Computer Science & Information Technology 82

Natarajan Meghanathan
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

7th International Conference on Natural Language Processing
(NLP 2018) January 27~28, 2018, Dubai, UAE

AIRCC Publishing Corporation

Volume Editors

Natarajan Meghanathan,
Jackson State University, USA
E-mail: nmeghanathan@jsums.edu

Dhinaharan Nagamalai,
Wireilla Net Solutions, Australia
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403
ISBN: 978-1-921987-80-9
DOI : 10.5121/csit.2018.80201 - 10.5121/csit.2018.80212

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

The 7th International Conference on Natural Language Processing (NLP 2018) was held in Dubai, UAE, during January 27~28, 2018. The 7th International Conference on Software Engineering and Applications (JSE 2018), The 5th International Conference on Foundations of Computer Science & Technology (CST 2018), The 7th International Conference on Signal & Image Processing (SIP 2018) and The 5th International Conference on Artificial Intelligence & Applications (ARIA 2018) was collocated with The 7th International Conference on Natural Language Processing (NLP 2018). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The NLP-2018, JSE-2018, CST-2018, SIP-2018, ARIA-2018 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, NLP-2018, JSE-2018, CST-2018, SIP-2018, ARIA-2018 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the NLP-2018, JSE-2018, CST-2018, SIP-2018, ARIA-2018.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

Natarajan Meghanathan
Dhinaharan Nagamalai

Organization

General Chair

David C. Wyld	Southeastern Louisisna University, USA
Jan Zizka	Mendel University in Brno, Czech Republic

Program Committee Members

Abbas Jalilvand	Islamic Azad University, Iran
Ahmad Qawasmeh	The Hashemite University, Jordan
Alessio Ishizaka	University of Portsmouth, United Kingdom.
Ali Selamat	Universiti Teknologi Malaysia, Malaysia
Amizah Malip	University of Malaya, Malaysia
Anazida Zainal	Universiti Teknologi Malaysia, Malaysia
Antonio De Nicola	ENEA, Italy
Ayush Singhal	University of Minnesota, USA
Bouchra Marzak	Faculty of Sciences - Hassan II University, Morocco
Carmen Martinez	University of Jaen, Spain
Chih Lai	University of St.Thomas, USA
Christophe Nicolle	Université de Bourgogne Franche-Comté, France
Da Yan	The University of Alabama at Birmingham, USA
Edwin Lughofer	Johannes Kepler University Linz, Austria
Fulvia Pennoni	University of Milano-Bicocca, Ialy
Gabor Kiss	Obuda University, Hungary
Gahangir Hossain	Texas A&M University, USA
Gamini Wijayarathna	University of Kelaniya, Sri Lanka
Georgia Garani	TEI of Thessaly, Greece
Gheorghi Guzun	The University of Iowa, USA
Grigorios N. Beligiannis	University of Patras, Greece
Habibollah Haron	Universiti Teknologi Malaysia, Malaysia
Hadi Amirpour	Universidade da Beira Interior, Portugal
Haibo Yi	Shenzhen Polytechnic, China
Hamid Ali Abed AL-Asadi	Basra University, Iraq
Hamido Fujita	Iwate Prefectural University, Japan
HASSINI Noureddine	University of Mostaganem,Algeria
Hedieh Sajedi	University of Tehran, Iran
Hiromi Ban	Nagaoka University of Technology,Japan.
Hojjat Rakhshani	UHA University, France.
Irena Patasiene	Kaunas University of Technology, Lithuania
Ireneusz Czarnowski	Gdynia Maritime University, Poland
Israa Sh. Tawfic	University of Gaziantep,Turkey.
Issa Atoum	The World Islamic Sciences and Islamic Studies, Jordan
Iyad Alazzam	Yarmouk University, Jordan
Jia Zhu	South China Normal University, China
Jonice Oliveira	Universidade Federal do Rio de Janeiro (UFRJ), Brazil

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Networks & Communications Community (NCC)

Information Technology Management Community (ITMC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

7th International Conference on Natural Language Processing (NLP 2018)

7th International Conference on Software Engineering and Applications (JSE 2018)

5th International Conference on Foundations of Computer Science & Technology (CST 2018)

7th International Conference on Signal & Image Processing (SIP 2018)

5th International Conference on Artificial Intelligence & Applications (ARIA 2018)

TOWARDS MAKING SENSE OF ONLINE REVIEWS BASED ON STATEMENT EXTRACTION

Michael Rist, Ahmet Aker and Norbert Fuhr

Department of Computer Science and Applied Cognitive Science,
University Duisburg-Essen, Duisburg, Germany

ABSTRACT

Product reviews are valuable resource for information seeking and decision making purposes. Products such as smart phone are discussed based on their aspects e.g. battery life, screen quality, etc. Knowing user statements about aspects is relevant as it will guide other users in their buying process. In this paper, we automatically extract user statements about aspects for a given product. Our extraction method is based on dependency parse information of individual reviews. The parse information is used to learn patterns and use them to determine the user statements for a given aspect. Our results show that our methods are able to extract potentially useful statements for given aspects.

KEYWORDS

Aspect-based opinion extraction, dependency parse trees, dependency patterns

1. INTRODUCTION

Product reviews are a valuable resource for information seeking and decision making purposes. Online shops such as Amazon allow customers and users to add reviews to every product in their online catalogue. Certain reviews are direct feedback about a product and can guide new customers in their buying process. However, in most cases due to the immense number of reviews it is impossible for a customer to digest all of them. Several automatic processing methods of such reviews have been proposed. One intensively studied approach is the extraction of sentiment from reviews and summarization of positive and negative instances within the review set [6,7]. For instance, for the following review the sentiment approach would say that the review is positive:

The display is bright, colourful and has a high resolution.

However, sentiment extraction has a clear limitation in this case because it is not able to provide why the review is positive. Obviously, this review provides some pieces of evidence about a product and claims implicitly that the product is good because of those premises. Thus, a better approach than sentiment extraction would be a pipeline that is able to extract those premises and provide a quality judgment or claim about the product based on the extracted premises such as: *The product X is good because many think it has an excellent display.*

We present such a pipeline in Figure 1. Our pipeline shows the process of evaluating all reviews from one certain product. In the first step, it extracts premises which entails aspects and personal

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 01– 12, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80201

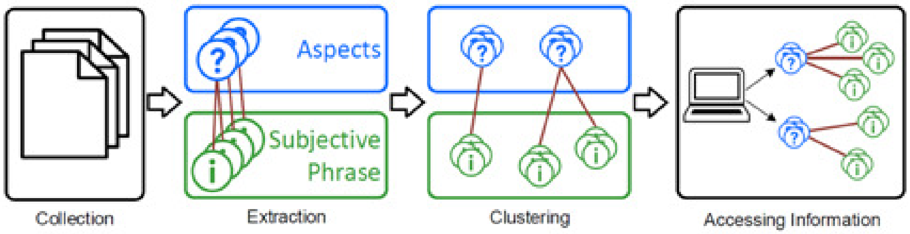


Figure 1. Information extraction pipeline

statements. Next, because the same aspect can be expressed in different ways (display, screen) it groups the different expressions of an aspect together to a broader one (e.g. display). The same is performed for all statements provided about an aspect. In the final step, it generates a summary about the product based on the aspects and statements. Our goal is to have a pipeline such that product reviews from arbitrary categories can be summarized. In this work, we focus at step one (extraction of premises) of the pipeline and leave the remaining steps for future work.

In our case a premise consists of an aspect and one or more personal statements. For instance, for the earlier example we have the aspect display and three statements: *bright*, *colourful* and *high resolution*. We assume in this work that aspects within reviews are already known and focus only on the automatic extraction of subjective phrases. Our statement extraction method is based on dependency parse trees. From the parse tree, we obtain generalized patterns that highlight the boundaries of statements and link them to an aspect within a review.

Patterns generated from dependency parse trees have been already investigated for extracting information from well-formed text [8,9,10] as well as in combination with aspect-based opinion mining [4,11,14]. However, to the best of our knowledge such patterns have not been applied to extract statements for given aspects.

The remainder of the paper is structured as follows. First, we take a short look at other approaches and methods used to process reviews for information extraction. After that we introduce the data we work with. Section 4 presents our technical solution to aspect relevant subjective phrase extraction, followed by Section 5 describing our experimental settings. Results are described and discussed in Section 6. We conclude our paper in Section 7.

2. RELATED WORK

Opinion mining and sentiment analysis is a wide research field and can be divided into different areas [3]. In terms of product reviews there has been a focus on aspect-based sentiment analysis [13,6,15]. In our work, we concentrate on aspect-based opinion mining and aim to extract statements for given aspects rather than sentiments. On this line, the work of Sauper et al. applies an LDA [1] model to simultaneously extract aspects and statements. Unlike us they are using rather clean data with one aspect per sentence and by only considering argumentative sentences, thus preemptively eliminating any noise in the data. Xu et al. also used LDA to jointly extract aspects and sentiments, however they also limit the aspects per sentence to one and extract them both at once. In our case the sentences can have more than one aspect as well as more than one statement in a sentence. We also do not assume that our sentences are argumentative. Furthermore, we apply patterns learned from dependency trees instead of LDA.

Dependency parse inspired patterns were used before in order to extract information from general texts [8,4,10] as well as online reviews [11,15,17]. In some of these studies the patterns are manually generated [4,11,15] and others learn them automatically from the data [8,10,17]. Fixed patterns are used both for learning or extracting aspects [11] and link aspects to statements [4]. Qiu et al. [15] apply relation patterns to find new aspects and statements. Their use of relations patterns is quite successful, but unlike ours has the clear restrictions of static patterns. Unlike our study generated patterns from previous studies have not been applied to extract personal statements for given aspects but rather used, for instance, to extract entity or sentiment related information.

Other approaches use the opposite direction, meaning that they search for aspects given certain, ambiguous, statement. Yauris at al. [19] for example apply the methods used in [15] to extract aspect from game reviews, however they statements are limited to adjectives only while our statements can be whole phrases. Hu et al. [20] uses a frequency based approach to extract aspects or features. The sentiment is given by an orientation and not the actual information like done here. The same underlying method was later also used and enhanced by Marrese-Taylor et al. [18] where they conducted a user study with a visual overview over the sentiment for each aspect.

3. DATA

The raw data is taken from Amazon reviews provided by [5]. These consist in total of 142.8 million reviews from which we annotated 400 randomly selected reviews. The reviews come from 4 different categories or representatively 4 different products with a sufficient review count. We annotated aspects and personal statements within the reviews. Statements are defined as certain assertions given by the reviewer. These can also be seen as a stated opinion or sentiment about some part of a product.

The aspect describes what part or characteristic of the product is being discussed. Aspects are also seen as an opinion target, like the ones used in [16]. All the reviews in our data were annotated by one single expert. Altogether we found 1,666 aspects and 1,987 statements within the annotated reviews. Among the reviews there are a few cases where the review contains only the aspect annotation and does not convey any statement. In our application scenario, we filter out such cases and focus only on reviews entailing both aspect and statements. The total number of reviews containing both annotation types is 1,966. In most cases a review contains only a single aspect and one or more statements. In this case all statements are linked to the single aspect. However, there are also cases where a review contains more than one aspect as the following example shows: *The keyboard and trackpad of this notebook is quite sturdy but not well designed*. This example sentence contains the aspects *keyboard* and *trackpad*. The statements are quite sturdy and *not well designed*. Since there are two aspects both statements are regarded as connected to each of the aspects. We use these 400 reviews to learn patterns based on dependency parse information. These patterns are in turn used to automatically extract subjective statements as well as to link them to aspects.

Table 1. Annotated Data

Categories	Claims	Premises	Relations
SD Card	333	399	396
Earphones	456	549	549
Keyboard	427	535	517
E-reader	450	504	504
all	1,666	1,987	1,966

4. METHODS

The task of extracting the complete statements is split into two successive steps. First, we identify the position of a statement within the sentence and afterwards we limit the borders of the statement. This limitation is needed because a statement might not be given in a single word and can consist of a certain part of the containing sentence. Looking at the previous example sentence: *The keyboard and trackpad of this notebook is quite sturdy but not well designed*, the statements are limited to the words *quite sturdy* and *not well designed*. When retrieving only a partial statement its meaning might be drastically altered. By excluding, for instance, the word *not* in the second statement the meaning is inverted and the actual information is lost.

For learning patterns, we use dependency parse trees, which we obtain using the DKPro framework [2], and word types (POS) for each word. Example dependency trees are shown in figures 2 to 4. Note that for POS tags with multiple variations, like nouns, we abstracted them to one general form. For instance, the sentences *The display is bright* and *The displays are bright* have the noun (aspect) *display* described by and adjective, the statement *bright*. When looking at the POS-tag of the aspects we have a tag for a singular and plural noun. Using the specific POS-tags would generate two different patterns. To avoid this, we simply use an abstract NOUN as word type for this node in the aspect.

Note, the information of the quantity is not needed for our purposes, as in the extraction and outcome each of these nouns are connected to the adjective, giving us the information how each aspect is described by its statement. This means that the correct noun, whether plural or singular will be linked to the adjective and so the information of the quantity is still present.

Figures 2 to 4 show basic examples of possible patterns and relations between the aspect and statement. Aspects are shown in blue, statements in green, the relations are marked as red edges. In general patterns are stored as a list of 3-tuples, $\langle wt(w_l), d, p(w_r)\rangle$, where each tuple describes an edge in the tree. The direction d is used to identify if the word on the right-hand side w_r of the tuple is a predecessor or successor of the current word w_l. This is done as each pattern has a leftmost node which is used as a kind of starting point or anchor for this pattern. $wt(w)$ simply gives the word type of the actual word w. $p(w)$ is a function to evaluate the word connected via the described edge and is described as:

$$p(x) = \begin{cases} \langle wt(w), d, \{p(s_1), p(s_2), \dots\}\rangle, & \text{if } w \text{ has sub-nodes} \\ wt(w), & \text{otherwise} \end{cases}$$

with s_x being the x-th succeeding word after word w in the pattern.

