14th ICCRTS

"C^2 and Agility"

Shannon Entropy as a Metric of Situational Awareness in C^2 Structures

Topic 5 – Experimentation and Analysis

André Luiz Pimentel Uruguay, Major, MSc*[1]
Celso Massaki Hirata, PhD [2]

[1] Instituto de Estudos Avançados
Rodovia dos Tamoios, Km 5,5
São José dos Campos, SP 12228-001 - Brazil
auruguay@ieav.cta.br

[2] Instituto Tecnológico de Aeronáutica
Pça Marechal Eduardo Gomes, 50
São José dos Campos, SP 12228- - Brazil
hirata@ita.br

*Point of Contact: André Luiz Pimentel Uruguay - Major

+55-12-3947-5345 / +55-12-9121-5695 (mobile)

auruguay@ieav.cta.br

Shannon Entropy as a Metric of Situational Awareness in C^2 Structures

Abstract

The present work is intended to investigate the possibility of using the concept of entropy to create metrics of situational awareness in command and control systems. The organizations are represented in its structural and functional forms, using MOISE+ model. Two metrics of situational awareness are proposed, based on the Shannon's entropy. The first metric measures the degree of uncertainty on the position of entities dispersed in a geographic area based on a belief diffusion model that takes into account the sensors in the proximity. For the second metric, the difference between the maximum value of entropy and the entropy after the use of the information obtained from the sensors is used to measure the information gain provided by a C^2 system. The difference between the information gains of two air forces in conflict is conceived as a simple metric of information superiority. A simulation study is presented, and its results show a fair correlation between the information superiority and the performance of an air force. This correlation increases if the global value of the information superiority, accumulated during the conflict period, is considered.

Keywords: metrics, entropy, uncertainty, information gain, multiagent simulation.

1. Introduction

In many application domains a set of entities organize themselves in order to operate in a distributed way in a dynamic, uncertain and partially observable environment, aiming at a common objective. These organizations can be considered as information processing units [Galbraith, 1974], where sensors obtain perceptions from the environment, transmit informations to decision making entities which, by their turn, communicate to entities capable of acting over the environment.

The present article aims to present evidence of the relationship between the information value, as perceived by the decision makers, and the performance of the organization in the accomplishment of its global goal. These value can be determined using the concept of entropy as a metric of the degree of uncertainty over random events and, thus, of the degree of lack of information. Initially the concept of entropy is defined, created with the advent of the Information Theory [Shannon, 1948]. Two examples of application in the military domain are presented. Two entropy-based metrics are defined. The first measures the information gain of an organization, expresses by the amount of information provided by its sensors, and calculated by the consequent reduction of entropy of the beliefs of a decision maker over the geographic position of the other agents in the environment. The second metric, denominated information superiority, measures the difference between the information gains obtained by two opposing organizations. To exemplify the application, an agent-based simulation study is presented, involving two C^2 systems (and its respective organizations), in a situation of air warfare.

2. The Concept of Entropy

In the 20th century, the appearance of the telephone and the expansion of the telegraph imposed some challenges on the possible constraints to the information transmission rates conveyed over noisy channels. It was believed, at that time, that the probability of error at the reception of a message could only be reduced by decreasing the transmission rate, i.e., an error-free message could only be successfully received if the transmission ceased completely.

Claude Shannon [1948] disagreed and proved that probabilities of error arbitrarily low could be reached, provided that the communication channel had non-null capacity (calculated with the noise included). Also, Shannon stated that the random processes like speech or music had an irreducible complexity below that no signal compression would be possible.

It's these reference to random processes that makes Shannon theory so attractive to Statistics in the attempt to measure uncertainty and its complement, i.e., what is known, based on the acceptance of a probability distribution as a complete and sufficient representation of the uncertainty present in mutually exclusive and collectively exhaustive events.

