UNIVERSITÀ DI PISA

DIPARTIMENTO DI INGEGNERIA DELL'ENERGIA DEI SISTEMI
DEL TERRITORIO E DELLE COSTRUZIONI

RELAZIONE PER IL CONSEGUIMENTO DELLA
LAUREA MAGISTRALE IN INGEGNERIA GESTIONALE

Text driven technological mapping of C4ISTAR: defining a fuzzy defence related domain

SINTESI

RELATORI

Prof. Ing. Gualtiero Fantoni
Dipartimento di Ingegneria Civile e Industriale, Università di Pisa

Dott. Filippo Chiarello
Dipartimento di Ingegneria dell'Energia, dei Sistemi, del Territorio e delle Costruzioni, Università di Pisa

Prof. Jean Belin
Department of Social and Human Sciences, University of Bordeaux (GREThA UMR CNRS 5113) and Institut des hautes études de défense nationale (IHEDN) Chair in Defence Economics, France

IL CANDIDATO

Irene Spada
irenespada@hotmail.it

Sessione di Laurea Magistrale del 18/06/2020
Anno Accademico 2019/2020
Consultazione non consentita

Text driven technological mapping of C4ISTAR: defining a fuzzy defence related domain

Irene Spada

Sommario

Questo lavoro di tesi si inserisce nell'ambito del progetto europeo ASSETs+[1], che è volto a progettare e implementare dei corsi di formazione per le professioni del settore della Difesa sulla base delle competenze emergenti legate alla robotica, all'Intelligenza Artificiale, alla cybersecurity e al dominio C4ISTAR[2]. L'oggetto della presente ricerca è la mappatura del dominio C4ISTAR attraverso metodologie di analisi testuale. Nello specifico, partendo dall'analisi dei sistemi della Difesa, è stato possibile identificare i documenti relativi all'ambito C4ISTAR. Al fine di comprendere il dominio tecnologico, individuarne le parole chiave e descrivere il complesso di elementi che lo compongono, tali documenti sono stati elaborati con l'analisi delle occorrenze in dizionari di settore, con l'analisi della distanza tra stringhe (per similarità con lista di parole chiave estratte automaticamente), con la mappatura delle co-occorrenze e con l'esplorazione delle strutture sintattiche nei documenti collezionati.

Abstract

This thesis work is related to the European project ASSETs+[1], set up to identify the existing and emerging skills needs for professions in the Defence sector, in order to build up training courses for the labour market. The technological domains to address are robotics, Artificial Intelligence, cybersecurity and C4ISTAR[2]. The goal of this research is the mapping of the C4ISTAR domain through text driven methodologies. Specifically, the identification of the documents related to the C4ISTAR area is based on the analysis of the Defense systems. Then, with the purpose to understand the technological domain, identify its keywords and describe the elements that compose it, these documents are elaborated with the occurrences analysis in several dictionaries of fields related to the C4ISTAR area, with the analysis of the strings distances (looking for the similarity with a list of keywords automatically extracted), with the mapping of the co-occurrences and with the exploration of the syntactic structures within the collected documents.

[1] "Alliance for Strategic Skills addressing Emerging Technologies in Defence".

[2] "Command, Control, Communication, Computer, Information/Intelligence, Surveillance, Target Acquisition, and Reconnaissance".

Introduction

C4ISTAR: a fuzzy domain

C4ISTAR is an acronym used in Defence for Command, Control, Communication, Computer, Information/Intelligence, Surveillance, Target Acquisition, and Reconnaissance. It is a wide concept that states both for the functions related to the military missions and the systems that allow the actions and the coordination. It can be used to refer also to the infrastructure, the role of individuals or military units and the procedures used [1]. Therefore, the concept lends itself to an all-inclusive interpretation. This aspect leads to a multidisciplinary of the field itself, that encompass military doctrine and technological systems to support defence operations. Within all the procedures and the applications of C4ISTAR systems (also known as Command and Control systems or C2 systems), it is fundamental the use of an agreed terminology and definition [2].