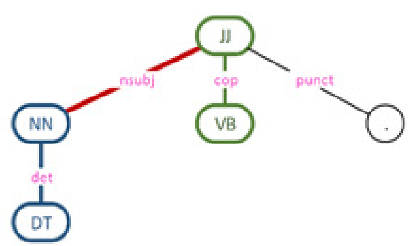


Figure 2. Dependency tree: *The display works great*.
Extracted linking pattern: $\langle NN, +, JJ\rangle$
Extracted limiting pattern: $\langle JJ, -, VB\rangle$

A simple, single pattern like seen in figure 2 contains the word type of the two connected nodes $wt(w_l)$ and $wt(w_r)$ and the direction of the link d. More complex patterns, as those in figure 3 and figure 4, are nested. This indicates which succeeding edges are needed to link the aspect to the statement in these cases.

Besides describing the relationships between an aspect and statement, the patterns can also be used to describe a statement. This allows us to minimize the complexity of our pattern as not only to find a link between an aspect but also to extract a complete statement. The whole process of extracting statements is divided into two steps: head identification of a statement, the linking to an aspect, and boundary detection, the limiting of a statement. Both limiting and linking steps are detailed in the following sections.

4.1 Limiting

The limiting of a statement defines its length and content. In order to extract only the relevant information, we need to distinguish the relevant part of a sentence from the irrelevant ones. To achieve this, we use the underlying dependency within a statement. A statement consists of several words forming a logical and rhetorical structure and have one certain root node. By determining this root node, we can extract a subtree containing all the words from the statement. Based on this subtree we create a pattern describing the word types and dependencies of the words within the statement. In figure 3 we can see the statement with one root node, the noun(NN) *colours*. The adjective *bright* is linked as an adjectival modifier (*amod*) to this noun. We can use these pieces of information to limit our pattern to $\langle NN, -, JJ \rangle$. Apart from determining the boundaries of patterns we can also use these root nodes as a clear target for the preceding step, the linking between an aspect and its statements which we describe in the next section.

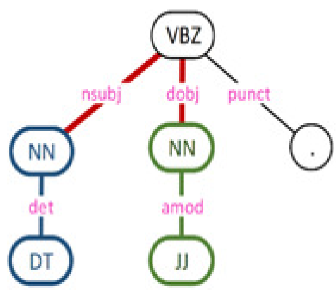


Figure 3. Dependency tree: *The Display has bright colours*.
Extracted linking pattern: $\langle NN, +, \langle VBZ, -, NN \rangle \rangle$
Extracted limiting pattern: $\langle NN, -, JJ \rangle$

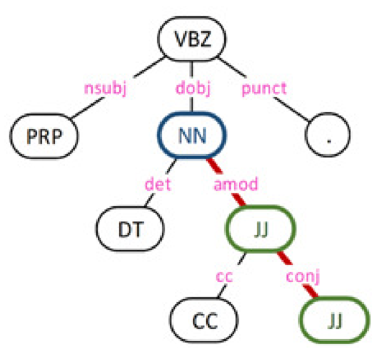


Figure 4. Dependency tree: *It has a bright and colourful display.*
Extracted linking patterns: $\{\langle NN, -, JJ \rangle, \langle NN, -, \langle JJ, -, JJ \rangle \rangle\}$
Extracted limiting patterns: $\{\langle JJ \rangle, \langle JJ \rangle\}$

4.2 Linking

Patterns are also used for linking an aspect to its statements. Similar to the statement extraction we also determine the root nodes of the aspects as a start of our patterns. For each pair of a given aspect and extracted statement we now have two root nodes for which we extract a linking pattern. In most aspects in the review there is only one statement given. For these cases, the path from the aspect to the head of the statement is taken as the pattern. These cases as well as the extracted linking patterns are shown in Figures 2 and 3. However, reviews might have complex structures such as containing more than one subjective phrases (see Figure 4). For these cases, we generate several patterns where each pattern captures only one path between the aspect and the head of each existing subjective phrase.

4.3 Selecting patterns

For both of our steps we have to select the right patterns to apply. This is needed because the extracted patterns can partially overlap each other. When looking again at the example linking patterns from figure 2, *<NN, +, VBZ>,* and figure 3, *<NN, +, <VBZ, -, JJ>>,* we can see that both patterns describe the first edge identical. The first pattern however ends after this edge, while the second pattern continues with another edge. In cases where we can apply the second pattern we could also apply the first one. Therefore, we have to prefer some patterns over others to increase overall performance as well as to have general patterns as much as possible. To achieve this, we use support and accuracy, as well as a combination of both computed over the patterns.

Support The support of a pattern states how often this pattern is observed. Quite common is a linking pattern like seen in figure 3 which is extracted from the sentence: *The display has bright colours*. This pattern is received from every sentence that has a sentence structure: *ASPECT* VERB *STATEMENT*. Instead of adding a new pattern each time we increase the support of the first pattern. The support of a limiting pattern is calculated similarly. Each occurrence of a pattern increases the support of it.

Accuracy The accuracy of a pattern is calculated by evaluating how often a pattern can be correctly applied in our data. When we apply a linking pattern we only know the head node of the aspect. When we look at the patterns from figure 2 and 3 the aspect head node has the same type. Assuming we only have those two linking patterns, we can apply the first pattern not only in the first example but also in the second as we have the same edge from the noun to the verb. This would result in one correct linking and one false linking and would achieve an accuracy of 0.5 For the limiting patterns, we proceed similarly.

Average accuracy and support These support and accuracy values are used to rank the patterns in order to determine the best ones. Additionally, we propose a third ranking by averaging over the normalized accuracy and support. The normalized accuracy a_n and support s_n are calculated by

$$a_n(p_x) = \frac{a(p_x) - \min_{p \in P} a(p)}{\max_{p \in P} a(p) - \min_{p \in P} a(p)} \text{ and}$$

$$s_n(p_x) = \frac{s(p_x) - \min_{p \in P} s(p)}{\max_{p \in P} s(p) - \min_{p \in P} s(p)},$$

with $a(p)$ as the accuracy, $s(n)$ as the support and P as the set of all patterns.

Threshold For the task of limiting a statement we use the best pattern (most highly ranked pattern) to select a single statement. However, when we want to link the statements to the aspects we have the problem that there can be multiple links per aspect and using the most highly linking pattern does not resolve the problem. Figure 4 for example has two statements *bright* and *colourful*. When we select only one linking pattern we can only retrieve one of the statements. To retrieve both statements we, have to apply more than one linking pattern. We determine the number of patterns that need to be applied using an adaptive threshold t_a This threshold is calculated by $t_a = rank\ (p_h)\ .\ (1 - r)$, where $rank\ (p_h)$ is the value of the highest matching pattern and r is the percentage of decline which we allow. For our linking patterns, we allow a 10% decline in performance.

Table 2. Results of the predicted links

Ranking	P	R	F1	P@10	P@20	P@50
LinkBaseline	.43	.43	.43	-	-	-
Accuracy	.54	.47	.50	.69	.64	.64
Support	.41	.34	.37	.28	.32	.39
Acc. & Sup.	.48	.44	.46	.29	.33	.38

5. EXPERIMENTAL SETTINGS

As we mentioned in Section 4 we separate our approach for extracting the statements into two elementary steps: linking to the location of a statement and limiting the extracted statement. For each step, we compare our results with a different straightforward and robust baseline. For obtaining patterns, as well as for the evaluation of both steps, we use the gold standard data described in Section 3. To evaluate the significance of our results we use a pairwise McNemar test[12] with Bonferroni correction.

5.1 Evaluation setup

To evaluate the performance of our statement extraction we apply 10-fold cross validation. Note that we keep in each fold only the patterns that occur at least twice. Patterns occurring less frequently in our training set, are not used for statement extraction. This is done to eliminate possible annotation and grammatical errors from our reviews. We compute precision, recall and F1-measure to quantify the performance of our pattern extraction approach. Additionally, as we can rank our retrieved patterns, we calculate the precision at 10, 20 and 50 to evaluate for the quality of the used ranking methods.

5.2 Baseline for linking (LinkBaseline)

As a baseline for finding the statements, we extract the nearest adjective and determine if this adjective is contained in the searched statement. This is a rather simple approach as we do not have any means of limiting a statement based on the adjective, but it will be sufficient enough for detecting the general area where a statement is located. For our previously chosen example from figure 2 we assume the linking is correct if for the aspect *display* the adjective *great* is chosen as the link target.

5.3 Baseline for limiting statement (LimitBaseline)

For the limiting step, we decided to use the dependency subtree of the root node as a baseline. More precisely we extract every word directly or indirectly dependent from the root node as part of the statement. This is again a quite simple baseline and therefore we allow for some noise. We define noise as additional words retrieved in an extracted statement. For instance, for the example sentence: *The display has bright colours* in figure 3 instead of only allowing the statement *bright colours* for the aspect display we also allow *has* leading to *has bright colours* as the statement for this baseline.

6. RESULTS

As previously mentioned we first look at the results of the individual steps and then regard the performance of the whole statement extraction step.

Table 3. Performance of the statement limitation methods

Ranking	**P**	**R**	**F1**	**P@10**	**P@20**	**P@50**
LimitBaseline	.51	.48	.50	-	-	-
Accuracy	.46	.46	.46	.64	.68	.73
Support	.26	.21	.23	.64	.46	.24
Acc. & Sup.	.35	.29	.32	.55	.62	.49

Table 2 shows the performance of our linking step. From the results, we see that best performance is achieved when accuracy alone is used to rank patterns. The support ranking performs overall worse than all the others, including the baseline. When we look at the precision at position 10, 20 and 50 we see that the accuracy ranking has only a small drop in the precision from precision@10 to precision@20. The support and acc. & sup have an increased precision for position 10 to 20, but nevertheless they are still vastly outperformed by the accuracy ranking.

Results for limiting a statement are shown in Table 3. We evaluated only the exact matches between the extracted statements and the gold standard. As we see from the table the baseline performs quite well and is, overall better than our best results. The accuracy ranking outperforms our other rankings by more than 10% in the precision, recall and F1 score. This may be contrary to the intuition, as the support of a pattern indicates its popularity, and therefore should improve the recall. Relying on the most frequent pattern should also receive the most correct results. The data however shows a significantly (p value) worse performance for the support compared to the accuracy ranking.

Table 4 shows the results for the complete extraction process (statement extraction and aspect linking) with different noise levels. Noise, as described in chapter 5.3, is additional words extracted along our statements. In our testing data, we have aspect and statement pairs. In the complete extraction process we aim to determine such pairs too. If the extracted pair is correct according to our evaluation criteria then we have a positive extraction, otherwise the extracted pair is considered as incorrect. From the results, we see again that the accuracy performs best for all the metrics. Contrary to the previous results, overall performance drops noticeably, from a F1 score of about .50 to only .31 for the precise results without the noise. However, when we allow more noise, our results improve by .07 points in the precision, recall and F1 score. When comparing the results from the different noise levels incrementally, we have a significant improvement (p-value $< .01$) between each noise level step. Furthermore, we can see that our ranked results perform quite stable with a precision of over 50%.

Table 4. Results for extracting statement

Noise	Ranking	P	R	F1	P@10	P@20	P@50
0	Accuracy	.32	.30	.31	.52	.53	.50
	Support	.09	.07	.08	.00	.05	.06
	Acc. & Sup.	.14	.12	.13	.00	.01	.14
1	Accuracy	.36	.34	.35	.52	.53	.51
	Support	.09	.07	.08	.00	.05	.06
	Acc. & Sup.	.22	.20	.21	.00	.01	.20
2	Accuracy	.39	.37	.38	.52	.53	.52
	Support	.10	.08	.09	.00	.05	.06
	Acc. & Sup.	.24	.22	.23	.00	.01	.22

6.1 Discussion

Our results show that ranking of the pattern has an enormous influence on the performance of the extraction methods. The large performance drop between the separate steps and the complete extraction indicates that, although the individual patterns perform rather good, the selection of the correct pair of patterns can be improved. Increasing the noise level in the statements largely improves our results. It shows us however that improving the patterns and their selection could lead to further improvement as either the link is not complete or the patterns are too vague for a better extraction. Either way, this shows that there is room for improvement.

On this line we performed an error analysis. We manually inspected statements which were extracted by our patterns. Table 5 shows some of these statements. Most of the shorter statements, with one or two words, are correct and even the longest and most complex one is extracted completely. Some extracted statements like *great and reasonable* for the aspect *price* were most likely extracted by the wrong pattern. The whole sentence is the following: *Besides that this card is great and very reasonable price of $50*. The statement *great* references to the aspect *card*, but without the knowledge about the first part of the sentence this statement could also be related to the aspect price.

Another area that requires further Sattention is the problem with erroneous reviews. We have seen several reviews that were problematic and yielded wrong dependency parse trees. We aim to implement detection methods for these erroneous cases, so that we can exclude them from processing.

7. CONCLUSION & FUTURE WORK

In this work, we described the extraction of aspect-based statements from product reviews through patterns extracted from dependency parse trees. We introduced methods for identifying the head of a statement and detecting the boundary for the statement given the head. Our evaluation results show that the best method for choosing reliable patterns in both steps separately, as well as at once, is the accuracy of the pattern.

Above, we already discussed some venues for improvement. In addition to these we also want to tackle the automatic extraction of aspects. Finally, we aim to use the aspects as well as all their assigned statements to generate summaries. Such summaries can be used by customers to satisfy their information needs and help them in their decision making purposes.

Table 5. Example extracted subjective phrases

Aspect	Extracted statement	Correct statement
price	fair	fair
	very reasonable	very reasonable
	great and reasonable	*very reasonable*
	low	low
	n't beat the price	*can't beat*
	price matches quality	*Matches the quality well*
battery life	really good	really good
	lasted through the movie and several episodes of a tv show	lasted through the movie and several episodes of a tv show
	awesome	awesome

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant No. GRK 2167, Research Training Group "User-Centred Social Media".

