y}_o, z_o, a_{oj}\right). \end{aligned}$$

The first term is:

$$\begin{aligned} \hat{\mathcal{R}}_{\text{inc}}\left(\hat{x}_o, z, \mathcal{S}_o^+\right) \quad &= \quad \int_{t=0}^{t_{\max}} \hat{p}\left(t, z, \mathcal{S}_o^+\right) \\ &\quad \left[\mathcal{U}_{\mathcal{S}_o^+}\left(\epsilon\left(\hat{x}_o\left(t\right)\right)\right) - \hat{\mathcal{C}}\left(t, z, \mathcal{S}_o^+\right)\right] \mathrm{d}t \\ &\quad - P_f \hat{\mathcal{C}}\left(t_{\max}, z, \mathcal{S}_o^+\right). \end{aligned}$$

This calculation involves the utility that the larger team is expected to have and the costs that are expected to be made to communicate the z. In detail it involves integrating over the delay t until $t_{\max}$. $t_{\max}$ is reached at the moment $\mathcal{U}_{\mathcal{S}_o^+}\left(\epsilon\left(\hat{x}_o\left(t\right)\right)\right) - \hat{\mathcal{C}}\left(t, \hat{x}_o, \mathcal{S}_o^+\right) <= 0$. In other words, when the expected costs of communication are higher than the expected utility of sharing. For each t the expected probability of successful communication, $\hat{p}$, is multiplied with the utility minus the costs. The probability of failed communication at $t_{\max}$ is expressed by P_f. The costs that would be made times this probability, $P_f \hat{\mathcal{C}}\left(t_{\max}, \hat{x}_o, \mathcal{S}_o^+\right)$, need to be subtracted to finalize the calculation.

The expected reward for exclusion consists of two terms, the first determining the expected reward of the state without the detection for the current team S_j:

$$\hat{\mathcal{R}}_{\text{exc}}\left(\hat{x}_o, \emptyset, \mathcal{S}_o\right) = \mathcal{U}_{\mathcal{S}_o}\left(\epsilon\left(\hat{x}_o\left(t=0\right)\right)\right). \tag{8}$$

And the second the utility of the local state estimate $\hat{y}_o$ of the object:

$$\hat{\mathcal{R}}_{\text{exc}}\left(\hat{y}_o, z_o, a_{oj}\right) = \mathcal{U}_j\left(\epsilon\left(\hat{y}_o\left(t\right)\right)\right). \tag{9}$$

VI. Exclusion

If there has not been any ATF moment for n seconds it is time for the *next ATF*, and all the agents, a_{oj}, comprising all the current teams, $\mathcal{S}_o$, calculate whether they should remain or be excluded from their team. For a single agent a_{oj} the function is therefore defined as the expected change of reward when the team would have without that agent instead:

$$\varphi(\mathcal{S}_o, \hat{x}_o, \mathcal{Z}_o) \equiv \Delta\hat{\mathcal{R}}\left(\mathcal{S}_o, \mathcal{S}_o^-\right) >= 0 \rightarrow \mathcal{S}_o^- : \mathcal{S}_o \tag{10}$$

If $\Delta\hat{\mathcal{R}}$ is positive, it would probably have been better to work without a and results in the new team being $\mathcal{S}_o^-$, if negative the agent remains $\mathcal{S}_o$. The difference in reward is calculated as follows:

$$\Delta\hat{\mathcal{R}}(\mathcal{S}_o, \mathcal{S}_o^-) = \hat{\mathcal{R}}_{\text{exc}}\left(\hat{x}_o, \emptyset, \mathcal{S}_o^-\right) + \hat{\mathcal{R}}_{\text{exc}}\left(\hat{y}_o, \mathcal{Z}_o, a_{oj}\right) - \hat{\mathcal{R}}_{\text{inc}}\left(\hat{x}, \mathcal{Z}_j, \mathcal{S}_j\right) \quad (11)$$

As Fig. 1 shows, the *current request*, all previously local contacts that associated with $\hat{x}$ between $t-m$ and t that is set $\mathcal{Z}_o$, and the EDDs and ECCs are needed for this calculation. The first and second term connote the expected reward when agent a_{oj} would have been excluded from the team—that is the reward of state $\hat{x}_o$ for $\mathcal{S}^-$ without $\mathcal{Z}_o$ plus the reward that $\mathcal{Z}_o$ would have brought to agent a_{oj} locally, hence without sharing. The third term of (11) signifies the expected reward that the previous detections $\mathcal{Z}_o$, would have brought to the shared track state $\hat{x}_o$.