Methodology overview

Understanding the evolution of the C2 systems facing with its complexity calls for a novel approach that can deal with this multifaceted domain. For this reason, a lexical approach both with automatic techniques and human supports has been used to provide a definition and a dictionary of keywords for this fuzzy defence related domain. The methodology includes the following steps:

1. Defence systems delineation is an exploration of the whole context of use of the C4ISTAR systems;
2. C4ISTAR field delineation consists in an investigation on the descriptions of this concept, with the purpose to provide a robust definition of this fuzzy domain;
3. C4ISTAR dictionary delineation is a lexical analysis on the keywords related to the C4ISTAR domain, aiming the identification of the main features of the field;
4. C4ISTAR mapping is a graphical analysis on the C4ISTAR dictionary using data visualization to understand the main feature of the domain.

Operative environment

The operative environment is the European transnational project Sector Skills Alliances, named ASSETs+, acronym of “Alliance for Strategic Skills addressing Emerging Technologies in Defence”.

It seeks to identify the skills needs for Defence professions, in order to build up training courses for the labour market. The technological domains to address are robotics, artificial intelligence, cybersecurity and C4ISTAR. The idea is to analyse the trends in technologies, identify the skills required to use them and understand what the requirements of the training programmes are. The collaboration between labour market, education and training stakeholders and social partners is the key to achieve this goal, so the ASSETs+ project involves 30 partners of 8 different countries (i.e. Italy, France, Denmark, Germany, Poland, Spain, Sweden and United Kingdom).

1. Defence systems delineation

1.1 Analysis of world military expenditure

Defence is the set of systems and methods used to protect the integrity of a State's territory, the security of its population and the maintenance of civil institutions. It includes Armed Forces and Diplomatic Activities. In every Nation there is a Defence Industrial Base, known as DIB, a group of firms which provide defense related equipment to the Defense Ministry [3].

Gross Domestic Product (GDP) is the monetary measure of the market value of all the final goods and services produced in a State in a specific time period. It can be used as a statistical indicator to understand the national economic health and the dynamic of development and progress. It can be broken down into the contribution of each sector of the economy, such as Defense or education services provided by the Government [4, 5].

On account of that, a State's military expenditure as a share of GDP — also known as the military burden — is one of the measures of economic commitment of a country in Defence. As registered in World Bank repositories [6], the military burden all over the World experienced a decreasing after the end of the Cold War. As noted in SIPRI Military Expenditure Database[3] [7], U.S. are the largest spender in the World, whose spending in 2018 is the 36% of global expenditure, so that this country has a dominant position in the world scene. Europe as a geographical region is the third largest spending zone, that contributes for 20% of the global value, and its top 5 of spender Member States are France, United Kingdom (before the Brexit), Germany, Italy and Spain.

1.2 Analysis of the Defence Industrial Base

Since the DIB develops, produces and provides goods and services for the Military Department of a State, in each Nation there is a solid group of core contractors and various suppliers at the lower levels of the supply chain [8]. The Government supports its DIB and it is also its main consumer. As a result, the Defence Industry is shaped on the National policy and financial decisions. With the purpose to give evidence of this point, it is possible to analyse the SIPRI TOP 100[4] [9].

Firms in SIPRI TOP 100 [9] can be classified in relation to the main operative sector, i.e. defence and/or civil, looking at their arms sales as a percentage of their total sales. During the last decades firms shift towards mix production (more dual firm than pure firms). This is the effect of the reduction of public military expenditure that each Nation experienced after the Cold War.

[3] SIPRI is an independent international institute dedicated to research into conflict, armaments, arms control and disarmament. It was established in 1966.

[4] SIPRI TOP 100 derives from the SIPRI Arms Industry Database, that contains information on arms-producing and military services and is based on data from OECD as well as open sources, such as company annual reports and articles in journals and newspapers. SIPRI TOP 100 is updated every year since 2002.

In U.S. “in 1993, Secretary of Defense Aspin and Deputy Secretary Perry concluded that some industrial consolidation was necessary and announced this to executives from major defense companies at a Pentagon dinner. … The audience was receptive and as Norman Augustine (Martin Marietta CEO at the time) put it later: «You weren’t going to survive unless you were willing to combine. So, there was not much of a choice»” [8].