Let's consider a set of n possible events of a random variable $X = \{x_1, x_2, \ldots, x_n\}$ with probabilities of occurrence, respectively, $p_1, p_2, \ldots, p_n$. These probabilities are known but this is all that is known about which event will occur. Let $H(p_1, p_2, \ldots, p_n)$ be a measure of how much "choice" is involved in the selection of the event, or of how much uncertainty is present. Shannon proved that, to H be continuous, monotonically increasing with n and with maximum value to equally probable events, such quantity, denominated *Entropy*, should be of the following form:

$$H = -K \sum_{i=1}^{n} p_i \log p_i ,\ p_i > 0 \ , (1)$$

which unit, if using base 2 in the logarithm and $K=1$, is the *bit*.

As defined, the entropy has limited values. The lower limit is reached when there is maximum certainty on which event will occur, represented as $p_j=1$ to the certain event, and $p_i=0$, $i \neq j$, for the other events. By definition it's considered 0 log 0 = 0, which yields

$$H(X) = -1 \log 1 - (n-1) 0 \log 0 \ , (2)$$

The upper limit occurs in the situation of maximum uncertainty, i.e., with all events equally probable. In this case, for n possible events, $p_i = 1/n$, and

$$H = -\sum_{i=1}^{n} \left(\frac{1}{n}\right) \log\left(\frac{1}{n}\right) = \log n \ . \ (3)$$

Thus, $H(X)$ ranges from 0 to $\log n$.

3. Entropy and Information in the Military Domain

One essential part of Command and Control is information processing, aiming to reach a decision. The decision quality is given by quality of shared information provided by sensors, people and other equipments. Thus, the performance of a C2 system is related to the information gain provided by these entities, being relevant to determine metrics to dynamically measure that gain. Barr and Sherrill [1996] considered these gain as the difference between the *a priori* uncertainty and the uncertainty after the event.

Considering entropy an adequate measure of uncertainty, if an event I affects the state of another random variable T, the information gain resultant of knowing I is

$$\delta(T|I) = H(T) - H(T|I) \qquad (4)$$

Applying Shannon's formula for entropy,

$$\delta(T|I) = \Sigma P(T|I) \log P(T|I) - \Sigma P(T) \log P(T) \qquad (5)$$

By discretizing the geographic space, it's possible to calculate the entropy of the information about the position of entities generated from many detection assets over a determined area [Barr and Sherrill, 1996].

Let T be a target positioned in a cell C_i , from a set of cells $C_1, C_2, \ldots, C_n$, and a sensor with probability of detection $p_D = P[I(j)|T(j)]$. Let $p_j = P[T(j)]$ be the *a priori* probability of T being at the cell C_j and $I(j)$ the event of detecting T at the cell j. Supposing null probability of false alarm $p_{FA} = P[I(j)|\neg T(j)]$,

$$P[T(j)|\neg I(j)] = \frac{(1-p_D)p_j}{1-p_D p_j} \qquad (6)$$

and

$$P[T(i)|\neg I(j)] = \frac{p_i}{1-p_D p_j}, i \neq j \ . \qquad (7)$$

where $\neg I(j)$ represents the event of non-detection of the target at cell C_j.

For the case when the probability of detection p_D varies as a function of the sensor and the cell, let $D_{s,c}$ be this probability. Thus, if no sensor detects the target in a certain period of time, the update p_{t+1} of the probability distribution p_t of the target's position is

$$p_{t+1}=\frac{p_t \circ d_{t+1}}{|p_t \circ d_{t+1}|} \qquad (8)$$

where $d_{t+1}=\prod_s (1-D_{s,c})$, $\circ$ represents component-wise product between vectors, and $|.|$ represents the sum of all elements of a vector. After these update, the entropy as a function of time is determined by

$$H(T,t)=-\sum_C p_t \log p_t \qquad (9)$$

We have to consider, especially in the air power domain, the situation of detecting moving targets. In this case, at the moment after a sensor detect a target, these information begins to degrade. Several models exist to describe such degradation [Shupenus and Barr, 1998]: uniform distribution square, uniform circular and exponential cone.

In the case of uniform circular, the probability over the position of the target is uniformly distributed along a circle (or, in this case, a polygon, composed of cells, that best approximate a circle) of radius D_p, being D_p the distance that the target can potentially travel for each time step after the last detection.