The first and second term are broken down as follows:

$$\hat{\mathcal{R}}_{\text{exc}}\left(\hat{x}_o, \emptyset, \mathcal{S}_o^-\right) = \sum_{z \in \mathcal{Z}_o} \left[\mathcal{U}_{\mathcal{S}_o^-}\left(\epsilon\left(\hat{x}_o\left(t=0\right)\right)\right)\right] \quad (12)$$

$$\hat{\mathcal{R}}_{\text{exc}}\left(\hat{y}_o, \mathcal{Z}_o, a_{oj}\right) = \sum_{z \in \mathcal{Z}_o} \left[\mathcal{U}_a\left(\epsilon\left(\hat{y}_o\left(t=0\right)\right)\right)\right] \quad (13)$$

In both cases there is no communication, so there is no delay and costs. The first term is the summed utilities of the current state at the times when detections $\mathcal{Z}_o$ were measured, but without the detections actually included. The second term is the summed local utility of the local track state at those moments.

The third term is dissected as follows:

$$\begin{aligned}\hat{\mathcal{R}}_{\text{inc}}\left(\hat{x}_o, \mathcal{Z}_o, \mathcal{S}_o\right) \quad = \quad & \sum_{z \in \mathcal{Z}_o} \int_{t=0}^{t_{\max}} \hat{p}\left(t, z, \mathcal{S}_o\right) \\ & \left[\mathcal{U}_{\mathcal{S}_o}\left(\epsilon\left(\hat{x}_o\left(t\right)\right)\right) - \hat{\mathcal{C}}\left(t, z, \mathcal{S}_o\right)\right] \, \mathrm{d}t \\ & -q\hat{\mathcal{C}}\left(t_{\max}, z, \mathcal{S}_o\right)\end{aligned}$$

For every $z \in \mathcal{Z}_o$ the expected reward is calculated and summed. The expected reward of a detection consists of two statements. The first is an integral over the delay interval $[0, t_{\max}]$. For every t the difference between zs' utility for state $\hat{x}_o$ and cost of communication is multiplied with the expected success probability of sharing z, $\hat{p}$. However, there is a probability, P_f, that communication fails to a moment, $t_{\max}$, where the cost becomes higher than the value so that it would not be rewarding anymore to share z. This also signifies the probability that costs have been made during $[0, t_{\max}]$. Therefore the last term is subtracted from the integral.

In conclusion, the loss or the gain of reward by excluding agent a_{oj} is estimated by comparing the effectiveness on the object of $\mathcal{S}_o$ cooperating with that of $\mathcal{S}_o^-$ and agent a_{oj} working separately. The combined utility of $\mathcal{S}_o$ is, as a rule, higher for the same state, but the delay may deteriorate the state so severely of the costs or communication may be so high, that the estimated reward is equal or lower. In that event the agent is excluded from the team $\mathcal{S}_o$.

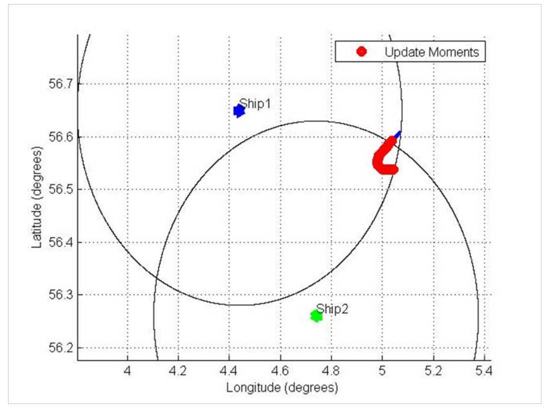

Fig. 2. A snapshot of the first scenario in matlab. Two ships with the black circled (overlapping) detection ranges observe a single flying object. The routes of object is shown in blue where the red rots express the update moments by the tracker.