Almost half of the firms in SIPRI TOP 100 [9] are American and as well the group of firms that gains the most part of the sales (General Electric, Boeing, United Technologies Corp., Lockheed Martin Corp. and Honeywell International), then there are also Trans-European firms (Airbus Group), German firms (ThyssenKrupp) and Japanese firms (Mitsubishi Electric Corp.). This outcome confirms the dominant position of the U.S. and the positions of power described in the Paragraph 1.1.

1.3 Analysis of Innovation in Defence

Innovation is a key factor for economic growth and produces positive effects for the society and the richness of a country [10]. It allows firms to maintain and improve their competitive positions in the market. In a similar way, the innovation in Defence allows a Nation to gain supremacy in the global panorama. Since it is a key public sector, innovation leads to strategic results and they require public protection. Nowadays technological development is faster[5] and more complex[6] than in the past and these effects produce changes also in the public guidance of the innovation in Defence.

For example, in U.S. the Department of Defence defines policy and guidelines to acquire and exploit new technologies and capabilities to support U.S. Army [12]. The main action of control is carried on by DARPA, the Defence Advanced Research Project Agency. Today it coordinates with its Programmes Managers the innovation ecosystem composed by the National agencies, the academic partners and the corporations whose activities are defence-related [13].

In France there is a strong cooperation between defence companies and research centres, both at National and International level, always under the aegis of the French Defence Procurement Agency, that is the Direction Générale de l’Armement (DGA), that supports innovation policy and project about innovative technologies with potential military applications with the actions of the Agence de l’Inovation de Défense (Defence Innovation Agency), and with the investment fund DefInvest [14].

At International level there are several agencies in NATO committed in innovation, research and development, such as NATO Science and Technology Organization (STO) and Science for Peace and

[5] The number of scientific and technical publications, as well as the Government Expenditure in Research and Development (GERD) increase all over the World.

[6] Technological complexity is linked to the need of different competencies to exploit the new technological knowledge, to the combination of different technologies in one product, and to the number of actors involved in R&D process. In the context of Defence to give evidence of the complexity increase there are fighter aircraft (combination of knowledge from engines, electronics, weapon-system, structural design, material) or the massive use of ICT in weapon-system [11].

Security (SPS) Programme [15, 16]. Then, there is the RAND corporation, a non-profit research organization that carries on research and analysis to realize report on critical topics tackled by public policy, to help institutions around the World make communities safer and more secure, healthier and more prosperous [17, 18].

1.4 Conclusion

In conclusion, the fuzzy nature of the C4ISTAR domain and the connection with the Defence systems necessitate multiple resources for a comprehensive analysis with wider perspective, that are academic documentations from Research Institutions (e.g. scientific publications); official documents from National or International Institutions (e.g. publications by Ministry of Defence and by NATO); companies reports (e.g. annual reports, press release) and others (e.g. thematic articles, surveys).

2. C4ISTAR field delineation

2.1 Collection of documents

The collection includes 166 documents related to the C4ISTAR domain and dated from 2000 to 2020. The choice of the utilization of the different sources listed in the Paragraph 1.4 is linked to the fuzzy nature of the domain in analysis and brings other advantages. First, the reduction of biases sources effects. In fact, each source may be influenced by various factors, for example official documents by National or International Institutes may be influenced by political decisions. Then different sources allow also to increase recall (i.e. the fraction of relevant instances obtained in the research [19]), so the more documents are included, the more this parameter increases. Finally, finding the same information in different documents ensures the quality of data retrieval.

2.2 Exploration of descriptions and keywords

A part of the collected documents is manually analysed with the purpose to find the descriptions related to the C4ISTAR concept, so 67 descriptions are identified and listed. The documents to be read are chosen giving priority to the scientific publications and to the recent available versions. In addition, there are 171 entries related to the term of the C4ISTAR acronym from the NATOterm database[7].

Moreover, during the manual analysis 1.474 terms and expressions related to the C4ISTAR systems are identified and listed.