A similar work was presented by Beene [1998] who, beyond discretizing a 2D space in cells, inserted in the Shannon's formula a factor related to the resolution of the sensor in question, i.e., a value of area beyond which the position of a detected target cannot be refined. The formula stated by Beene is

$$H(X)=-\sum_i p_i \ln\left(\frac{p_i}{A}\right), \qquad (10)$$

where A is the 2D resolution of the considered sensor.

One advantage of the Beene's formula is the ability of analyzing the dynamics of the position of an entity in the environment by assessing the use of sensors with different capabilities (range and resolution). Also, Beene created a model to describe how the belief about the position of a target degrades after a sensor loses detection.

4. A Metric of Information Gain

Based on the work of Shupenus, Barr and Beene, it's possible to create a process to quantify the degree of uncertainty as seen by a decision maker who needs the information about the position of many targets in a geographic space of interest, discretized in cells [Uruguay, 2006].

For this study a discretization based on a bidimensional hexagonal lattice was employed. There are only three regular polygons able to cover a plan: triangles, squares and hexagons. Hexagons have the advantages of being the most compact form, providing the best angular resolution and discretizing the space with the least average error [Sahr et al., 2003]. Consequently, the propagation of any spatial attribute along hexagonal cells has the least *anisotropy*, i.e., has the best uniformity in all directions. This is valuable characteristic for the accuracy of the calculation of belief degradation.

Situated in this lattice there is a C2 system composed, among other elements, of spatially situated sensors with defined maximum detection ranges.In general these sensors don't have total coverage over the area of interest, i.e., as commented at the introduction, the decision maker has only partial observability.

If the information about the position of a given target is not updated by the sensors, it's supposed that the corresponding beliefs of the decision maker will begin to degrade. These degradation can be expressed by spatially diffusing the probability distribution along the set of cells potentially covered by the target. These set grows with a rate proportional to the velocity of the target.

To better comprehend these diffusion process it's relevant to recall the Principle of Maximum Entropy. The principle states that, for a given random event of unknown probability, one shall adopt the distribution that maximizes the entropy of such event, with the condition that such distribution shall satisfy all known hypothesis about the event. Otherwise, any other distribution would convey incorrect bias. A simple pictorial example of the belief degradation about a target's position can be seen at fig.1.

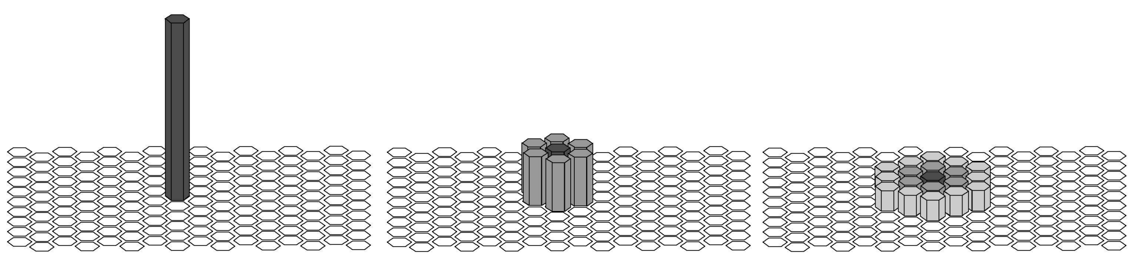

Figure 1: A pictorial example of the belief degradation of a target's position in the case of no sensors in the neighborhood.

Supposing that what is known about the target's position is certain, i.e., that the target is in only one cell, such belief is graphically expressed as a single bar at the left, representing the unitary value. As time passes, it's reasonable to assume that the target can also be in other cells. Thus, the initial belief of a certain event (and consequent unitary probability) shall be shared with other neighbor cells. If no other sensor is present at the neighborhood, the probability distribution would be such that all cells in a given space would have the probability value, I.e., the entropy would be maximum.