When an agent did not have any visibility or effector range on the object during the previous time-window the two terms will be equal: $\hat{\mathcal{R}}_{\text{exc}} == \hat{\mathcal{R}}_{inc}$ and therefore it will be excluded. First of all, the agent did not collect any detections associated with the track, which makes the costs zero and obviously causes no delay. Therefore $\hat{\mathcal{R}}_{\text{inc}} = \mathcal{U}_{\mathcal{S}_o}\left(\epsilon\left(\hat{x}_o\left(t=0\right)\right)\right)$ Then, since the state lies out of its effector range, the shared awareness does not have an added utility for the agent. This means that $\mathcal{U}_{\mathcal{S}_o}\left(\epsilon\left(\hat{x}_o\left(t=0\right)\right)\right) == \mathcal{U}_{\mathcal{S}_o^-}\left(\epsilon\left(\hat{x}_o\left(t=0\right)\right)\right)$. Locally the agent does not have an awareness of the object and makes the local utility logically $0, \mathcal{U}_a\left(\epsilon\left(\hat{y}_o\left(t=0\right)\right)\right) == 0$. In conclusion: $\hat{\mathcal{R}}_{\text{exc}}\left(\hat{x}_o, \emptyset, \mathcal{S}_o^-\right) = \mathcal{U}_{\mathcal{S}_o^-}\left(\epsilon\left(\hat{x}_o\left(t=0\right)\right)\right) = \hat{\mathcal{R}}_{\text{inc}} = \mathcal{U}_{\mathcal{S}_o}\left(\epsilon\left(\hat{x}_o\left(t=0\right)\right)\right)$

VII. Experimental Results

Fig. 2 shows the first scenario. Both scenario's consisted of a DSS. Scenario 1 comprised two ships j_1, j_2 and scenario 2 three ships j_1, j_2, j_3. The ships used radars to keep objects in their visual range under surveillance and the level 1 agents together maintained an ISA of the tracks of the objects in the area. Level 0 agents served as a local source of noisy detections. A single unidentified objects, which was in truth hostile fighters, flew first into the detection and effector range of ship 1 and later into these ranges of ship 2 as well. In the second scenario the object flow first in the detection and effector range of ship 1 then in those of ship 3 and finally in that of ship 2. Ship 3 is placed in between ship 1 and 2 and the object moves in a straight line.

This scenarios represent a DSS in need of timely and accurate information about the objects in the environment in a communication constrained network. The ships needed to surveil the environment for hostile fighters. However, transmitting contacts in high-frequency strained limited communication resources, and resulted in high communication costs and significant delays. Delay decreased the accuracy of the ISA, hence the effectiveness of surveillance task. In conclusion, these scenarios are suited for testing to what extent ATF can maintain a high quality ISA.

The resulting tracks are measured in how much offline reward the tracks have brought in relation to the currently relevant features of information. We measure the offline reward—utility minus costs—at the moments that measurements have been made.

The ships get the utility function of Fig .3 as an information-request:

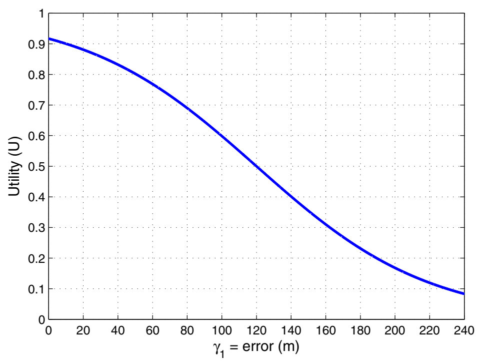

Fig. 3. The accuracy request of agents a_2 and a_2 take the form of this utility function.

Furthermore, there are some important parameters to be set:

g The number of detections used to determine whether a new entity should be included is 2.

n The interval after which there is a new exclusion moment is 10 s. Each entities generates contacts every second. This means 10 contacts per interval are used.