[7] It is the Official NATO Terminology Database, it "contains non-classified military terminology, as well as non-military terminology relevant to NATO. NATOTerm is available in both NATO official languages, i.e. English and French, and is maintained by the NATO Terminology Office, which is part of the NATO Standardization Office". It is available online at https://nso.nato.int/natoterm/content/nato/pages/home.html?lg=en

2.3 Analysis of descriptions

The 1999 version of the U.S. Army Field Manual defines Command and Control as "the exercise of authority and direction by a properly designated commanding officer over assigned and attached forces in the accomplishment of a mission". Then, the focus has moved to the "unity of effort" [20] and about "what to do … for creating value" [21] and "increasing Situational Awareness, giving decision-makers the information they need as fast as possible, and using the right materials, equipment and systems to make that happen. All the components of C4ISR MUST work together smoothly to achieve mission success" [22]. A quantitative measure of these considerations is given by the analysis on the frequency of terms used in the dataset of descriptions and by the examination of the taxonomy of the domains[8] in NATOterm database. So that, C4ISTAR includes science and technology applications and military operations. Considering the technological perspective C4ISTAR is a complex of communications systems, navigation systems, vehicles, drones, robots, autonomous systems, cyber systems, visual, electronic, photographic means, connected with the Battle Management System. Considering the architectural perspective C4ISTAR is the network to support land, air and naval troops and systems during missions and daily operations, based on the so-called Network Centric Approach, the C2 Agility and the OODA loop.

At this point it is possible to outline the definition of the C4ISTAR domain.

C4ISTAR stands for Command, Control, Communication, Computer, Information/Intelligence, Surveillance, Target Acquisition, and Reconnaissance. It indicates systems, structure, processes and all entities involved in missions and in military daily operations. As technological domain it includes science and technology applications related to military operations, sensor systems, as well information technologies and telecommunication for information sharing and data fusion. So that all the systems, tools and devices that create Situational Awareness (known as Common Operational Picture – COP) and support decision making as well as actions. But only technology is not enough, C4ISTAR needs also a well-constructed architecture to promote interoperability and integration in the context of Network-Centric-Warfare. Therefore, C4ISTAR embraces system engineering, components architecture, military doctrine and procedures both in missions and daily operations.

This outcome allows to point out the main field related to C4ISTAR as technological domain, i.e. communication, information, sensor systems in relation to data fusion and defence matter; and the main elements of the C4ISTAR systems, i.e. technology, user, application, purpose, requirement.

[8] Domains are groups to which terms relates as defined by NATOterm. The definition of the domains "deals with a broad range of issues but with an emphasis on political, security and military matters and the assets (personnel, equipment and infrastructure required to perform its activities)" [23].

3 C4ISTAR dictionary delineation

3.1 Selection of keywords

The list of keywords identified during the exploration of the collection of documents is compared with several dictionaries of the fields related to the C4ISTAR domain, as identified in the Paragraph 2.3. The dictionaries have been selected using the approach of the heterogeneity of sources, as already done for the documents collection: official institutions, academic resources, companies and others. An automatic lexical analysis, executed with RStudio and based on match finding [24, 25], measures the comprehensive of manual selection and filters the initial dataset.

The examination of the units of text (e.g. the full expression “personal computer” consists of two units of text, namely “personal” and “computer”) shows the correspondences between the words obtained by the list of keywords and the dictionaries mentioned above (both in terms and in definition). The dataset is composed by 1.252 unique words used in the list of keywords. Almost the 33% of words, precisely 412 out of 1.252, are not in any dictionary, so they are not related to the fields considered in this comparative analysis; the remaining 67% is mentioned in at least one dictionary. This outcome proofs the integrity and the comprehensive of the manual selection.

The examination of the full expressions gives more precise results for the filtering process. The dataset considered is composed by 1.468 unique expressions used in the list of keywords. Almost the 77% of the expressions, precisely 1.126 out of 1.468, are not in any dictionary. The remaining 23% is mentioned in at least one dictionary. Therefore, only those expressions are kept the list.

In both analysis, most of the terms are linked to information and communication systems and of course to defence, confirming the initial consideration about C4ISTAR as the technological domain of sensors, information and communication systems to detect and share data in warfare and defence operations. Additionally, the dictionaries themselves have been used for dataset expansion. In fact, the most recurrent terms in these dictionaries have been added to the list. Finally, it contains 446 expressions related to the C4ISTAR field.

3.2 Classification of keywords

Then the terms are manually classified using a top-down and a bottom-up approach, giving the foundation for the C4ISTAR dictionary. The former is a definitions comparison and consist in assigning a term to a class based on the correspondence between the term definition and the class definition. The latter is a semantic comparison, since how terms are written could suggest information to understand the main feature of each class.