REFERENCES

[1] Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of machine Learning research 3(Jan), 993–1022 (2003)

[2] Eckart de Castilho, R., Gurevych, I.: A broad-coverage collection of portable nlp components for building shareable analysis pipelines. In: Proceedings of the Workshop on Open Infrastructures and Analysis Frameworks for HLT. pp. 1–11. Association for Computational Linguistics and Dublin City University, Dublin, Ireland (August 2014), http://www.aclweb.org/anthology/W14-5201

[3] Chhabra, S., Bedathurb, S.: Summarizing entities: A survey report

[4] Gindl, S., Weichselbraun, A., Scharl, A.: Rule-based opinion target and aspect extraction to acquire affective knowledge. In: Proceedings of the 22nd International Conference onWorld Wide Web. pp. 557–564. ACM (2013)

[5] He, R., McAuley, J.: Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. CoRR abs/1602.01585 (2016), http://arxiv.org/abs/1602.01585

[6] Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 168–177. KDD '04, ACM, New York, NY, USA (2004), http://doi.acm.org/10.1145/1014052.1014073

[7] Kiritchenko, S., Zhu, X., Cherry, C., Mohammad, S.: Nrc-canada-2014: Detecting aspects and sentiment in customer reviews. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). pp. 437–442. Association for Computational Linguistics and Dublin City University, Dublin, Ireland (August 2014), http://www.aclweb.org/anthology/S14-2076

[8] Li, P., Jiang, J., Wang, Y.: Generating templates of entity summaries with an entity-aspect model and pattern mining. In: Proceedings of the 48th annual meeting of the Association for Computational Linguistics. pp. 640–649. Association for Computational Linguistics (2010)

[9] Li, P., Wang, Y., Jiang, J.: Automatically building templates for entity summary construction. Information Processing and Management 49(1), 330 – 340 (2013), http://www.sciencedirect.com/science/article/pii/S0306457312000568

[10] Lippi, M., Torroni, P.: Context-independent claim detection for argument mining. In: Proceedings of the 24th International Conference on Artificial Intelligence. pp. 185–191. IJCAI'15, AAAI Press (2015), http://dl.acm.org/citation.cfm?id=2832249.2832275

[11] Maharani,W.,Widyantoro, D.H., Khodra, M.L.: Aspect extraction in customer reviews using syntactic pattern. Procedia Computer Science 59, 244–253 (2015)

[12] McNemar, Q.: Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12(2), 153–157 (Jun 1947), https://doi.org/10.1007/BF02295996

[13] Moghaddam, S., Ester, M.: Opinion digger: An unsupervised opinion miner from unstructured product reviews. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management. pp. 1825–1828. CIKM '10, ACM, New York, NY, USA (2010), http://doi.acm.org/10.1145/1871437.1871739

[14] Moghaddam, S., Ester, M.: On the design of lda models for aspect-based opinion mining. In: Proceedings of the 21st ACM international conference on Information and knowledge management. pp. 803–812. ACM (2012)

[15] Qiu, G., Liu, B., Bu, J., Chen, C.: Opinion word expansion and target extraction through double propagation. Computational linguistics 37(1), 9–27 (2011)

[16] YING, D., Yu, J., Jiang, J.: Recurrent neural networks with auxiliary labels for cross-domain opinion target extraction (2017)

[17] Zhuang, L., Jing, F., Zhu, X.Y.: Movie review mining and summarization. In: Proceedings of the 15th ACM International Conference on Information and Knowledge Management. pp. 43–50. CIKM '06, ACM, New York, NY, USA (2006), http://doi.acm.org/10.1145/1183614.1183625

[18] Edison Marrese-Taylor, Juan D. Velásquez, Felipe Bravo-Marquez, A novel deterministic approach for aspect-based opinion mining in tourism products reviews, In Expert Systems with Applications, Volume 41, Issue 17, 2014, Pages 7764-7775, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2014.05.045.

[19] K. Yauris and M. L. Khodra, "Aspect-based summarization for game review using double propagation," 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA), Denpasar, Indonesia, 2017, pp. 1-6. doi:10.1109/ICAICTA.2017.8090997, http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8090997&isnumber=8090950

[20] Hu, M., & Liu, B. (2004, July). Mining opinion features in customer reviews. In AAAI (Vol. 4, No. 4, pp. 755-760).

[21] Xu, X., Cheng, X., Tan, S., Liu, Y., & Shen, H. (2013). Aspect-level opinion mining of online customer reviews. China Communications, 10(3), 25-41.

AUTHORS

M.Sc. Michael Rist,
Research & Teaching assistant,
Workgroup Information Engineering,
Department of Computer Science and Applied Cognitive Science
University Duisburg-Essen

Dr.Ahmet Aker,
Research & Teaching assistant,
Workgroup Information Engineering,
Department of Computer Science and Applied Cognitive Science
University Duisburg-Essen

Prof. Dr. Norbert Fuhr,
Full Professor,
Workgroup Information Engineering,
Department of Computer Science and Applied Cognitive Science
University Duisburg-Essen

A DOMAIN INDEPENDENT APPROACH FOR ONTOLOGY SEMANTIC ENRICHMENT

Tahar Guerram and Nacima Mellal

Departement of Mathematics and Computer Science,
University Larbi Ben M'hidi of Oum El Bouaghi - ALGERIA

ABSTRACT

Ontology automatic enrichment consists of adding automatically new concepts and/or new relations to an initial ontology built manually using a basic domain knowledge. In a concrete manner, enrichment is firstly, extracting concepts and relations from textual sources then putting them in their right emplacements in the initial ontology. However, the main issue in that process is how to preserve the coherence of the ontology after this operation. For this purpose, we consider the semantic aspect in the enrichment process by using similarity techniques between terms. Contrarily to other approaches, our approach is domain independent and the enrichment process is based on a semantic analysis. Another advantage of our approach is that it takes into account the two types of relations, taxonomic and non taxonomic ones.

KEYWORDS

Ontology, Ontology learning, Semantic enrichment, Natural language processing.

1. INTRODUCTION

According to AI community, ontology is a formal explicit specification of a shared conceptualization. [1]. Ontology enrichment is the task of extending an existing ontology with additional concepts and semantic relations and placing them at the correct position in the ontology [2]. Ontology learning is a wide research area that contains ontology enrichment, Ontology population and inconsistency resolution [2]. Ontology construction and maintaining, is a fastidious knowledge acquisition task which gives always a bottleneck problem, namely when the dynamicity of the ontology domain is high. In the other hand, because of the development of the world wide web, textual information is available with huge quantities. Hence, It will be very useful if this task is achieved automatically or semi automatically from textual sources. But automating ontology enrichment is not an end in itself but the objective is to preserve the coherence of the ontology after the enrichment process and the best way to achieve this is to consider the semantics of used texts.

In this paper, we propose an approach for semantic ontology enrichment. We begin by building an initial (or basic) ontology using a basic knowledge about a target domain. The semantic enrichment of this basic ontology is done through both syntactic and semantic analysis of a corpus of texts relating to the same target domain. Syntactic analysis is accomplished using natural language processing tools to obtain a POS tagged and named entity annotated corpus. We mention that before applying NLP tools, preprocessing operations of the studied corpus are

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 13– 19, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80202

applied like stop words eliminating and words stemming. For each sentence of this annotated corpus we extract a short sentence obeying to the form **<*S_t V_t O_t*>** where S_t is the subject, V_t is the verb and C_t is the object. This short sentence plays the role of a relation $\mathbf{V_t(S_t, O_t)}$ which will be matched to the content of the basic ontology < ***S_o V_o C_o*** >. This matching is achieved using WorldNet resource by means of a semantic similarity measure and allows us to enrich the basic ontology as depicted in Figure 1 below.

The remainder of this paper is organized as follow. Section two is devoted to clarify the ontology learning bases. Section three gives a summary of related work; section four presents our approach. Finally, we conclude and we give some future work in section five..

2. ONTOLOGY LEARNING

In computer science, ontologies aim explaining and describing the world around us. However, in reality, they only focus on a part of the world, what is called a domain. The knowledge representation community defines ontology in accordance as follows: “Ontology is a formal, explicit specification of a shared conceptualization” [1]. ‘Conceptualization’ refers to an abstract model of phenomena in the world by having identified the relevant concepts of those phenomena. ‘Explicit’ means that the type of concepts used, and the constraints on their use are explicitly defined. ‘Formal’ refers to the fact that the ontology should be machine readable. ‘Shared’ reflects that ontology should capture consensual knowledge accepted by the communities.” Simply, ontology represents the knowledge by a set of the concepts within a domain of interest and the relationships between those concepts. For that aim, ontologies play a central role in knowledge extraction, they can be learnt from various sources, be it databases, structured and unstructured documents or even existing preliminaries like dictionaries, taxonomies and directories.

To a large extent, the ontology learning system is understood in a variety of ways, it can be ontology extraction, ontology generation, or ontology acquisition. Nevertheless, ontology learning can be defined as an automatic or semi-automatic creation of ontologies, including extracting the corresponding domain's terms and the relationships between the concepts that these terms represent from a corpus of natural language texts, and encoding them with an ontology language for easy retrieval.

Ontology enrichment is one of the important objectives for the ontology learning process. It consists of adding automatically new concepts and new relations to an initial ontology constructed manually using a basic knowledge relating to a given domain. Concepts and relations have to be placed in the relevant place in the initial ontology. However, numerous approaches and applications focus only on constructing taxonomic relationships (is-a-related concept hierarchies) rather than full-fledged formal ontologies [3]**.** For that, we are interesting, in our work to develop an approach for the ontology enrichment taking in account both taxonomic and non-taxonomic relationships between concepts. Generally, the process of enrichment attempts to facilitate text understanding and automatic processing of textual resources, moving from words to concepts and relationships. It starts by extracting concepts/relationships from plain text using linguistic processing such as part-of-speech (POS) tagging and phrase chunking [4]. The extracted concepts and relationships are then arranged in the initial ontology, using syntactic and semantic analysis techniques.

3. RELATED WORK

The first paper [5] presents a methodology called PACTOLE (Proprietary And Class Characterization from Text for Ontology Enrichment) for the enrichment of an initial ontology from a collection of texts relating to the astronomic domain. The first step is analyzing the collection of texts using NLP techniques in order to extract objects of the domain and their properties using predefined syntactic patterns then in the second step FCA technique is applied to the couples (object, property) in order to generate a concept lattice where each concept is a collection of a maximum number of objects sharing the maximum number of properties. The third step consists of expressing existing celestial objects data base by a second lattice of concepts using FCA technique as well as. The fourth step consists of merging the two lattices of concepts to obtain a resulting concept hierarchy. In step five this concept hierarchy is represented in FLE description language to be able to do reasoning tasks on it. The methodology was applied on a high number of Astronomy Abstracts journals and with the existing SIMBAD celestial objects database and the score of precision is high (74.71%) meaning that objects are classified in adequate classes and the score of recall is low because, mainly, the number of properties associated with objects is not sufficient for classification.

The second paper [6] presents a framework called Ontorich allowing the enrichment and the evaluation of ontologies using RSS Feeds. The enrichment of an otology is proceeded using OpenNLP API, which is a natural language processing Library, and WordNet [7] resource. RSS feeds are an important source of information as they provide permanently updated web information. To extract relations and concepts from RSS feeds, statistical and syntactic methods are applied using OpenNLP API. After the enrichment phase, the author(s) used several metrics to measure how the initial ontology is modified. Ontorich was compared against two ontology enrichment systems, which are Kaon and Neaon, and compared also against two other ontology evaluation systems, which are OntoQA and Romeo, relating to a certain number of functionality criteria and the results show that Ontorich is a more powerful tool for ontology enrichment and evaluation.

The third paper [8] presents a framework based on machine learning strategy to the automatic extraction of non taxonomic relations which remain a great challenge for ontology learning systems community. The framework proposed, initially extracts a set of causation contextual constructs (CCC) from annotated corpus and WordNet [7] to be used as initial indicators that can locate the good candidate sentences that may hold causation relation in text. In the second step, a new algorithm (graph based semantics GBS) is applied to indicate the real existence of causation in the sentences and if so, label both relation parts (cause, effect). To achieve this, sentences are divided into two parts and the most representative word in each part is searched based on the hypernym structure. The main steps of this algorithm are:

1. Specify causation relations direction according to CCC (cause-effect or effect-cause).

2. Extract the window for each relation part (for both cause and effect).

3. Build a graph for each window.

4. Specify the RDB (relation data base containing examples of cause, effect) semantic pair, suitable for the window.

5. Process each window graph to find best candidate semantic feature from the graph.

6. Extract a representative noun in the window that corresponds to the semantic feature.

To evaluate their system, precision, recall, and F-measure are computed based on a set of a total of 1213 used sentences and the results were as follow: precision = 78 %, recall = 68% and F-measure = 73 %.

The fourth paper [9] proposes an automatic process for ontology population from a corpus of texts. The proposed process is independent from the domain of discourse and aims to enrich the initial ontology with non taxonomic relations and ontology class properties instances. This process is composed of three phases: identification of candidate instances, construction of a classifier and classification of the candidate instances in the ontology. The "Identification of instance candidates phase applies natural language processing techniques to identify instances of non taxonomic relationships and properties of an ontology by annotating the inputted corpus. The "Construction of a Classifier" phase applies information extraction techniques to build a classifier based on a set of linguistic rules from ontology and queries on a lexical database. This phase has a corpus and an ontology as inputs and outputs a classifier used in the "Classification of Instances" phase to associate the extracted instances with ontology classes. Using this classifier, an annotated corpus and the initial ontology, the third phase consisting of the classification of these instances, produces a populated ontology.

Implementation of this process applied to the legal domain show results of 90% as precision 89.50% as Recall and 89.74% as F-measure. Authors conducted others experiments of their process on the touristic domain and obtained the results of 76.50% as precision 77.50% as Recall and 76.90% as F-measure.