However, assuming that the only constraint over the target is its limited capability of movement, the set of cells where the presence of the target is possible is also limited. In the search and detection theory these set of cells is sometimes denominated *datum*. As time passes, if the target is not re-observed, new steps of degradation shall occur, increasing the size of the datum and, also, the dynamic value of entropy. It's possible to assess the degradation of the information about a target's position taking these different snapshots of probability distributions and applying Shannon's entropy as a metric of, in this case, increasing uncertainty.

Now, let's consider the presence of a sensor in the neighborhood of the target. Sensors are supposed to have an effective range of detection over a 2D space, so all cells within these range have a non-null probability of detection. Such cells are represented in fig.2 in gray color.

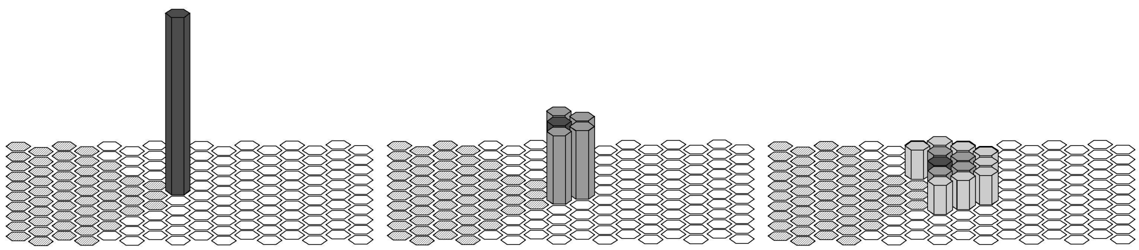

Figure 2: Example of the belief degradation of a target's position in the case of a sensor in the neighborhood. Cells covered by the sensor are in gray color.

The process of belief degradation is similar, with the remark that, in the present case, there is an additional hypothesis to the Principle of Maximum Entropy: considering the case of a missing target and a perfect sensor in the neighborhood, the *a posteriori* probability of the target being in a cell covered by the sensor is null. Thus, it doesn't make sense to diffuse the probability distribution over these cells, as shown in the last images of fig.2.

5. Information Superiority

As shown, the concept of entropy enables the calculation of the information gain about a random event from the knowledge about the occurrence of another, related, event. It's possible, also, to determine, between two organizations, which one had, at a certain moment, the best information gain, i.e., in relative terms, a simple metric of its *information superiority*.

Let's take the maximum value of entropy H_{MAX} for the situation corresponding to the total absence of sensors, I.e., null information gain. Considering O_1 and O_2 two C^2 organizations, its information gains are, respectively,

$$\delta(O_1) = H_{MAX} - H(O_1) \text{ and } \delta(O_2) = H_{MAX} - H(O_2) \qquad (11)$$

It's possible to consider the *information superiority S* between two organizations, as the difference between its respective information gains. Thus,

$$S(O_1,O_2) = \delta(O_1) - \delta(O_2) = H(O_2) - H(O_1) \qquad (12)$$

6. Example of Application

A simulation study was conceived to study the dynamics of the information gain (and, consequently, uncertainty). The simulated scenario involves an air campaign theater, composed of radar sites, strike and air defense aircrafts, communication stations and command centers. These assets were modeled as artificial cognitive agents, organized in two opposing forces, BLUE and RED.

All agents have their behaviors constrained by their respective organizational roles. Also, aircrafts and radars have limited capabilities of sensing and movement. Strike and air defense aircrafts can attack, respectively, only ground targets and other aircrafts. These actions can only occur by orders given by decision makers at the command centers. Also, the results of the engagements are determined by a simple random model where either attackers or defenders can be destroyed.

Each force aims to win by attacking ground targets in the enemy territory and intercepting enemy aircrafts invading its own airspace. As each agent is inserted in an explicit organizational model, its destruction can generate effects over other agents belonging to the same organization. For instance, depending on the organizational structure, the destruction of a communications station can cease the information flow between radar sites and command centers, causing a decrease in the information conveyed to the corresponding decision makers.

In order to verify the effects of the structure on the performance of an organization in such an uncertain environment, two different structures were modeled. The MOISE+ organizational model [Hübner,2003] was adopted to express both structure and functional specifications. The two organizational structures are depicted in fig.3.