A. results scenario 1

We have run the same scenario with three different settings:

1) $\alpha = 0$ and using only RCBE.
2) $\alpha = 0$ and using ATF as well as RCBE.
3) $\alpha = 0.5$ and using ATF as well as RCBE.

Recall from Eq. (2) that α is the exponent for cooperation. What we would like to show is that ATF rightly chooses not to cooperate in case $\alpha = 0$ because the sum of local rewards will by definition always be higher than the shared reward. We also would like to show that ATF rightly chooses to cooperate in case $\alpha = 0.5$.

In Fig. 4 plotted the offline rewards of:

1) The local tracks of the object when using ATF with $\alpha = 0$. Over time there is no cooperation between the ships since ATF chooses to do so, hence the ships do not have a shared track. The offline reward is simply the same as the offline utility of the track error in relation to the ground truth.
2) The sum of the offline rewards of the local tracks.
3) of the shared track when only using RCBE with $\alpha = 0$,
4) the shared track when using ATF when $\alpha = 0.5$.

What can be seen is that the reward of the track when only using RCBE is equal or lower than the sum of the

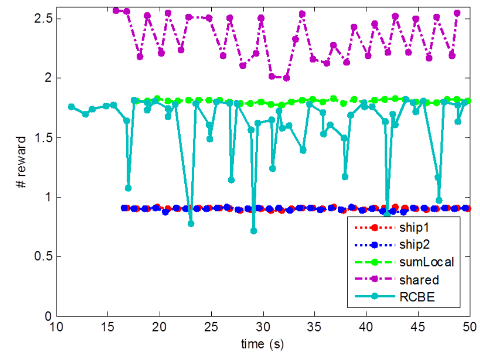

Fig. 4. The offline rewards of the local estimates in blue and red, of the sum of the two local rewards in green, of when using only RCBE in light blue and in purple when ATF is used with an $\alpha = 0.5$

local rewards. This is logical since $\alpha = 0$ and because less detections are used for the shared track—some detections are not rewarding enough for communication. Moreover, the delay and costs of communication also influence the the offline reward negatively.

Taking these things into account and the fact that communication costs are made makes it clear that in this situation ATF has merits.

In the next experiment α is raised to 0.5 to increase the effect of cooperation. This to show that ATF does choose to cooperate when the sum of local expected rewards is always lower than the shared reward. The purple line shows this offline reward of the shared track. It can be seen that ATF has done well to choose cooperation since the shared reward always tops the sum of the local offline rewards. Apparently this is enough to compensate for the costs of communication.

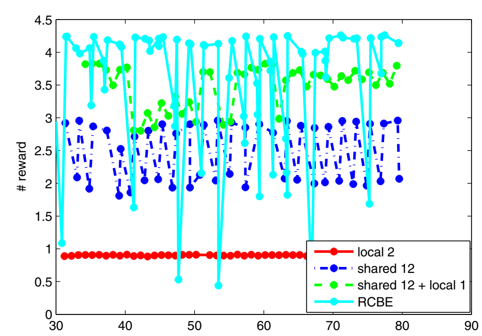

Fig. 5. Case $\alpha = 0.4$: The offline reward of the local estimate of ship2 in red, of the shared estimate between ship 1 and 3 in blue, of when using only RCBE in light blue and green the sum of the first two

B. results scenario 2

We have run the same scenario with three different settings:

1) $\alpha = 0.4$ and using only RCBE.
2) $\alpha = 0.4$ and using ATF as well as RCBE.

3) $\alpha = 0.7$ and using ATF as well as RCBE.