The fifth and the last, but not least, paper [10] presents a pattern based approach of ontology enrichment by antonymic relations extracted from Arabic language corpora. Ontology of "seed" pairs of antonyms is used to extract lexicon - syntactic patterns in which pairs of antonyms occur. These patterns are then used to find new antonym pairs in a set of Arabic language corpora. The approach is tested on three different Arabic corpora: classical Arabic corpus (KSUCCA) [11], the contemporary Arabic corpus (CCA) [12] and the mixed Arabic corpus (KACSTAC) [13]. The correctly extracted patterns are used to enrich an ontology based lexicon for Arabic semantic relations called SemTree [14]. The developed system has as input the set of patterns and the KSUCCA corpus. First the given corpus is preprocessed in order to clean diacritics from the texts and by pattern matching antonyms are extracted and evaluated by an expert evaluator and new antonym pairs are added to the SemTree ontology. The system is evaluated using three measures which are pattern reliability, precision and performance of the system. Pattern reliability is the ratio of correct antonyms extracted using the pattern to the total extracted ones using the same pattern. System precision is the ration of the total correct extracted antonyms to the total extracted ones, while system performance is the measure of the increase in ontology size. The obtained results show that despite the fact that system performance is high (42, 3 %), system precision computed is about 29, 45 % as a mean of all obtained precisions relating to all used corpuses (KSUCCA, CCA and KACSTAC).

To summarize, we can say that firstly, all the above approaches consider only one type of relations, taxonomic or non taxonomic. Secondly, the performance of the above approaches depends on the target domain.

Our proposed approach aims to consider the two kinds of relations, taxonomic or non taxonomic, and to preserving the coherence of the enriched ontology by using the semantic similarity measure techniques offered by WordNet [7] technology, and this, independently of the domain of discourse.

4. THE PROPOSED APPROACH

We propose an approach for automatic ontology enrichment giving a corpus of texts relating to a target domain. First, a basic knowledge related to this target domain is predefined and represented in an initial or a basic ontology through a set of concepts and relationships between these concepts. The objective is to enrich the initial ontology by the content of texts relating to the same target domain through semantic analysis.

The proposed enrichment process is composed of three phases. In the first phase, we proceed to the annotation of the texts using a morpho-syntactic analysis of the given texts by means of natural language processing tools in order to provide a first level of understanding of the given texts. We parse texts to extract syntactic relations between terms as well as the part of speech tags of these terms. This annotation phase is followed by a second phase consisting of a simplification of complex sentences to simple clauses. It consists of a semantic analysis of the annotated text where clauses obeying to the form V(S, O) (S: subject, V: verb, O: object). The Third phase called semantic enrichment phase, consists of the comparison of each extracted relationship to the content of the initial ontology using a similarity measure. According to this comparison, we decide whether each extracted relationship will be candidate to enrich our initial ontology or not. The similarity measure is based on WordNet [7] and the enrichment process aims to identify new concepts or new relationships or just concept or relationship instances already existing. In this step, we study for each relation $V_t(S_t, O_t)$, extracted from the text, the semantic similarity of this last with existing ontology relations $V_o(S_o, O_o)$ to identify new concepts and relations enriching the ontology. Figure 1 gives the semantic enrichment algorithm of the basic ontology (Phase 3 of the semantic enrichment framework given by the figure 2). In figure 1, *SimilarityThreshold* is a parameter of the algorithm fixed by the user or by the domain expert.

Case 1 : Semantic_Similarity (V_t , V_o) **> =** SimilarityThreshold , **THEN**

*/ * study of the similarity between S_t & S_o, and between O_t & O_o, */*

1- **IF** S_o & S_t are locally linked in the ontology, **THEN** do nothing.
ELSE use **WordNet** technique to extract this link between S_o & S_t ;
IF link exists, then add it to the ontology, **Else,** add S_t as a concept (class) which will be the domain of V_t ;

2- **IF** S_o & S_t are locally linked in the ontology, **THEN** do nothing.
ELSE use **WordNet** to extract this link between S_o & S_t ;
IF link exists then add it to the ontology, **Else** add S_t as a class, which plays the role of V_t domain;

3- **IF O_t** & O_o are locally linked in the ontology, **THEN** do nothing.
ELSE use **WordNet** to extract this link between O_t et O_o;
IF link exists then add it to the ontology, **Else** add O_t as a class, which plays the role of V_t Codomain ;

Case 2 : Semantic_Similarity (V_t , V_o) < SimilarityThreshold **THEN**
add **V_t** to the ontology as a relation and using **WordNet** find **Link(V_t ,V_o).**
IF Link (V_t , V_o) found then add it to the ontology

*/ * study the similarity between S_t & S_o, and between O_t & O_o, */*

: 1- **IF** Semantic_Similarity (S_t , S_o) **> =** SimilarityThreshold (same appellation), **THEN** define S_o as domain of V_t .**ELSE** use **Wordnet** to find Link (S_t, S_o), **IF** Link**(S_t, S_o)** found **THEN** add it to the ontology, **ELSE** add S_t as Class in the ontology (represents the domain of V_t).

2- **IF** Semantic_Similarity(O_t, O_o) **>=** SimilarityThreshold (same appellation), **THEN** define O_o as codomain of **V_t** . **ELSE** using **Wordnet** find Link(O_t, O_o), **IF** Link(O_t, O_o) found **THEN** add it to the ontology, **ELSE** add O_t as a Class in the ontology (represents the codomain of V_t).

Figure 1: Semantic ontology enrichment

Our objective is to get an enriched ontology giving an extended semantic coverage of a target domain. We give below in Figure 2, the semantic enrichment framework of the proposed approach.

5. CONCLUSION

In this paper we have proposed an approach for ontology enrichment. It is composed of three phases. The first phase consists of the annotation of the corpus of texts relating to a given domain using natural language processing tools. The second phase allows extracting knowledge from the annotated corpus of texts in the form of basic binary relations. The third phase consists of the semantic enrichment of the basic ontology using WorldNet similarity techniques. Besides the consideration of all types of relations, our approach presents two main advantages, compared to the existing approaches. The fist advantage of our approach is that it is independent from the domain of discourse and the second one is that the enrichment process is done using semantic similarity between relations and concepts which allows preserving the coherence of the enriched ontology. Actually, we are building a basic ontology relating to Small and Medium sized Enterprises (S.M.E) domain in the aim to validate our approach and we expect to obtain promoting results.

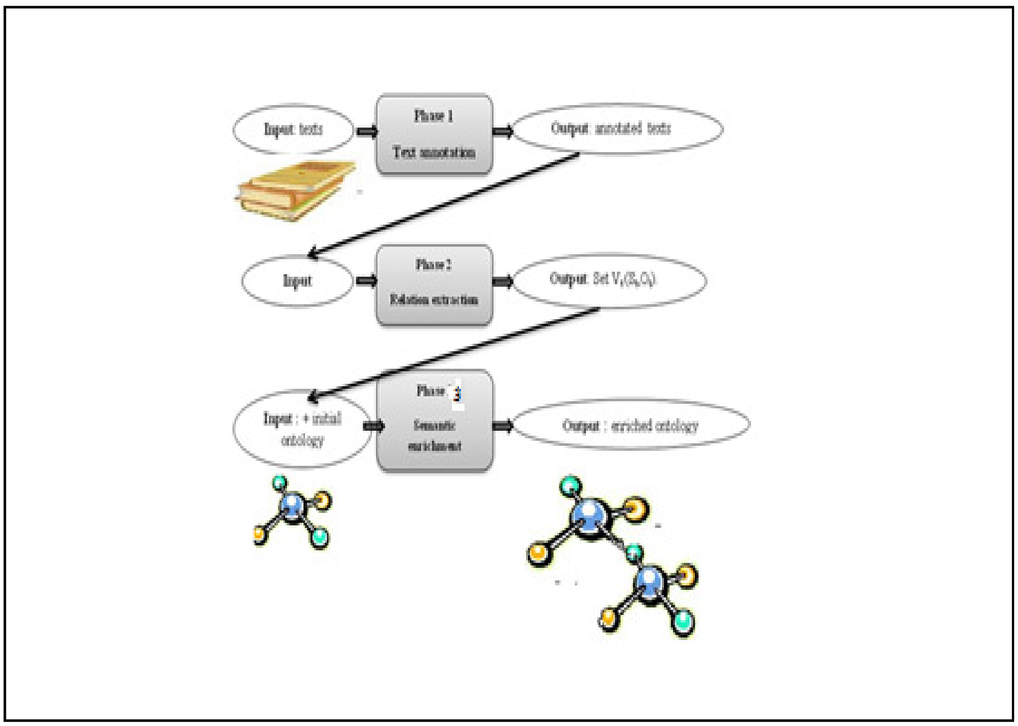


Figure 2: The semantic enrichment framework of the proposed approach

REFERENCES

[1] T. R. Gruber, "A translation approach to portable ontology specifications," Knowledge acquisition, vol. 5, no. 2, pp. 199-220, 1993.

[2] G. Petasis, V. Karkaletsis, G. Paliouras, A. Krithara, and E. Zavitsanos, "Ontology population and enrichment: State of the art, " In Knowledge-driven multimedia information extraction and ontology evolution, Springer-Verlag, January 2011, pp. 134-166

[3] C. Biemann, " Ontology learning from text: A survey of methods,". In LDV forum , Journal of Computational Linguistics and Language Technology, vol. 20, no. 2, pp. 75-93. 2005

[4] A. Gómez-Pérez, and D. Manzano-Macho,. "An overview of methods and tools for ontology learning from texts," The knowledge engineering review, vol. 19, no. 3, pp. 187-212, 2004

[5] R. Bendaoud, Y. Toussaint and A . Napoli, " Pactole: A methodology and a system for semi-automatically enriching an ontology from a collection of texts," Lecture Notes in Computer Science, vol. 5113, pp. 203-216, 2008.

[6] G. Barbur,., B. Blaga, and A. Groza, " OntoRich; A support tool for semi-automatic ontology enrichment and evaluation," In IEEE International Conference on Intelligent Computer Communication and Processing, 2011, pp. 129-132,

[7] Princeton University, "WordNet : A lexical Data Base for English" , Wordnet Princeton University , 2010. [online]. Available: http://wordnet.princeton.edu/wordnet, [Accessed : June, 10 th , 2016]

[8] A. S. Al Hashimy and N. Kulathuramaiyer, "Ontology enrichment with causation relations, " In IEEE Conference on Systems, Process & Control (TCSPC 2013), 2013, pp. 186-192.

[9] C. Faria, I, Serra, and R. Girardi, "A domain-independent process for automatic ontology population from text," Science of Computer Programming, vol. 95, pp. 26-43, 2014.

[10] M. Al-Yahya, S. Al-Malak, and L. Aldhubayi , "Ontological lexicon enrichment: The BADEA system for semi- automated extraction of antonymy relations from Arabic language corpora," Malaysian Journal of Computer Science. vol. 29, no. 1, 2016.

[11] Classical Arabic corpus (KSUCCA), [online]. Available: http://www. ksucorpus.ksu.edu.sa. [Accessed: july, 13, 2017] .

[12] Contemporary Arabic corpus (CAC) . [Online]. Available: http://www.comp.leeds. ac.uk/eric/ latifa/ research.htm . [Accessed: july, 13, 2017] .

[13] Mixed Arabic corpus (KACSTAC) . [online]. Available : http://www.kacstac.org.sa/pages / Default.aspx. [Accessed: july 13, 2017] .

[14] A. Al-Zahrani,., Al-Dalbahie, M., Al-Shaman, M., Al-Otaiby, N., and W. Al-Sultan,. SemTree: analyzing Arabic language text for semantic relations. PhD Thesis, IT Department, KSU, Saudi Arabia,, 2012.

AUTHORS

Tahar Guerram is assistant professor of Computer Science at the Department of Mathematics and Computer Science at the university of Larbi Ben Mhidi of Oum El Bouaghi and head of the same department. His areas of interest include mutli agent systems, text mining, ontologies and natural language processing.

Nassima Mellal is assistant professor of computer science at the Department of Mathematics and Computer Science at the university of Larbi Ben Mhidi of Oum El Bouaghi. Her areas of interest include semantic web and e-learning.

INTENTIONAL BLANK

A NEW METHOD OF TEACHING FIGURATIVE EXPRESSIONS TOIRANIAN LANGUAGE LEARNERS

Leila Erfaniyan Qonsuli and Mostafa Bahraman

Assistant Professor, Kashmar Higher Education Institute, Iran

ABSTRACT

In teaching languages, if we only consider direct relationship between form and meaning in language and leave psycholinguistic aside, this approach is not a successful practice and language learners won't be able to make a successful relation in the real contexts. The present study intends to answer this question: is the teaching method in which salient meaning is taught more successful than the method in which literal or figurative meaning is taught or not? To answer the research question, 30 students were selected. Every ten people are formed as a group and three such groups were formed. Twenty figurative expressions were taught to every group. Group one was taught the figurative meaning of every expression. Group two was taught the literal meaning and group three was taught the salient meaning. Then three groups were tested. After analyzing data, we concluded that there was a significant difference in mean grades between classes and the class trained under graded salience hypothesis was more successful. This shows that traditional teaching methods must be revised.

KEYWORDS

Teaching, Graded Salient Hypothesis, Figurative expressions, Language

1. INTRODUCTION

A strong foundation in teaching languages is psycholinguistics and cognitive linguistics so that language teaching and psycholinguistics are not detachable. The significance of psycholinguistic and cognitive research is best known when the result of researches show that Most of the time we speak, we speak not straight and usually we say something but we mean something else. For example we say "it's cold today" but we mean" please close the door". This has gone so far that most psycholinguists and cognitive linguists believe that even our thinking is metaphorical. Therefore, in teaching languages, if we only consider direct relationship between form and meaning in language and leave psycholinguistic aside, this approach is not a successful practice and language learners won’t be able to make a successful relation in the real contexts. According to Graded Salience Hypothesis (Giora, 1997), more salient meanings–coded meanings foremost on our mind due to conventionality, frequency, familiarity, or prototypicality–are accessed faster than and reach sufficient levels of activation before less salient ones. even rich and supportive contexts which are biased in favor of less salient meanings do not inhibit activation of salient meanings. This hypothesis predicts that in teaching languages, it’s more successful to teach the

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 21– 27, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80203

salient meaning of figurative expressions, not figurative or literal meaning. The present study intends to answer this question: is the teaching method in which salient meaning is taught more successful than the method in which literal or figurative meaning is taught or not?