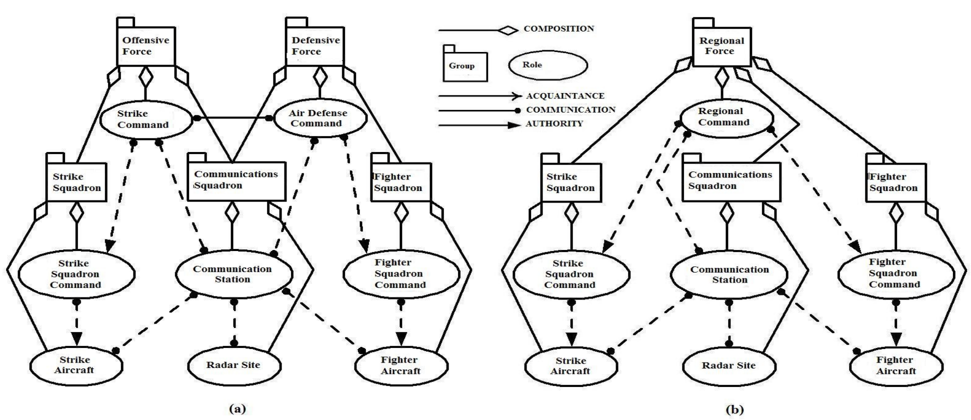

Figure 3: Structural specifications of the organizations modeled: (a) Centralized; (b) Regional.

The first structure is such that all relevant decisions are taken by a central command center. There are two specialized commands: one, dedicated to strike missions, and another, responsible for air defense.

The second structure considered for study involves a regional, geographic division of organizational functions. Instead of specialized commands, there are two regional commands, with identical responsibilities and independence of one each other.

An agent-based simulation was implemented, based on the Belief-Desire-Intention (BDI) cognitive architecture. With the purpose of facilitating the design of the goals of each agent, the IDEF0 technique [USA, 1998] was applied to express the organizational functional model, in a modified version of MOISE+ model [Uruguay and Hirata, 2006]. Fig. 4 presents the most relevant functions, in IDEF0 language. Both organizations modeled had the same functional model.

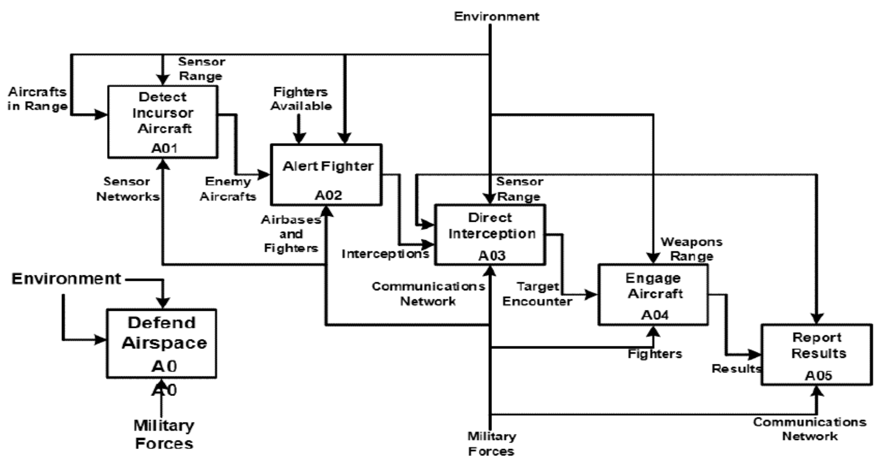

Figure 4: Example of functional specification, in IDEF0, of the modeled organizations.

A screenshot of the scenario, as shown in the developed simulation tool can be seen in fig.5.