The results of the first two settings are displayed in Fig. 5. There the offline rewards are plotted of:

1) the local track by ship 2 of the object with—red,
2) the shared track between ship 1 and 3 of the object with—blue,
3) the sum of the previous—green,
4) the shared track when only using RCBE—light blue,

In case $\alpha = 0.4$, ship 1 and ship 3 are cooperating, but ships 2 is not. This can be explained as follows: The object moved into the effector and detection range of ship 1 first and then moved into the ranges of ship 3. At that point in time ship 3 received the track from ship 1 and evaluated itself to be rewarding enough to cooperate with ship 1. Hence they became a team: $\mathcal{S}_o = j_1, j_3$. When the object moved into the ranges of ship 2, it received the track from ship 1, since that is the leader of team $\mathcal{S}_o$. It evaluated itself to be negatively influencing the accuracy of the track, hence not including itself in the team. Observing the offline rewards, this did make sense. The light blue line represents the offline reward when only using RCBE and not ATF. This means that all ships are team members by default. The green line represents the offline reward when ATF was used, which is the sum of the blue and red line. The green line does not top the blue line all the time, but the valleys are so large and below the lowest parts of the green line that the average offline reward is lower. These valleys indicate that the communication costs where really high. We can conclude ATF was useful, since it determined cooperation between three ships would not be rewarding due to the high communication costs.

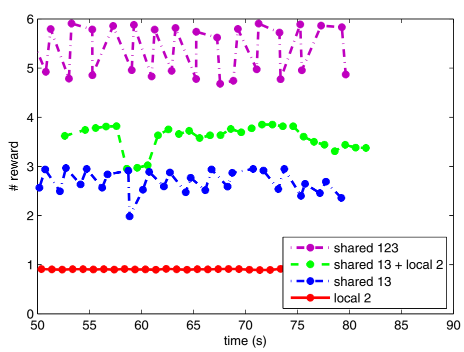

Fig. 6. Case $\alpha = 0.7$: The offline reward of the local estimate of ship2 in red, of the shared estimate between ship 1 and 3 in blue of the sum of the previous two and the shared estimate between ship 1, 2 and three in purple

If $\alpha = 0.7$ ship 2 did decide to cooperate, see Fig. 6. The offline reward of cooperation topped the green line, representing the sum of the offline reward of cooperation between 1 and 3 and the local offline reward of ship 2.

In conclusion, ATF is able to adapt the team formation to the degree that cooperation is appreciated (this is indicated by α) and to the communication costs.

VIII. Conclusion

In this article we introduced the method of adaptive team formation. This method determines run-time which entities should be part of the team to create shared awareness. Results are presented for creating shared awareness with two and with three ships. From these results it is shown that ATF, depending on the benefit of coordination of actions, automatically selects the proper entities in the team that should cooperate in tracking objects. In future experiments we would like to test the impact on team formation of varying communication capabilities. Moreover, we would like to research the scalability of ATF by having larger groups of entities and multiple objects to track. Finally, we want to examine the adaptivity to run-time chances such as the communication capabilities or information-requests.

References

[1] E. van Foeken, L. Kester, J. Bergmans, M. Kwakkernaat, and F. Groen, "Maintaining an identical shared awareness in communication constrained sensor systems," *Localized Algorithms for Information Fusion in Resource-Constrained Networks*, 2011, to be published.

[2] F. Bolderheij, F. Absil, and P. van Genderen, "A risk-based object-oriented approach to sensor management," in *the 8th International Conference of Information Fusion*, Seoul, Republic of Korea, 2005.

[3] M. T. J. Spaan and P. U. Lima, "A decision-theoretic approach to dynamic sensor selection in camera networks," in *Int. Conf. on Automated Planning and Scheduling*, 2009, pp. 279–304.

[4] M. Howard and R. E. D. Payton, "Coalitions for distributed sensor fusion," in *Proceedings of the Fifth International Conference on Information Fusion 2002*, vol. 1, 2002, pp. 636–642.

[5] A. Steinberg and C. Bowman, "Rethinking the jdl data fusion levels," in *Proceedings of the Military Sensing Symposia (MSS) National Symposium on Sensor and Data Fusion (NSSDF)*, vol. 2, Colombia, 2004.

[6] S. Eswaran, D. Shur, and S. Samtani, "A metric and framework for measuring information utility in mission-oriented networks," *Pervasive and Mobile Computing*, 2011.

[7] E. van Foeken and M. Kwakkernaat, "Low-complexity wireless communication modeling for information flow control in sensor networks," in *14th International Conference on Information Fusion 2011*, Chicago, Illinois, USA, 2011.