The classes are the elements of the C4ISTAR systems, i.e. technology, user, application, purpose, requirement, as identified in the Paragraph 2.3. Moreover, the classification itself leads to the definition of two other classes, that are information and approach.

Finally, the similar expressions (i.e. terms related to the same concept) are grouped under the same label and a definition is provided for each label. At this point, in the C4ISTAR dictionary there are 343 entries distributed among the defined classes, 240 unique labels and the relative definitions.

3.3 C4ISTAR dictionary validation

The validation is executed on RStudio with a sematic comparison among the terms in the C4ISTAR dictionary and a list of keywords automatically extracted from the collection of documents. The 1.547 extracted keywords are compared with the C4ISTAR dictionary with an automatic orthographic analysis based on string distances, i.e. a computational algorithm that explores similarities and semantic equivalations among expressions, finding which keywords of the two lists are written in a similar way [26]. There are 138 equal expressions, 1.365 similar terms and only 44 different ones. Therefore, it is possible to validate both the adopted approach and the obtained results described in the Paragraphs 3.1 and 3.2. Moreover, the similar terms are used to expand the C4ISTAR dictionary. In conclusion, there are 412 terms and 273 unique labels and definitions distributed among the 7 classes.

4 C4ISTAR mapping

4.1 Exploration of relationship

A co-occurrence analysis, executed with RStudio, allows to find the pairs of terms most frequently used in the text corpus (i.e. a large and structured set of texts) in order to understand the relationships among the terms in the C4ISTAR dictionary to map the field. This is an extensions of occurrence analysis since it is based on the word counting on a higher dimension: for each given pair of words, called w_1 and w_2, an algorithm counts how many times w_1 is present in a text which contains w_2. The result is the co-occurrence matrix, a square matrix where rows and columns contain the words and each cell the number of co-occurrence between the two words. This matrix allows to define a weighted and directed graph where the words are the vertices and the probabilities that 2 terms co-occurred in a document ($\frac{number\ of\ documents\ where\ w_1 and\ w_2 co-occurred}{number\ of\ documents\ where\ w_1\ occurred + number\ of\ documents\ where\ w_2\ occurred} * 100$) are the weights of the edges [27]. Therefore, the results are exported for the software Gephi, an open-source software for graph visualization and analysis. This software proposes many layout algorithms, that set the graph shape according to the features of the topology to highlight [28]. The ForceAtlas2 algorithms emphasizes the complementarities among nodes to detect communities and to visualize in the network the circulation of the meaning. This is a force directed layout based on the attraction, the repulsion and the gravity forces: nodes repulse each other and edges attract their nodes and the movement converges to a balanced state, the gravity force compensates the repulsion preventing disconnected components to move away during the spatialization [29, 30]. Then the nodes are assigned to different communities with the Modularity algorithm. It is a method to detect communities in a network based on the collective proximity, so called modularity: the densely connected nodes are grouped together and the

nodes belonging to different communities are sparsely connected [31]. This algorithm detects 3 communities. Therefore, this representation, visualized in the Figure 1, could be interpreted as the map of the C4ISTAR domain, because it describes the principal functions of C4ISTAR systems, that are "detection", "elaboration" and "execution", respectively in purple, orange and green in Figure 1. In fact, the C4ISTAR systems are used to collect information from the context where missions and defence operations take place, to elaborate the collected information creating the Situational Awareness among all the actors involved, and to allow them to execute the orders derived from the comprehension of the context and finally to communicate the results. Then this last step makes the cycle starts again.