2. LITERATURE REVIEW

Kecskes(1999)claims that pragmatic skills in an L2 do not necessarily reflect conceptual fluency in the target language properly because individual variables rather than conceptual fluency play a decisive role in the selection and use of these pragmatic units. In order to investigate the validity of this hypothesis a survey was conducted with 88Non-Native Speakers (NNSs) and 33 Native Speakers of English (NSs) who were given three types of tests: two discourse-completion tests, a problem-solving test and a dialog interpretation test. Data were analyzed for two variables: cultural specificity of SBUs and individual learner strategies. The findings of the survey demonstrated the existence of three developmental stages which are characterized respectively by strong L1-culturetransfer, false generalization, individual choice. Students in the third stage tend to choose SBUs on affective grounds and reject those pragmatic units which they find too culture specific.

Kecskes(2006) discusses three claims of the Graded Salience Hypothesis presented in Rachel Giora's book 'On our mind'. It is argued that these claims may give second language researchers the chance to revise the way they think about word meaning, the literal meaning–figurative meaning dichotomy and the role of context in language processing. Giora's arguments are related to recent second language research and their relevance is explained through examples. There are also several suggestions made for further research.

Kecskes(2000) considered a particular type of formulaic expressions called situation-bound utterances (SBUs). Since the meaning of these pragmatic units is shaped by the interplay of linguistic and extra linguistic factors, they can be best accounted for in a theoretical framework which represents a knowledge-for-use conception. A cognitive-pragmatic approach to SBUs reveals that in many cases cognitive mechanisms such as metaphor and conventional knowledge are responsible for the unique situational meaning of SBUs. In this respect, SBUs are similar to other formulaic expressions such as idioms and conventional implicatures whose meaning structure can also be better accounted for if the underlying cognitive mechanisms are examined. It will be claimed that the relationship of SBUs to socio-cultural concepts resembles that of words and concepts as described in Cruse (1992). SBUs will be classified according to their formula-specific pragmatic properties which are either encoded in the expression or charged by the situation. The investigation of the characteristic features of SBUs and the development of their situational meaning necessitates the review of two important theoretical issues: the creative aspect of language use and the role of formulaic expressions in the development of syntax. It will be argued that the formulaic-creative dichotomy makes sense only at sentence level, whereas it loses its significance at discourse level. Not all types of formulaic expressions contribute to syntactic development in an L2, because some of them (including SBUs) are almost never split into constituents by L2 learners. Errors in the use of SBUs can mainly be due to the lack of native-like conceptual fluency and metaphorical competence of adult L2 learners, who rely on their LI conceptual system when producing and comprehending SBUs in the target language.

Abel(2003) In two empirical studies, investigated judgments that native speakers of German make about the decomposability of English idioms . A decomposable idiom is an idiom whose individual components contribute to its figurative meaning, whereas the constituents of a non decomposable idiom do not make such a contribution. The findings were analysed and compared to native judgments. The Model of Dual Idiom Representation is introduced in order to explain the differences between the two groups. At the lexical level, the model postulates the parallel existence of idiom entries and constituent entries. The degree of decomposability and the frequency with which the idiom is encountered determine its lexical representation. If there is no idiom entry for a particular idiom, conceptual representations are accessed during comprehension. Because non native speakers encounter idioms less often than native speakers, the first language (L1) and second language (L2) lexicon vary with regard to the number of idiom entries.

In a research by Bortfeld(2003), Speakers of three different languages (English, Latvian, and Mandarin) rated sets of idioms from their language for the analyzability of the relationship between each phrase's literal and figurative meaning. For each language, subsets of idioms were selected based on these ratings. Latvian and Mandarin idioms were literally translated into English. Across three experiments, people classified idioms from the three languages according to their figurative meanings. Response times and error rates indicate that participants were able to interpret unfamiliar (e.g., other languages') idioms depending largely on the degree to which they were analyzable, and that different forms of processing were used both within and between languages depending on this analyzability. Results support arguments for a continuum of analyzability (Bortfeld & McGlone, 2001), along which figurative speech ranges from reflecting general conceptual structures to specific cultural and historical references.

Cies´licka(2013) aimed to explore the role of compositionality in the course of processing idioms by second language users. The study employed a cross-modal priming technique in which English decomposable and non decomposable idioms were embedded in sentences (e.g. 'George wanted to bury the hatchet soon after Susan left') and presented auditorily via headphones to Polish fluent speakers of English. While participants were listening to the sentence, a target word related figuratively (e.g. FORGIVE) or literally (e.g. AXE) to the idiom was presented on the computer screen for a lexical decision either at the end of the idiom or before the last word of the idiom. Contrary to the predictions of the Idiom Decomposition Hypothesis (Gibbs and Nayak 1989; Gibbs et al. 1989), figurative meanings of decomposable idioms were not available faster than those of non decomposable idioms. In addition, strong activation was found for literal meanings of idiom constituents, in line with previous L2 processing research (Kecskes 2000; Liontas 2002; Abel 2003).

3. THEORETICAL FRAMEWORK

In this part, the graded salience hypothesis (Giora, 1997) will be discussed.

3.1 Literal meaning-figurative meaning

- Literal meaning has been defined as linguistic meaning, i.e., as nonfigurative, coded, fully compositional, context-invariant, explicit, and truth conditional (Katz, 1977, in Ariel, 2002).

- Figurative meaning is seen as its counterpart, i.e., as extra linguistic, indirect, inferred, noncompositional, context-dependent, and cancelable (Ariel, 2002).

3.2 Graded Salience Hypothesis

According to the graded salience hypothesis (Giora, 1997), for information to be salient—to be foremost on one's mind—it needs to undergo consolidation, that is, to be stored or coded in the mental lexicon.

Stored information is superior to unstored information such as novel information or information inferable from context: while salient information is highly accessible, nonsalient information requires strongly supportive contextual information to become as accessible as salient information. Salience is not an either-or notion, however. Rather, it admits degrees. The more frequent, familiar, conventional, or prototypical/stereotypical the information in the mind of the individual or in a certain linguistic community, the more salient it is in that mind or among the community members (Giora, 2003:15).

3.3 Graded Salience Hypothesis Predictions

The Graded Salience Hypothesis has predictions regarding the first phase and the second phase of comprehension. It predicts that comprehension involves an initial phase in which contextually appropriate and salient meanings are activated—the latter automatically and independently of contextual information, the former as a result of a predictive context—and an immediate subsequent phase of integration in which the activated meanings are either retained for further processes or suppressed as contextually disruptive. This holds for any stage of the comprehension processes (Giora, 2003: 37).

3.4 Familiarity

The more familiar the meaning, the quicker it is to retrieve (Blasko&Connine, 1993; Gernsbacher, 1984; Hintzman& Curran, 1994). Given enough exposure and individual experience, any information can become foremost on our mind to the extent that it resists contextual information (see Zajonc, 2000); so the salience depends on the experiential familiarity an individual has with the stimulus in question.

3.5 Figurative Language

In the psychological literature, eight distinct types of nonliteral language have emerged (Kreuz& Roberts, 1993). In figurative expressions, figurative meaning is intended, not literal meaning.

4. RESEARCH METHOD

To answer the research question, 30 Iranian students of English translation field of study, aged 18-22, and female gender were selected. Every ten people are formed as a group and three such groups were formed. In a week, twenty figurative expressions were taught to every group. The materials used were 20 indirect requests or idioms taken from naturally occurring conversations; such as the following: « it is cold! »(It means: close the door!). A pretest was performed. Tennative English speakers participated. In this pretest, participants rated the expressions on a 1–

7 familiarity scale, in which 7 = highly familiar and 1 = entirely unfamiliar. Then, expressions were divided into familiarity groups (familiar, less-familiar, unfamiliar). Therefore the salient meaning was obtained. For the main test, Group one was taught the figurative meaning of every expression. Group two was taught the literal meaning and group three was taught the salient meaning. Then three groups were tested and the meaning of expressions was asked.

5. DATA ANALYSIS

Chi-Square Testwas used for the pretest. Then, expressions were divided into familiarity groups (familiar, less-familiar, unfamiliar). Foranalyzing the results of main test, t-Test was used and sig =0.022<0.05. This indicated thatthere was a significant difference in mean grades between classes.

Table 1.Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
v3	12.3000	20	3.15561	.70562
v4	14.9000	20	2.66853	.59670

Table 2.Samples Correlations

	N	Correlation	Sig.
v3 & v4	20	-.271	.247

Table 3.Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
v3 - v4	-2.60000	4.65267	1.04037	-4.77752	-.42248	-2.499	19	.022

6. DISCUSSION AND CONCLUSION

After analyzing data, we concluded that there was a significant difference in mean grades between classes and the class trained under graded salience hypothesis was more successful. This shows that traditional teaching methods must be revised.

Graded salience hypothesis provides new insights supported by empirical evidence that prompt the reader to revise his or her views about L2 language processing, vocabulary acquisition, pragmatics and reading. It has much to offer to applied linguists. This Hypothesis has already begun and has produced some interesting results (Abel, 2003;Bortfeld, 2002; 2003; Cies´licka, forthcoming). The investigation of salience in second language acquisition should allow us to understand how emerging new conceptual knowledge blends with existing L1-based conceptual knowledge, resulting in the development of a new complex language system and modifications in the operation of the L1. The GSH may help us develop an intake theory that focuses not only on input and output as the inter language approach does, but that attempts to explain what happens in the mind of language learners as they process new input(Kecskes,2006).

REFERENCES

[1] Abel, Beate, (2003), English idioms in the first language and second language lexicon: a dual representation approach, Second Language Research 19,4 ; pp. 329– 358.

[2] Ariel, M. (2002).the Demise of a Unique Literal Meaning. Journal of Pragmatics, 34, 361-402.

[3] Blasko, G. D. &Connine, C. (1993).Effects of Familiarity and Aptness on Metaphor Processing. Journal of Experimental Psychology: Learning, Memory, and Cognition,19, 295–308.

[4] Bortfeld, H., &McGlone, M. (2001). The continuum of metaphor processing.Metaphor and Symbol, 16,75-86.

[5] Bortfeld, H.,(2002), What native and non-native speakers' images for idioms tell us about figurative language. In Heredia, R. and Altarriba, J., editors, Bilingual Sentence Processing.North Holland/Elsevier, 275–95.

[6] Bortfeld, Heather, (2003), Comprehending Idioms CrossLinguistically, Experimental Psychology; Vol. 50(3): 217-230.

[7] Cies´licka, Anna B. , (2013), Second Language Learners' Processing of Idiomatic Expressions: Does Compositionality Matter? Second Language Learning and Teaching, Series Editor MirosławPawlak, KrystynaDroz´dział-Szelest. , Studies in Honor of WaldemarMarton,springer, 115-136.

[8] Cruse, D.A., (1992), Antonymy revisited: Some thoughts on the relationship between words and concepts. In: A. Lehrer and E. Kittay, eds., Frames, fields and contrasts, Hillsdale, 289-309: Erlbaum.

[9] Gernsbacher, M. A. (1984). Resolving Twenty years of Inconsistent Interactions between Lexical Familiarity and Orthography, Concreteness, and Polysemy. Journal of Experimental Psychology: General, 113, 256–281.

[10] Gibbs, R. W., N. P. Nayak and C. Cutting. (1989), How to kick the bucket and not decompose: Analyzability and idiom processing. Journal of Memory and Language28: 576-593.

[11] Gibbs, R. W. and N. P. Nayak.(1989), Psycholinguistic studies on the syntactic behavior of idioms. Cognitive Psychology21: 100-138.

[12] Giora, R. (1997). Understanding Figurative and Literal Language: the Graded Salience Hypothesis. Cognitive Linguistics, 8(3), 183-206.

[13] Giora, R. (2003). On Our Mind: Salience, Context, and Figurative Language. Oxford University Press.

[14] Hintzman, D. L. & Curran, T. (1994). Retrieval Dynamics of Recognition and Frequency Judgements: Evidence for Separate Processes of Familiarity and Recall. Journal of Memory and Language, 33, 1–18.

[15] Katz, A. N. (1977). Propositional Structure and Illocutionary Force.NewYork: Thomas Y.Crowell.

[16] Kecskes, Istvan, (1999), Conceptual fluency and the use of situation-bound utterances in L2, Links & Letters 7, 145-161.

[17] Kecskes, Istvan, (2000),A cognitive-pragmatic approach to situation-bound utterances, Journal of Pragmatics 32 , 605-625.

[18] Kecskes, Istvan, (2006), On my mind: thoughts about salience, context and figurative language from a second language perspective, Second Language Research 22,2; pp. 1–19.

[19] Kreuz, R, J. & Roberts, R. M. (1993).the Empirical Study of Figurative Language in Literature. Poetics, 22, 151-169.

[20] Liontas, J. 2002. Context and idiom understanding in second languages.In EUROSLA Yearbook, eds. S. H. Foster-Cohen, T. Ruthenberg and M. L. Poschen, 155-185. Amsterdam: John Benjamins Publishing Company.

[21] Zajonc, R. B. (2000). Closing the Debate Over the Independence of Affect. In J. P. Forgas(Ed.), Feeling and thinking: the Role of Affect in Social Cognition(31–58). Cambridge: Cambridge University Press.