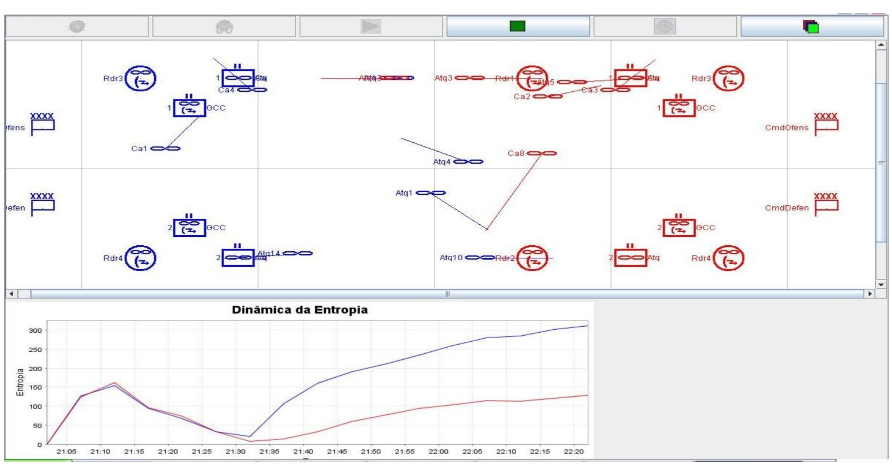

Figure 5: Screenshot of the simulated scenario.

To observe the dynamics of the entropy values calculated during simulation, the organizational structure and radar detection range of blue forces were varied.

The measured values of information gain refer to the entropy over the known position of the aircrafts (the only mobile entities in the scenario) of both forces, consistent with the beliefs of the decision makers, located at the command centers. These beliefs are based on the information generated by the radar sites and transmitted to the command centers via communication stations.

If a command center has an updated information about the position of an aircraft, its entropy is null, corresponding to a maximum information gain. But once this information is not updated, the belief of the commander begins to degrade in the way presented at section 4, and the entropy (information gain) begins to increase (decrease).

The entropy of each force is understood as the sum over the entropies of all aircrafts, friendly or enemy, as perceived by the commander. An important premise was that the communications infrastructure would perfectly reflect the information exchange requirements of each organizational structure. Also, for the sake of simplicity, no collaboration processes between the agents were adopted for the case of lack of information.

For each set of parameters 10 simulation runs were executed. Each run consisted of 6 simulated hours, and the total entropies of each force were registered at periods of 5 minutes. The total numbers of ground and air targets destroyed by each force were also registered.

7. Results and Analysis

Fig. 6 presents an instance of the dynamics of entropy, where one can note that the RED force (continuous line) kept lower values of entropy, which corresponds to greater information gain.

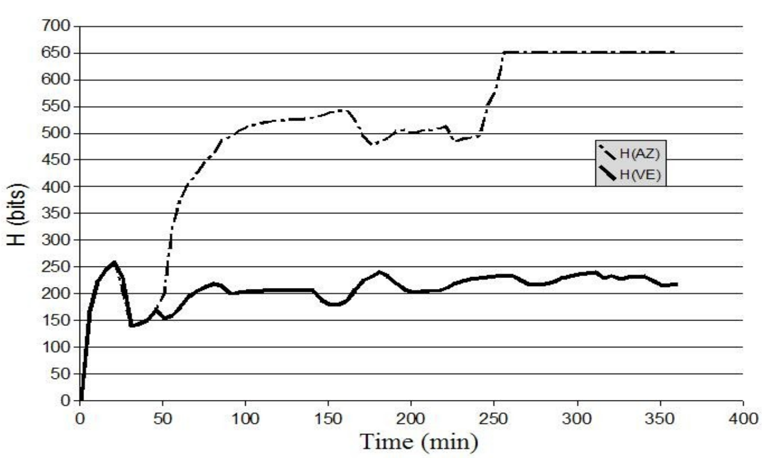

Figure 6: Example of the dynamic of entropy of the two forces, as a function of time.

Fig. 7 presents the relationship between the cumulative value of the information superiority between BLUE and RED forces and the number of RED and BLUE targets destroyed, for three different values of BLUE radar detection ranges.

It's possible to note that the points corresponding to low detection ranges (white triangles) by BLUE are related to a negative difference between RED and BLUE targets destroyed. Also, as expected, these points correspond to negative values of information superiority. Conversely, it seems, from figure 7, that high detection ranges (black squares) allow the system to reach regions of performance otherwise unreachable, especially in the centralized case.