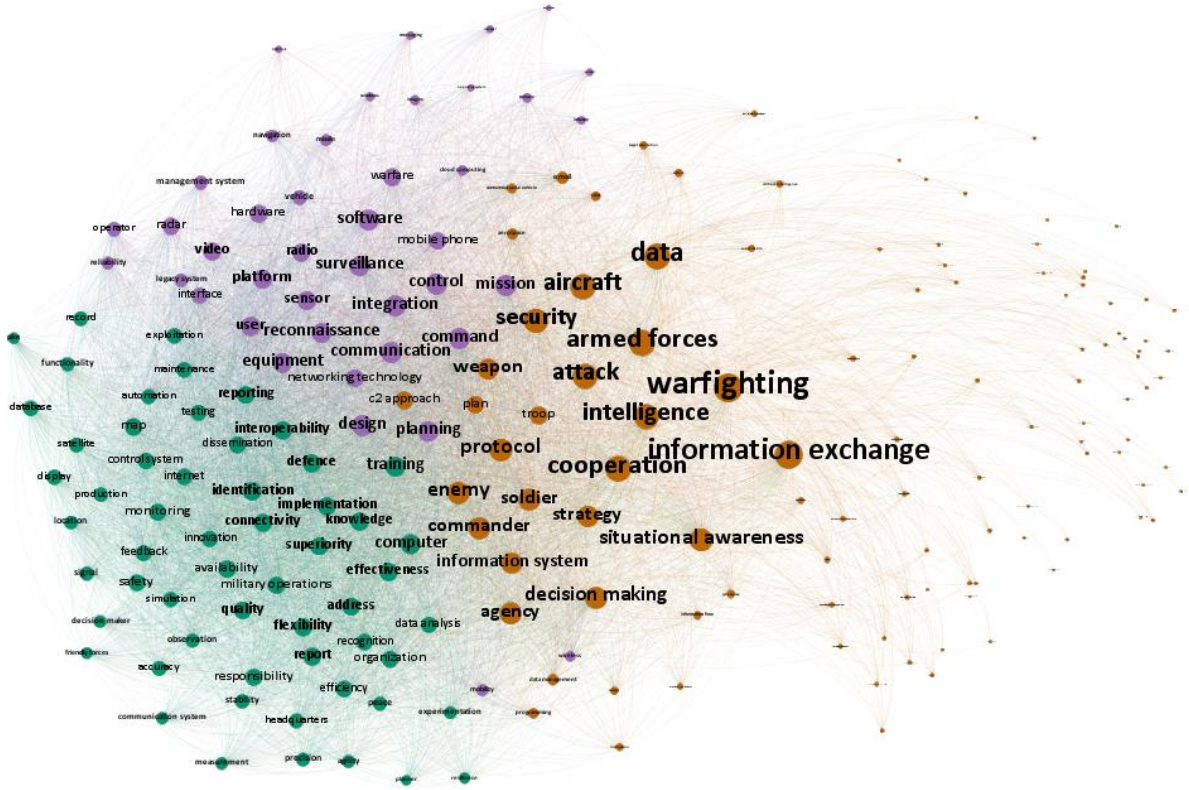

Figure 1 - Network of the C4ISTAR domain.

4.2 Exploration of structures

4.2.1 Lexical pattern analysis

A lexical pattern analysis, executed with automatic text analysis tools and manual examination, is used to identify the patterns of interaction among the entities involved in the C4ISTAR systems, exploring a dataset of 9.807 sentences in the collection of documents. The sentences contain the terms of at least 2 different classes, those with a greater co-occurrence and the verbs frequently used.

A sample of 126 sentences is manually analysed. First, the structure of the relationship is identified in each sentence and then it is defined as following: 2 entities (as classes in the C4ISTAR dictionary), one acts as starting point and the other as ending point, and a verb or a construct, that is the pattern of interaction; an example is presented in the Table 1. There are different types of patterns, grouped as their functional meaning: connection, improvement, influence, intention, location, source,

specification and usage. This result explains the interactions among entities (i.e. the relationships among classes of terms). In conclusion, the C4ISTAR systems are a combination of technologies and applications that allow the cooperation of users, influencing and improving their intentions; the users can get all the necessary information for the Situational Awareness; those systems are developed for certain usage in some context with some specification.

Sentence	Structure of relationships among entities	Starting entity	Pattern of interaction	Ending entity	Type of pattern
Decision makers need access to actionable intelligence **in** peace **coping with** interoperability constraints.	user **need** access to technology **in** purpose **coping with** requirement	user	needs	technology	**usage**
		user	in	purpose	**location**
		technology	coping with	requirement	**influence**

Table 1- Example of sentence with structures and patterns of interactions among entities in the C4ISTAR systems.