AUTHOR

Leila ErfaniyanQonsuli

EDUCATION

- B.A.: English Language and Literature(University of Sistan and Baluchestan, Zahedan, Iran)
- M.A.: General Linguistics(Ferdowsi University of Mashhad, Mashhad, Iran)
- Ph.D: General Linguistics(Ferdowsi University of Mashhad, Mashhad, Iran)

HONORS AND DISTINCTIONS

- Earn the rank of "A" (the first person in Iranian`s Ph.D. entrance exam)
- Earn the rank of "A" among Ph.D. students in Ferdowsi university of Mashhad (Top graduate student with the Top GPA. of A+).
- Earn the rank of "A" (the first person in Iranian`s MA. entrance exam)
- Earn the rank of "A" among MA. Students in Ferdowsi university of Mashhad (Top graduate student with the Top GPA. of A+).
- A member of Iranian elite center.
- Earn the third rank of the best Ph.D. thesis in Iran.
- Earn the excellent rank for MA. Thesis (GPA: A+).
- Earn the excellent rank for Ph.D. Thesis (GPA: A+).
- Earn the "Best professor" title from scientific and applied university.
- Earn the "Best married student" title from Ferdowsi university of Mashhad.
- Earn the "Best student" title from Ferdowsi university of Mashhad multiple times.
- A member of "applied psycholinguistics" journal reviewers.
- A member of "Khorasan`s dialects and linguistics" of Ferdowsi university journal reviewers.
- A member of "Literary and rhetorical studies" journal of PayameNour university.
- Supervisor and advisor and professor in MA. theses.
- Author of multiple published paper s in scientific journals

INTENTIONAL BLANK

SEMANTIC STUDIES OF A SYNCHRONOUS APPROACH TO ACTIVITY RECOGNITION

Ines SARRAY[1], Annie RESSOUCHE[1], Sabine MOISAN[1], Jean-Paul RIGAULT[1] and Daniel GAFFE[2]

[1]Universite´Co^ted'Azur, INRIA, SophiaAntipolis, France
[2]Universite´Co^ted'Azur, CNRS, LEAT, SophiaAntipolis, France

ABSTRACT

Many important and critical applications such as surveillance or healthcare require some form of (human) activity recognition. Activities are usually represented by a series of actions driven and triggered by events. Recognition systems have to be real time, reactive, correct, complete, and dependable. These stringent requirements justify the use of formal methods to describe, analyze, verify, and generate effective recognition systems. Due to the large number of possible application domains, the researchers aim at building a generic recognition system. They choose the synchronous approach because it has a well-founded semantics and it ensures determinism and safe parallel composition. They propose a new language to represent activities as synchronous automata and they supply it with two complementary formal semantics. First a behavioral semantics gives a reference definition of program behavior using rewriting rules. Second, an equational semantics describes the behavior in a constructive way and can be directly implemented. This paper focuses on the description of these two semantics and their relation.

KEYWORDS

Activity Recognition, Language, Synchronous Approach, Semantics

1. INTRODUCTION

Activity Recognition aims at recognizing sequences of human actions that follow the predefined model of an activity. Our research team mainly works on medical applications to help physicians detect abnormal behaviors or monitor patient activities such as serious games.

Recognition systems must satisfy stringent requirements: dependability, real time, cost effectiveness, security and safety, correctness, completeness... To enforce most of these properties, the chosen approach is to base the configuration of the system as well as its execution upon formal techniques. Therefore, these formal bases should permit static analysis, verification and validation, but also easy and direct implementation.

The aim is to build a generic recognition system for such activities. The authors chose to model these activities as synchronous finite automata. The synchronous paradigm ensures determinism and supports concurrency through parallel composition. In particular, critical races can be

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 29– 45, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80204

detected by static analysis. This model is also well-founded owing to formal semantics. Thus the recognition system benefits from the sound foundations of the synchronous approach and from the automata theory, allowing automatic proofs, static verification, powerful simulation, code generation, etc.

For the users to describe synchronous automata, languages such as Lustre, Esterel, Scade, and Signal [1] have been defined. These languages are for expert users. This paper proposes another language called ADeL (Activity Description Language). Building a complete generic recognition system involves many different aspects. The paper concentrates on the formal description of activities as synchronous automata and their mathematical semantics. However, the flavor of all these different aspects will begiven.

The paper is organized as follows. The next section is a short reminder of the synchronous model of reactive systems. An overview of the ADeL language is given in section 3. Section 4 is the core of the paper: it introduces the semantics and the mathematical concepts on which researchers rely to define and verify the behavior of programs and to compile them. Finally several related works are presented before concluding.

2. SYNCHRONOUS MODEL OF REACTIVE SYSTEMS

The Synchronous Paradigm relies on a discrete logical time composed of a sequence of logical instants, defined by the system reactions.

Reactive systems listen to input events coming from the external environment and react to them by generating output events towards the environment. Such systems can be complex. The synchronous model is a way to reduce the complexity of behavior description by considering their evolution along successive discrete instants. An instant starts when some input events are available. The output and internal events deriving from these inputs are computed until stability (fixed point) is achieved; the instant finishes by delivering the output events to the environment.

No inputs occurring “during” the instant are considered. Hence, instants are atomic, their sequence defines a logical time. In this model, instants take “no time” with respect to the logical time they define.

The synchronous paradigm is interesting because it ensures determinism and it supports concurrency through deterministic parallel composition. In particular, critical races are detected by static analysis. It supports a true notion of simultaneous events and provides not only a reaction to the presence of an event but also to its absence (to some extent). This model is also well-founded owing to formal semantics. Moreover, along the last decades, tool sets for simulation, verification, and code generation of synchronous automata have been developed.

The synchronous model has been applied to several different systems, from hardware design [2] to embedded real time systems [3]. In this work, the team proposes to apply it to another real time system, namely human activity recognition. Synchronous models can be represented as *Mealy machines*. The Mealy machines that they consider are 6-uples of the form: $<Q, q_{init}, I, O, \lambda, \tau>$, where Q is a finite set of states, $q_{init} \in Q$ is the initial state, I (resp. O)is a finite set of input (resp. output) events; λ: $(Q \times I) \rightarrow Q$ is the transition function and $\tau : (Q \times I) \rightarrow O$ is the output function. This is an explicit representation of Mealy machines as automata. Mealy himself introduced another representation as Boolean equation systems that calculate both the output event values and the next state from the input event values and the current state [4]. The authors call this representation “implicit” Mealy machines.

Synchronous languages such as Lustre, Esterel, Scade, and Signal[1] have been defined to describe synchronous automata. These languages are for expert users. This paper proposes another synchronous language that is easier to understand and to work with for non-computer scientists (e.g., doctors). To improve its acceptance and its ease of use by non-computer scientists, the authors are working in collaboration with ergonomists and doctors from Claude Pompidou hospital. This language is called ADeL (Activity Description Language) and is described in the next section.

3. ACTIVITY DESCRIPTION LANGUAGE (ADEL)

ADeL provides two different (and equivalent) formats: graphical and textual. It is a modular and hierarchical language, which means that an activity may contain one or more sub-activities. The description of an activity consists of several parts: first the user defines the participants in the activity, their types, their roles, as well as the initial state of the activity. Second, the user describes the expected behavior using a set of control operators detailed in table 1. These operators are the base of the ADeL language. They have a synchronous semantics and they deal with events coming from their environment.

Table1.ADeL operators. S, S_1 are events (received or emitted);p, p_1 and p_2 are instructions; condition is either an event or a Boolean combination of event presence/absence

nothing	does nothing and terminates instantaneously.
[**wait**] S	waits for event *S* and suspends the execution of the activity until *S* is present. Operator wait can be implicit or explicit.
p_1 **then** p_2	starts when p_1 starts; p_2 starts when p_1 ends; the sequence terminates when p_2 does.
p_1 **parallel** p_2	starts when p_1 or p_2 start; ends when both have terminated.
p_1 **during** p_2	p_1 starts only after p_2 start and must finish before $p_2$end.
while *condition* {*p*}	*p* is executed only if the *condition* is verified. When *p* ends, the loop restarts until the *condition* holds.
stop {*p*} **when** *S* [**alert** S_1]	executes *p* to termination as long as *S* is absent, otherwise when *S* is present, aborts *p*, sends an alert S_1, and terminates
if *condition* **then** p_1[**else** p_2]	executes p_1 if *condition* holds, otherwise executes p_2.

p **timeout** *S* {p_1}[**alert** S_1]	executes *p* ; stops if *S* occurs before *p* terminates and possibly sends alert S_1; otherwise executes $p_1$when *p* has terminated.
alert *S*	raises an alert.
Local (*events*){*p*}	declares internal events to communicate between sub parts of *p*.
Call (*activity*)	calls a sub-activity.

Some of these operators are "instantaneous" (**nothing, alert**) while others take at least one (synchronous) instant to process.

Compared to other synchronous languages where it is difficult or even impossible to treat the real clock time, ADeL can manipulate it thanks to the operator **"timeout"**. For example, deadlines are expressed as follows: P **timeout** *S* {P_1} (*S* is a timed signal). To compare with a classical approach, in Esterel, this operator should be written as:

```
abort{p} when S;
presentS then alert else P1;
```

This part of Esterel code seems easy for a programmer but it is not the case for non-computer scientists such as doctors. Indeed, it would be even more difficult to write this kind of code in a declarative synchronous language like Lustre. Moreover, it is more complex to use these languages to express the "**during**" operator.

The main issue of the synchronous paradigm is that the world is not synchronous in general. Thus it requires to transform asynchronous physical flows of events into a succession of discrete instants. The authors propose a synchronous transformer, called Synchronizer. The Synchronizer receives asynchronous events from the environment, filters them, decides which ones may be considered as "simultaneous", and groups them into a logical instant according to predefined policies. In general, no exact simultaneity decision algorithm exists but several empirical strategies may be used for determining instant boundaries, relying on event frequency, event occurrence, elapsed time, etc. To manage the real clock time, the Synchronizer considers the clock time as an event like others.

4. ADEL SEMANTICS AND COMPILATION

To provide the language with sound foundations, the authors turn to a formal semantic approach. First, logical rewriting rules are a classical and rather natural way to formally express the intuitive semantics. This form of behavioral semantics gives an abstract description of a program behavior and facilitates its analysis. However, it is not convenient as an implementation basis nor suitable for proofs (e.g., model-checking). Hence, an equational semantics, which maps an ADeL program to a Boolean equation system representing its finite state machine, was also defined. The ADeL compiler can easily translates this equation system into an efficient code. Using such a double semantics is somewhat traditional in the synchronous language area[5].

Since there are two different semantics, it is mandatory to establish their relationship. In fact the authors proved that the execution of a program based on the equational semantics also conforms to the behavioral semantics (see4.4).

4.1 Mathematical Context

One of the bases of ADeL semantics is the notion of an *environment*, which is a finite set of events. Environments record the status of events in each synchronous instant and the goal of the ADeL semantics is to compute the status of the output events for each reaction of a program. A 4-valued algebra ($\xi = \{\perp, 0, 1, \top\}$) is used to represent the status: $\perp$ means that the status of the event is not yet determined, 0 that the event is absent, 1 that the event is present, and $\top$ that the status of the event is over determined (error). Usually $\top$ Occurs because the event would have two incompatible status in the same instant (e.g., 0 and 1 in different parts of the program).

Previously, synchronous language semantics expressed the status of events using 3-valued algebras. Indeed, such semantics either cannot reflect how the information about event status grows when the algebra is a lattice ($0 \leq \perp \leq 1$), or fix the status of each event to 0 or 1 in each instant when the algebra is a complete partial order (c.p.o) ($\perp \leq 0, \perp \leq 1$)[6]. In this latter case, at the beginning of the computation of event status, all unknown status are set to $\perp$. Then, as soon as an event is present in a part of the program, its status grows to 1, otherwise it is set to 0. This way prevents any incremental compilation of activities. To compile a main activity without the knowledge of event status of a sub activity, these latter must be kept to $\perp$.An event never present in a program has 0 for status and if it is present in a sub program, its status cannot be changed to 1 when the already compiled sub program is included in the main one, because 1 and 0 are incomparable with respect to the c.p.o order and have no upper bound. It is why the authors consider a 4-valued algebra and a structure which allows us to rely on the semantics rules to compile programs in an incremental way. Moreover, this algebra provides us with a convenient means to compile activities.

As a consequence, the authors supply ξ with a bilattice structure [7]. Bilattices are mathematical structures having two distinct orders denoted $\leq_B$ (Boolean order) and $\leq_K$ (knowledge order) and a $\neg$ operation, such that both (ξ, $\leq_B$) and (ξ, $\leq_K$) are lattices for their respective orders. In ξ, $\leq_B$ represents an extension of the usual Boolean order and $\leq_K$ expresses the level of information about the presence of an event. These two orders are defined as follows: $\perp \leq_K 0 \leq_K \top$; $\perp \leq_K 1 \leq_K \top$; $0 \leq_B \perp \leq_B 1$; $0 \leq_B \top \leq_B 1$. These two orders play complementary roles in the ADeL semantics: the Boolean order is a means to calculate the event status while the knowledge order reflects the growth of information about event status when computing the equation system. As a consequence, four operations in ξ have been introduced: $\boxdot$ and $\boxplus$ are respectively the "meet" and "join" operations of (ξ, $\leq_B$) and $\sqcup$ and $\sqcap$ play the same roles for (ξ, $\leq_K$).

Finally, the $\neg$ operator is used to reverse the notion of truth from a Boolean point of view, but its role with respect to $\leq_K$ has to be transparent: actually, no more nor less information about x and $\neg x$ is known, then the authors give the following definitions for this $\neg$ operator: $\neg 1 = 0$, $\neg 0 = 1$, $\neg \top = \top$, and $\neg \perp = \perp$.

Another important feature of our approach is the ability to encode ξ elements into pairs of Boolean ones. There exist several possible encoding functions and the researchers choose one which is compatible with the $\leq_K$ order:

$$e:\xi \mapsto \mathbb{B}\times\mathbb{B} : x \in \xi,\ e(x) = (x_h, x_l) : \begin{cases} \bot & \mapsto & (ff, ff) \\ 0 & \mapsto & (ff, tt) \\ 1 & \mapsto & (tt, ff) \\ \top & \mapsto & (tt, tt) \end{cases}$$

Here $\mathbb{B}$ is the usual Boolean algebra $\{ff,\ tt\}$.