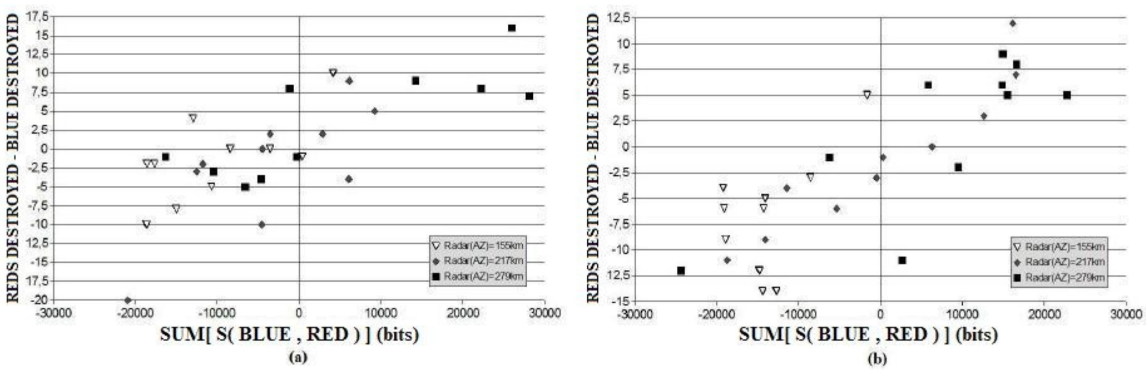

Figure 7: Relationship between cumulative information superiority of BLUE over RED and the difference of RED and BLUE targets destroyed:(a) centralized scenario; (b) regional scenario.

Table 1 presents the correlation index between the performance of BLUE and its information superiority, for the centralized scenario (centralized BLUE versus centralized RED organizational structures) and regional scenario (regional BLUE versus centralized RED). In this case “performance” was understood as the difference between the number of RED and BLUE targets destroyed.

Both final and cumulative (integrated) values of our information superiority metric were considered. As it can be seen, these results correspond to a moderate correlation between information superiority performance. Also, the correlation is a barely better for cumulative values, which considered the “history” of entropy dynamics, than for the final values of information superiority.

Table 1: Correlation Index (R^2) between the difference of RED and BLUE targets destroyed and information superiority of BLUE over RED, for the two simulated structures.

	Performance = $f(S_{final})$	Performance = $f(S_{cumulative})$
Centralized Scenario	0,50	0,55
Regional Scenario	0,65	0,70

Related to the organizational structures modeled, fig.8 shows no relevant difference observed in the values of information superiority, with a small advantage to the centralized structure. The figure includes all the points referring to the different values of radar detection ranges for BLUE force.