4.2.2 Experimentation and next steps

The exploration of structures can be executed also with the semantic hypergraph [32], a novel approach in text analysis, based on the recursive aspect of the natural language (i.e. the application of a grammar rule within its own definition [33]), that leads to the creation of a new concept from the existing ones. So, the knowledge can be intended as the relations among the concepts.

The full table of 9.807 sentences is processed with a demo of this tool. The result is a table where for each sentence there is the verb, the main and auxiliary concepts, the role[9] of the concept within the verb and some additional elements that define the relation, namely the triggers[10],the builders[11]. A manual exploration of the dataset allows to identify some new patterns of interactions (composed of the verbs and the triggers) and two measures for the structure, that are the percentage of the terms of one class related to the type of relationship and the percentage of the role of each term within its class related to the type of relationship.

These measures could be used for an automatic terms classification based on their roles within each type of relation. Moreover, the opposite approach, that is identifying the patterns of interaction based on the role of the terms within their class, may also be explored.

[9] The role of a concept within a verb could be: subject (s), passive subject (p), agent (a), subject complement (c), direct object (o), indirect object (i), specifier (x), parataxis (t), interjection (j) and clausal complement (r) [34].

[10] The triggers build an additional specification of a relationship, for example conditional ("We go *if* it rains."), or temporal ("John and Mary travelled to the North Pole *in* 2015"), or local ("Pablo opened a bar *in* Spain") [32].

[11] The builders are atoms that can combine several concepts, for example, the atomic concepts of "founder" and "psychoanalysis" can be combined with the builder atom "of" to produce the concept of "founder of psychoanalysis". The builders relate the main concepts and the auxiliary ones) [34].

An example of the obtained result is presented in the Table 2, where the blue cells indicate the entities involved in the relative relationships, while the red cells indicate the entities not involved in a specific relation. Several patterns involve some entities not yet classified.

Type of relation	Relation	Application	Technology	Requirement	Purpose	Approach	Information	User	Terms not classified
assess	assess								
	assess [by]								
	assess [in order to]								
	assess [in terms of]								
	assess [through]								
	examine								
	examine [in]								
	examine [through]								
	view								
	view [as]								
	view [through]								

Table 2 - Example from the map of patterns of interactions among entites in the C4ISTAR systems.

Conclusion

Facing with a fuzzy domain related to Defence requires a multi-dimensional approach for a comprehensive analysis. With the tools of the text mining it is possible to handle this issue overcoming the limits of the absence of a complete expert knowledge about this field. In fact, the various domains of knowledge that compose the C4ISTAR domain are not merely summed together, rather they are combined to generate a new domain with value added. Moreover, the connection with the Defence systems increase the complexity of this field, since it belongs to a strategic industry for each country.

The tool of the dictionary has been used in the context of the project ASSETs+ for the initial comprehension of the domain under analysis and for the technology roadmap as a source of keywords for technologies and applications in the C4ISTAR systems.

Future developments includes an analysis on the management tools used in Defence and applicable in other context, as a meaning of example the possible parallelism between the OODA loop and the PDCA cycle or the DMAIC methodology; the Network Centric approach as a base for the organisational structure in the context of the Industry 4.0 with the distributed system. Moreover, it is possible to go in deeper into a comparison between the evolution of Battlefield-of-Things (BoTs) and Internet of Things (IoT) to understand where the change take place first (i.e. in Defence applications or in civil application). Furthermore, there's a chance to perform an in-depth examination on the skills of the users involved in the C4ISTAR systems and a detailed analysis of the structures with the semantic hypergraphs theories.

References

[1] Vassiliou, M. S., Alberts, D. S., & Agre, J. R. (2014). *C2 re-envisioned: the future of the enterprise*. CRC Press.

[2] Commandant Australian Defence Force Warfare Centre (2009). *ADDP 00.1 Command and Control*. (first edition). Australia: Defence Publishing Service Department of Defence CANBERRA ACT 2600.

[3] Dunne, J. P. (1995). The defense industrial base. *Handbook of defense economics*.

[4] Callen, T. (2020). *Gross Domestic Product: An Economy's All*. Retrieved June, 2020 from https://www.imf.org/external/pubs/ft/fandd/basics/gdp.htm

[5] Lepenies, P. (2016). *The power of a single number: A political history of GDP*. Columbia University Press.