This encoding function extends to the ξ operators. The structure ($\mathbb{B}$, ≤) is a complete lattice for the $ff \leq\ tt$ order. Then, the structure: $\mathbb{B} \bigodot \mathbb{B} = (\mathbb{B}\times\mathbb{B}, \leq_B, \leq_K, \neg)$ defined as follows:

$$\begin{cases} (x_1, x_2) \leq_B (y_1, y_2) & \text{iff} \quad x_1 \leq y_1 \ and\ y_2 \leq x_2 \\ (x_1, x_2) \leq_K (y_1, y_2) & \text{iff} \quad x_1 \leq y_1 \ and\ x_2 \leq y_2 \\ \neg(x_1, x_2) = (x_2, x_1) & \end{cases}$$

is a bilattice and the following theorem holds:

Theorem 1 *(ξ, $\leq_B$, $\leq_K$, ¬) and* $\mathbb{B} \bigodot \mathbb{B}$ *are isomorphic.*

To justify this theorem, the authors show that the encoding *e* previously defined is an isomorphism between (ξ, $\leq_B$, $\leq_K$, ¬) and $\mathbb{B} \bigodot \mathbb{B}$ Indeed, the four binary operations and the negation one of the (ξ, $\leq_B$, $\leq_K$, ¬) bilattice are preserved in $\mathbb{B} \bigodot \mathbb{B}$ The proof is detailed in [8].

As a result of the theorem, the encoding *e* previously defined for ξ elements can be extended to the operators of the bilattice (ξ, $\leq_B$, $\leq_K$, ¬)[1]:

$$\begin{bmatrix} e(x \sqcup y) & = & (x_h + y_h, x_l + y_l) \\ e(x \sqcap y) & = & (x_h . y_h, x_l . y_l) \\ e(x \boxplus y) & = & (x_h + y_h, x_l . y_l) \\ e(x \boxdot y) & = & (x_h . y_h, x_l + y_l) \end{bmatrix}$$

Thus, one can efficiently convert ξ-equation systems into the Boolean universe.

4.1.0.1 Extension to Environments

Owing to the ξ algebra, it is now possible to formally introduce the notion of *environments*. Environments are finite sets of events where each event has a single status.

---

[1]In the following equations, + and .denote the join and meet operations of the lattice (B, ≤))

More formally, consider a finite set of events S = $\{S_0, S_1, \ldots S_n, \ldots\}$. A valuation $\mathcal{V} : \mathcal{S} \mapsto \xi$ is a function that maps an event $S \in \mathcal{S}$ to a status value in ξ. Each valuation $\mathcal{V}$ defines an environment : $E = \{S^x \mid S \in \mathcal{S}, x \in \xi, \mathcal{V}(S) = x\}$. The goal of the semantics is to refine the status of the events of a program in each instant from ⊥ to T according to the knowledge order ($\leq_K$).

Then, for each instruction *p*, built with ADeL operators, let us denote $\mathcal{S}(p)$ the finite set of its events and $\mathcal{E}(p)$ the set of all possible environments built from $\mathcal{S}(p)$.Operations in (ξ, $\leq_B$, $\leq_K$,¬) can be extended to environments[2]:

$$\begin{aligned} \neg E &= \{S^x | S^{\neg x} \in E\} \\ E \sqcup E' &= \{S^z | \exists S^x \in E, \exists S^y \in E', z = x \sqcup y\} \\ &\cup \{S^x | S^x \in E, \nexists y \in \xi, S^y \in E'\} \\ &\cup \{S^y | S^y \in E', \nexists x \in \xi, S^x \in E\} \end{aligned}$$

The order relation $(\preceq)$ on environments is defined as follows:

$$E \preceq E' \text{ iff } \forall S^x \in E, \exists S^y \in E' | S^x \leq_K S^y$$

Thus $E \preceq E'$ means that each element of E is less than an element of E' according to the lattice knowledge order of ξ. As a consequence, the $\preceq$ relation is a total order on $\mathcal{E}(p)$ and $\sqcup$ and $\sqcap$ operations are monotonic according to $\preceq$. Moreover, ($\mathcal{E}(p)$, $\preceq$) is a complete lattice, its greatest element is $\{S^\top \mid S \in \mathcal{S}(p)\}$ and its least element is $\{S^\perp \mid S \in \mathcal{S}(p)\}$. According to Tarski's theorem, each monotonic increasing function F has a least fixed point, computed by iteration of F from the least element [9]. This ensures that the behavioral semantics has solutions.

4.2 Behavioral Semantics

Behavioral semantics is a classical and formal way to describe behaviors in an axiomatic way. This semantics formalizes each reaction of a program by computing the output environment from the input one. To this aim, it defines a set of rewriting rules of the form:

$$p \xrightarrow[E]{E', term} p'$$

where *p* and *p'* are two instructions of ADeL, p' is the derivative of *p*, i.e. the new instruction that will react to the next input environment. *E* is the input environment, E' is the resulting output environment, and *term* is a Boolean flag which describes the termination of *p*, and which turns to true when *p* terminates. The rewriting rules of the whole program apply from the root instruction, structurally following the syntactic tree of the program.

[2]Only the operations needed to define both semantics are introduced. However, the five operators of ξ can be similarly extended.

Due to lack of space, the behavioral semantics of all the operators cannot be described. Only the rules for two operators are presented: **parallel** which is specific to synchronous languages and **timeout** that takes into consideration the synchronous time. Nevertheless, a complete description is detailed in[10].

4.2.0.1 Operator parallel.

Operator **parallel** has two argument instructions that are executed and computed concurrently, possibly broadcasting events between them. Thus the evolution of both instructions can have an impact on both environments. The operator ends when the two instructions terminate, i.e. when $term_{p}1$ and $term_{p}2$ become true, and the resulting output environment is the unification of the respective resulting environments computed for p_1 and p_2.

$$\frac{p_1 \xrightarrow[E \sqcup E_2]{E_1, term_{p_1}} p'_1 \quad , \quad p_2 \xrightarrow[E \sqcup E_1]{E_2, term_{p_2}} p'_2}{p_1 \| p_2 \xrightarrow[E]{E_1 \sqcup E_2,\ term_{p_1}\ and\ term_{p_2}} p'_1 \| p'_2} \tag{1}$$

4.2.0.2 Operator timeout.

The behavior of: p **timeout** $S\{p_1\}$ **alert** S_1 depends both on the behavior of its instruction p and on the status of S. If S is not present and p terminates, p_1 starts and the behavior of the operator turns out to be the behavior of p_1 (rule2).

$$\frac{p \xrightarrow[E]{E_1,1} \textbf{nothing}, S \notin E,\ p_1 \xrightarrow[E1]{E_2, term_{p_1}} p'_1}{\textbf{timeout}(p,\ S\ \{p_1\}) \xrightarrow[E]{E_2, term_{p_1}} p'_1} \tag{2}$$

If S is present (i.e., timeout elapsed), the computation of the operator stops the execution of p and finishes by generating **nothing** as final result and changing $term_p$ to true. The final environment is the output environment E, where the status of the event S_1 becomes true (rule3).

$$\frac{p \xrightarrow[E]{E_1, term_p} p' \quad , \quad S \in E}{\textbf{timeout}(p,\ S,\ p_1) \xrightarrow[E]{E \cup \{S_1^1\}, 1} \textbf{nothing}} \tag{3}$$

The behavioral semantics is a "macro" step semantics that gives the meaning of a reaction for each ADeL instruction. Nevertheless, a reaction is the least fixed point of a "micro" step semantics [6] that computes the output environment from the input one. As mentioned in subsection 2, for each instruction p, each monotonic increasing function from $\mathcal{E}(p)$ to $\mathcal{E}(p)$has a least fixed point which defines the semantics of the program of which p is the root instruction.

More precisely, $p \xrightarrow[E]{E', term} p'$ represents a sequence of micro steps such that:

$$p \xrightarrow[E]{E_1, term_1} p_1, p_1 \xrightarrow[E_1]{E_2, term_2} p_2, \ldots, p_n \xrightarrow[E_n]{E_{n+1}, term_{n+1}} p'$$

and where, at each step, $E_{i+1} = F(E_i)$ (F represents the application of one of the semantic rules to calculate the output environment from the input one). Since the F functions rely on the $\sqcup$ operator on environments, they are monotonic and increasing with respect to the $\preceq$ order. Then $\forall i,\ E_{i+1} \preceq F(E_i)$ and E' is the least fixed point of the F^n function application.

The behavioral semantics is a logical one based on rewriting rules. However, it cannot be really usable to build compilers because it requires the non-trivial computation of fixed points. Nevertheless, this semantics is the reference for the ADeL language and any other semantics must conform to it.

To get an efficient means to compile programs, the authors introduced another semantics based on *constructive* Boolean logic. Hence, this second semantics is also constructive: one can deduce the status of events by propagating the status of input events instead of computing fixed points[6].

4.3 Equational Semantics

Equational semantics allows us to make an incremental compilation of the ADeL programs by translating each root instruction of programs into a ξ-equation system. An equation system is definedasthe4-tuple $<I,O,R,D>$ where I are the input events, O are the output events, $R$are the registers, i.e specific variables acting as memories to record values useful to compute the next instant, and D is the definition of the equation system to calculate the status of each event.

The equational semantics computes the equation system of a program instruction. The authors defines it first for the operators of ADeL and then they extend these definitions to programs. The equational semantics is a function $\mathcal{S}_e$ which calculates an output environment from an input one. Let p be an ADeL instruction and E an input environment. Let us denote $\mathcal{D}(p)$ its equation system and $\langle p \rangle_E$ the resulting output environment, computed by $\mathcal{S}_e$ which is expressed as follows: $\mathcal{S}_e(p, E) = \langle p \rangle_E$ iff $E \vdash \mathcal{D}(p) \hookrightarrow \langle p \rangle_E$. From the event valuation of E, the equation system $\mathcal{D}(p)$ gives the event valuation of $\langle p \rangle_E$. Thanks to theorem 1, ξ-equation systems can be represented as Boolean ones to calculate output events value, and turn back into the 4-valued world to build the resulting environment. So, $\hookrightarrow$ means that $\langle p \rangle_E$ is deduced by applying well known Boolean algebra properties on Boolean equation systems. Then, for an ADeL program and a global input environment E (i.e an environment where output and local events have $\perp$ as status and the registers have 0 as value), the global output environment E' computed by the equational semantics is $\mathcal{S}_e(p, E)$, p being the root instruction of the program. Thus, the equation system of an instruction is deduced from semantic rules expressed for each operators of the language. To define these rules, three specific events are defined for each operator: START to start the instruction, KILL to kill the instruction, and FINISH to send the termination information to the enclosing instruction.

The operator equation systems are defined by operator semantic rules to compute the status of the FINISH, output, and local events, according to the status of START, KILL, input and local events.

As example, here follows the equational semantics of the two operators already considered in section 4.2.

4.3.0.1 Operator parallel

Operator **parallel** unifies (operation $\sqcup$) the output environments of its two operands. The output environment is computed according to the following rule:

$$\langle p_1 \rangle_E \sqcup \langle p_2 \rangle_E \vdash \mathcal{D}_{p_1 \| p_2} \hookrightarrow \langle p_1 \| p_2 \rangle_E.$$

The rule to define $\mathcal{D}_{p_1 \| p_2}$ (see Fig. 1) introduces two registers R_1 and $R_2$to memorize the respective statuses of the FINISH events of the two parallel arguments, since this operator ends when both of its two operands have finished their execution [3]. Note that the operands do not in general terminate in the same instant.

4.3.0.2 Operator timeout.

The output environment of p **timeout** $S\{p_1\}$ **alert** S_1 is calculated as follows:

$$\langle p \rangle_E \vdash \mathcal{D}_{timeout(p,S,p_1,S_1)} \hookrightarrow \langle timeout(p, S, p_1, S_1) \rangle_E.$$

The $\mathcal{D}_{timeout(p,S,p_1,S_1)}$ equation system (see Fig. 2) contains also two registers to record the way this instruction terminates: either with the normal termination of its argument (p) or when the timeout event becomes true. To express the rule for timeout operator, the same rules to denote events as in the previous operator are used.

[3]In operator parallel equation system, the specific signals of the considered operator(here parallel) are denotedSTART, KILL and FINISH while the specific signals of the arguments $p1$ and $p2$ are indexed with the argument respective name.

$$\mathcal{D}_{p1\|p2} = \left[\begin{array}{lcll}
R_1^+ & = & R_1 \boxdot \neg\text{FINISH}_{p_2} \boxdot \neg\text{KILL} \boxplus \neg R_2 \boxdot & \\
 & & \text{FINISH}_{p_1} \boxdot \neg\text{FINISH}_{p_2} \boxdot \neg\text{KILL} & (1) \\
R_2^+ & = & R_2 \boxdot \neg\text{FINISH}_{p_1} \boxdot \neg\text{KILL} \boxplus \neg R_1 \boxdot & \\
 & & \neg\text{FINISH}_{p_1} \boxdot \text{FINISH}_{p_2} \boxdot \neg\text{KILL} & (2) \\
\text{START}_{p_1} & = & \text{START} & (3) \\
\text{START}_{p_2} & = & \text{START} & (4) \\
\text{KILL}_{p_1} & = & \text{KILL} & (5) \\
\text{KILL}_{p_2} & = & \text{KILL} & (6) \\
\text{FINISH} & = & R_1 \boxdot \neg R_2 \boxdot \text{FINISH}_{p_2} \boxplus R_2 \boxdot \neg R_1 \boxdot \text{FINISH}_{p_1} \boxplus & \\
 & & \neg R_1 \boxdot \neg R_2 \boxdot \text{FINISH}_{p_1} \boxdot \text{FINISH}_{p_2} & (7)
\end{array}\right]$$

Figure 1.Equational semantics of parallel operator.

$$\mathcal{D}_{timeout(p,S,p_1,S_1)} = \left[\begin{array}{lcll}
R_1^+ & = & R_1 \boxdot \neg S \boxdot \neg\text{FINISH}_p \boxdot \neg\text{KILL} \boxplus & \\
 & & \neg R_1 \boxdot \neg R_2 \boxdot \text{START} \boxdot \neg\text{KILL} & (1) \\
R_2^+ & = & R_1 \boxdot R_2 \boxdot S \boxdot \neg\text{FINISH}_p \boxdot \neg\text{KILL} \boxplus & \\
 & & R_1 \boxdot R_2 \boxdot \neg S \boxdot \text{FINISH}_p \boxdot \neg\text{KILL} \boxplus & \\
 & & R_1 \boxdot R_2 \boxdot \neg S \boxdot \neg\text{FINISH}_p \boxdot \neg\text{FINISH}_{p_1} \boxdot \neg\text{KILL} \boxplus & \\
 & & R_1 \boxdot \neg R_2 \boxdot \neg S \boxdot \text{FINISH}_p \boxdot \neg\text{KILL} \boxplus & \\
 & & \neg R_1 \boxdot R_2 \boxdot \neg\text{FINISH}_{p_1} & (2) \\
\text{START}_p & = & \neg R_1 \boxdot \neg R_2 \boxdot \text{START} & (3) \\
\text{START}_{p_1} & = & R_1 \boxdot \neg R_2 \boxdot \neg S \boxdot \text{FINISH}_p & (4) \\
\text{KILL}_p & = & \text{KILL} & (5) \\
\text{KILL}_{p_1} & = & S \boxdot \neg\text{KILL} \boxdot R_1 \boxplus \text{KILL} & (6) \\
\text{FINISH} & = & R_1 \boxdot \neg R_2 \boxdot S \boxplus \neg R_1 \boxdot R_2 \boxdot \text{FINISH}_{p_1} & (7) \\
S_1 & = & R_1 \boxdot \neg R_2 \boxdot S \boxdot \neg\text{FINISH}_p & (8)
\end{array}\right]$$

Figure 2.Equational semantics of timeout operator.