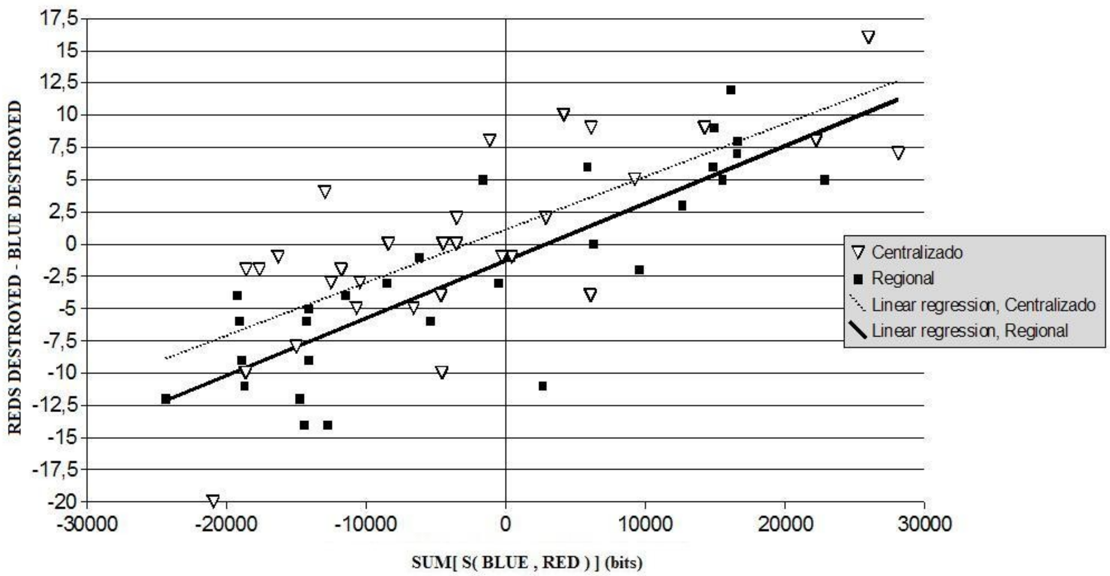

Figure 8: Global information superiority of BLUE over RED, for both organizational structures of BLUE.

By analyzing figures 7 and 8 we conclude that there is a monotonic relationship between information superiority and performance, although the correlation is, according with table 1, moderate, not sustaining any claim of linearity, as expected from a non-linear typical C^2 scenario.

8. Conclusions

The present work aimed at the possibility of employing the concept of entropy, as defined by Claude Shannon, to build metrics of situational awareness for C^2 organizations.

As an example, a simulation study was conducted, about the scenario of two opposing air forces of equal capabilities (similar aircrafts, radars and weapon systems). To express organizations the MOISE+ model was applied, which includes distinct representations both for structure and for function.

Two metrics were presented, based on entropy: the first, to determine the degree of information provided by a Command and Control system; and the second, to measure at which level a C2 system is superior to another in providing more information gain.

Results of applying the information gain metric showed that, in general, the winner side was able to keep low values of entropy.

Also, results point to a moderate correlation between the second metric, denominated information superiority, and the performance of a military C^2 organization, here defined, in a simplistic way, by the difference between targets destroyed.

One advantage of the entropy-based metrics is the possibility of dynamic application, i.e., its values can be computed as the whole system changes its state. Even in the case of the information superiority, its calculation is possible in practical terms, if one can estimate the actual information gain of the enemy C^2 system based on the current tactical picture and additional intelligence data (obtained from passive electronic warfare, remote sensing, Humint, etc.).

The ability of capturing the organizational and systemic effects is another advantage. Although in our experiments no relevant difference was found between the two organizational structures modeled, we believe that different C^2 architectures can be evaluated from the point of view of uncertainty, which is one of the factors impacting agility.

References

Barr, D. R. and Sherrill, E. T. (1996). Measuring Information Gain in Tactical Operations. Technical report, US Military Academy, West Point.

IEEE (1998). *IEEE STD 1320.1-1998 - Standard for Functional Modeling Language - Syntax and Semantics for IDEF0*. Institute of Electrical and Electronic Engineers, New Jersey.

Galbraith, J. R. (1974). Organization Design: An Information Processing View. *Interfaces*, 4(3):28–36.

Hübner, J. F. (2003). *Um Modelo de Reorganização de Sistemas Multiagentes*. PhD Thesis, Escola Politécnica da Universidade de São Paulo, São Paulo.

Sahr, K., White, D., and Kimerling, A. (2003). Geodesic Discrete Global Grid Systems. *Cartography and Geographic Information Science*, 30(2):121–134.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27:379–423, 623–656.

Shupenus, J. L. and Barr, D. R. (1998). Information Gain and Loss. Technical report, US Military Academy, West Point.

Uruguay, A. L. P. (2006). Uso da entropia como métrica de consciência situacional em estruturas de comando e controle. Master Thesis, Instituto Tecnológico de Aeronáutica, São José dos Campos.

Uruguay, A. L. P. and Hirata, C. M. (2006). Using IDEF0 to Enhance Functional Analysis in MOISE+ Organizational Modeling. In Sichman, J., Coelho, H., and Rezende, S., editors, *Proceedings of the 10th Ibero-American AI Conference (IBERAMIA-SBIA 2006).*, LNAI 4140, pages 78–87, Berlin. Springer-Verlag.