[6] World Bank (2020). *Military expenditure (% of GDP)* [database]. Retrieved June, 2020 from https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS

[7] SIPRI (2019). *Military expenditure by country, in constant (2017) US$ m., 1949-2018, ©SIPRI 2019*. SIPRI Military Expenditure Database [database]. Retrieved June, 2020 from https://www.sipri.org/databases/milex

[8] Hartley, K., & Belin, J. (Eds.). (2019). *The Economics of the Global Defence Industry*. Routledge.

[9] SIPRI (2019), *The SIPRI Top 100 arms-producing and military services companies in the world (excluding China).* SIPRI Arms Industry Database [database]. Retrieved June, 2020 from https://www.sipri.org/databases/armsindustry

[10] Schilling, M. A., & Shankar, R. (2017). *Strategic management of technological innovation*. (fifth edition). McGraw-Hill Education.

[11] Belin, J., Guille, M., Lazaric, N., & Mérindol, V. (2019). Defense firms adapting to major changes in the French R&D funding system. *Defence and Peace Economics*, *30*(2), 142-158.

[13] Dennis, M.A. (n.d.) *Defense Advanced Research Projects Agency UNITED STATES GOVERNMENT*. Retrieved June, 2020 from https://www.britannica.com/topic/Defense-Advanced-Research-Projects-Agency

[14] Belin, J., & Guille, M. (2019). Innovation Dynamics in Defence Industries. *Disruptive Technology and Defence Innovation Ecosystems*, *5*, 1-30.

[15] NATO Science and Technology Organization (2019), *2019 Highlights Empowering the Alliance's Technological Edge*.

[16] NATO Science for Peace and Security Programme (2018). *2018 Annual Report*. NATO Emerging Security Challenges Division.

[17] RAND Corporation (n.d.). *About the RAND Corporation*. Retrieved June 2020, from https://www.rand.org/about.html

[18] RAND Europe (2019). *SPOTLIGHT 2019-2020.*

[19] Precision and recall (2020). Retrieved June, 2020 from https://en.wikipedia.org/wiki/Precision_and_recall

[20] Alberts, D. S. (2007). *Agility, focus, and convergence: The future of command and control*. Office of the Assistant Secretary of Defense for Networks and Information Integration Washington DC.

[21] Alberts, D. S., & Hayes, R. E. (2006). Understanding Command and Control [Electronic Version]. *Washington, DC: Department of Defense Command and Control Research Program*.

[22] Ross B. (2018*). CS Command, Control, Communication, Intelligence, Surveillance and Reconnaissance*, New Jersey: Infosecforce.

[23] Jones, I. (2011). *The NATO Terminology Programme and NATOTerm.* Retrieved June, 2020 from http://www.mt-archive.info/Aslib-2011-Jones.pdf

[24] Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. O'Reilly Media, Inc.

[25] Grolemund, G., & Wickham, H. (2018). *R for data science*. O'Reilly Media, Inc.

[26] Van der Loo, M. P. (2014). The stringdist package for approximate string matching. *The R Journal*, *6*(1), 111-122.

[27] Bourgeois, N., Cottrell, M., Lamassé, S., & Olteanu, M. (2015, June). Search for meaning through the study of co-occurrences in texts. In *International Work-Conference on Artificial Neural Networks* (pp. 578-591). Springer, Cham.

[28] Gephi Tutorial Layouts (2011). Gephi documentation.

[29] Jacomy, M. Venturini, T.-Heymann, S.-Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE*, *9*(6).

[30] Badar, K., Hite, J. M., & Badir, Y. F. (2013). Examining the relationship of co-authorship network centrality and gender on academic research performance: the case of chemistry researchers in Pakistan. *Scientometrics*, *94*(2), 755-775.

[31] Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, *2008*(10), P10008.

[32] Menezes, T., & Roth, C. (2019). Semantic Hypergraphs. *arXiv preprint arXiv:1908.10784*.

[33] Recursion (2020). Retrieved June, 2020 from https://en.wikipedia.org/wiki/Recursion#In_language

[34] Semantic Hypergraph Notation (2020). Graphbrain. Retrieved June, 2020 from http://graphbrain.net/reference/notation.html