4.4 Relation between Behavioral and Equational Semantics

The behavioral semantics gives a meaning to each program: for each ADeL operator, it formally defines the computation of the output environment and of a Boolean termination flag. The equational semantics, by associating a ξ-equation system to each operator, provides a constructive way to perform the computation. It is important to establish the relation between the solutions obtained by both semantics. To this aim, the following theorem has been proved:

Theorem 2 *Let p be an ADeL instruction, O a set of output events and E an input environment. If* $\langle p \rangle_E$ *is the resulting environment computed by the equational semantics, then the following property holds:* $\exists p'$ *such that* $p \xrightarrow[E]{E',FINISH_P} p'$ *and* $\forall o \in O$, *o has the same status in* $\langle p \rangle_E$ *and* E'.

In short, the theorem means that if the equational semantics yields a solution, there exists also a behavioral solution with the same outputs. It is a proof by induction on the size of a program where the size of an instruction is roughly speaking the number of nodes in its syntax tree. The proof is detailed in[10].

4.5 Compilation and Validation

To compile an ADeL program, our system first transforms it into an equation system which represents the synchronous automaton as explained in section 2. Then it implements directly this equation system, transforming it into a Boolean equation system thanks to the encoding defined in section 2 and to theorem 1. The latter system provides an effective implementation of the initial ADeL program.

Since the equations may not be independent, a valid order (compatible with their interdependencies) is needed to be able to generate code for execution (recognition automata), simulation, and verification. Thus an efficient sorting algorithm has been defined [11], using a critical path scheduling approach, which computes all the valid partial orders instead of one unique total order. This facilitates merging several equation systems, hence, an incremental compilation can be performed: an already compiled and sorted code for a sub-activity can be included into a main one, without recompiling the latter.

The internal representation as Boolean equation systems also makes it possible to verify and validate ADeL programs, by generating a format suitable for a dedicated model checker such as our own BLIF CHECK[4]. The same internal representation also allows us to generate code for the off-the-shelf NuSMV model-checker[5].

4.5.0.1 Use Case.

To illustrate our purpose, a small use case in the domain of healthcare is detailed. The goal is to monitor the drug treatment of an Alzheimer person. The activity *medicine To Take* must check that the person is near a table, takes a glass, eats some drugs, and drinks. If the person does not drink before 2 minutes, a *danger* event is sent.

In the graphical format, users declare roles of actors in the declaration window of the graphical tool. Then, they declare sub-activities, and describe the steps of their activity along a ”timelined organigram” (see Fig. 3).

[4]http://www.unice.fr/dgaffe/recherche/outils blif.html
[5]http://nusmv.fbk.eu/

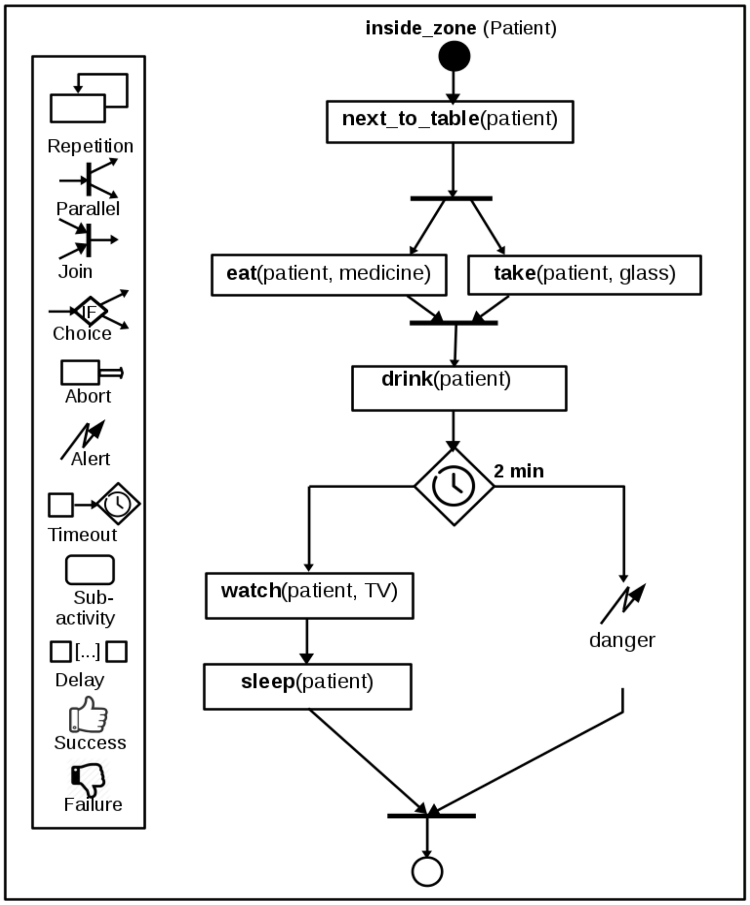


Figure 3. Graphical format of the activity description (organigram)

In the textual format of ADeL, users first declare types of actors. For this use case, there are 2 types: a Zone, a Person, and Equipment. Then, they have to assign roles to actors: in our case, a patient(Person), a medicine, a glass, a TV(Equipment) are needed. The declaration is as follows:

```
Type Person, Equipment;
Activity medicineToTake :
Roles
patient:Person;
medicine:Equipment;
glass: Equipment;
TV:Equipment;
```

After that, users define the name of the activity, its expected events and sub-activities:

```
SubActivities
next_to_table(Person);
take(Person, Equipment);
eat(Person,Equipment);
watch(Person,Equipment);
drink(Person); sleep(Person);
```

Finally, they describe the activity by defining the initial state, and by combining the sub-activities using operators of the language.

```
InitialState:inside_Zone(Patient);
start
{
next_to_table(patient)
then
eat(patient,medicine) parallel take(patient, glass)
then
drink(patient) timeout 2.0 minutes
{    watch(patient,TV)then alert ( danger)
then
sleep(patient)
}
End
```

This code is not correct because the alert should be sent went the timeout is reached. To prove that the alert works correctly, the *medicine To Take* activity is compiled and the input code for the NuSMVmodel-checkerisgenerated.ThentheLTLtemporalpropertycanbechecked: if *danger* is true then *cond_timeout_2_minutes* must have been previously true ("*danger* ⇒*O* *cond_timeout_2_minutes*")[6].

The property is false and a generated counter example allows to fix the problem. Hence, the correct body of the program is:

```
start
{
next_to_table(patient)
then
eat(patient,medicine) parallel take(patient, glass)
then
drink(patient) timeout 2.0 minutes
{
watch(patient,TV) then sleep(patient)
}alert(danger)
end
```

Now the property holds.

5. RELATED WORK

Synchronous languages such as Esterel [1] are meant to describe reactive systems in general and thus can be used to describe human activities. These languages and ADeL use a *logical* time which means that the recognition is performed only when something meaningful occurs. Although their syntax is rather simple, their large spectrum makes them difficult to master by

[6] *cond_timeout_2_minutes* is a Boolean variable true when the timeout is over

some end users. Being dedicated to activity description, a language like ADeL appears more ”natural” for its end users. All these synchronous languages have been given formal semantics. For instance, Esterel has several semantics, with different purposes. In particular, one of these semantics provides a direct implementation under the form of “circuits”. ADeL adopts a similar approach but it simplifies some operators whose semantics in Esterel is complex.

Message Sequence charts [12, 13], which are now introduced in UML, and Live Sequence Charts [14] are also specification languages for activities with a graphical layout that immediately gives an intuitive understanding of the intended system behavior. These languages may be given formal semantics liable to analysis. Message Sequence Charts (MSC) graphically represent the messages exchanged among the actors along time. It is possible to model a complex activity involving several different activities (i.e., MSCs) using High-level Message Sequence Charts (HMSCs). The HMSCs support also parallel composition. The MSC operators and the hierarchical composition of HMSCs are similar to our approach. However, [15] reveals “pathologies” in MSCs, due to defective MSC specifications. These pathologies mainly affect synchronization issues. For instance races may arise from discrepancies between the order of graphical description and system causalities. In our case, since the Synchronous Paradigm is meant to avoid these synchronization problems, race conditions are detected and the program is rejected at compile time. Another pathology comes from possible ambiguous choices between events. In the Synchronous approach, this kind of problems is avoided by producing deterministic systems, in particular mastering event simultaneity. MSCs address the pathology problems by using model checking and formal verification. In [13], the authors illustrate problems of the MSCs models verification for synchronous and asynchronous interpretations and suggest different techniques to fix these model checking problems in several kinds of MSCs representations. In our case, even though model checkers may be interfaced, it is not mandatory. Indeed, most of these pathologies are compile-time checked.

Live Sequence Charts (LSCs) [14] is another activity-based specification and modeling language. It is an extension of MSCs, more expressive and semantically richer. Similarly to ADeL, LSCs are used to specify the behavior of either sequential or parallel systems. They have a formal semantics and can be transformed to automataas ADeL. This allows analysis, verification, and testing using depth-first search methods. Model checking of LSCs is possible by translating them into temporal logic, but the size of the resulting formula, even for simple LSCs, makes it difficult. However, [16] proposes a more efficient translation, but only for a class of LSCs.

Many works in video understanding address the difficult task of extracting semantically significant objects and events from sequences of pixel-based images. A good survey of the corresponding techniques is presented in Lavee and all [17]. These techniques are based on well-founded mathematical methods such as hidden Markov models, (dynamic) Bayesian networks, finite state machines, Petri nets, constraint satisfaction, etc. The authors rely on tool based on such techniques to obtain reliable input events. These approaches allow a form of activity recognition (namely “composite events”) but ADeL addresses more complex activities with longer duration and involving variants, parallel behaviors and multiple actors. Moreover it loses the dependency on video sensors and proposes a more generic approach.

In [18], authors propose a natural and intuitive language to describe activity models using actors, sub activities, and a set of constraints. They also introduce a temporal constraint resolution techniques to recognize activities in real time. This approach is only dedicated to recognize activities using video interpretation, while the authors in this work aims to develop a generic approach that can be used in a large range of domains, by accepting basic events that can come not only from video interpretations but also from other sensors. On the other hand, as authors are working with video interpretation in real time, they can receive the same events (the same image frame) for a long lapse of time without any changements which makes the system awake and working for nothing. With the synchronous approach, the notion of logical time makes the system work only when it receives a significant event.

Researchers in [19] work in activity recognition in smart houses to provide Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) assistance for their users. They have developed a generic conceptual activity model which allows the modeling of simple and composite activities. To this aim, they propose an hybrid approach which combines ontological formalisms, which describes the link between the activities and their entities, and temporal knowledge representations which specify the relationships between sub-activities that form the composite activity. Then, they encode their characteristics and forms. In our case, ontologies are not used, the ADeL language has only semantics which help to generate the needed activity model to recognize simple and complex activities. Actually, a basic activity can be represented as an event or a simple activity. Activity models for complex/composite activities can be created by composing the sub-activity models which constitute them.

6. CONCLUSION AND PERSPECTIVES

This paper presents a formalization of a synchronous approach to describe (human) activities and to generate a computer recognition system. The Synchronous Paradigm offers several advantages in terms of expression power, ease of implementation, verification through model checking, etc. The authors endowed their own activity description language (ADeL) with two complementary formal semantics, one to describe the abstract behavior of a program, the second to compile the program into an automaton described as an equation system. They proved a theorem which establishes a consistency relation between these two semantics.

The first tests show that the current code that ADeL generates, basically composed of Boolean equations, is easy to integrate in a recognition system, produces compact code, and is efficient at run time. There remains a fundamental issue, common to all synchronous approaches: at the sensor level, the events are asynchronous and they must be sampled to constitute input environments and to define the synchronous "instants". No exact solution is available; several strategies and heuristics have been already tested but large scale experiments are still necessary.
Based on formal foundations, work remains to be done to complete a full framework to generate generic recognition systems and automatic tools to interface with static and dynamic analysis tools, such as model checkers or performance monitors.

REFERENCES

[1] N. Halbwachs. Synchronous Programming of Reactive Systems. Kluwer Academic,1993.

[2] Ge´rard Berry. Mechanized reasoning and hardware design. chapter Esterel on Hardware, pages 87–104. Prentice-Hall, Inc., Upper Saddle River, NJ, USA,1992.

[3] Esterel Technologies. Scade suite. http://www.ansys.com/products/embedded- software/ansys-scade-suite.

[4] G. H. Mealy. A method for synthesizing sequential circuits. Bell Sys. Tech. Journal, 34:1045– 1080, September1955.

[5] G. Berry. The Foundations of Esterel. In G. Plotkin, C. Stearling, and M.Tofte, editors, Proof, Language, and Interaction, Essays in Honor of Robin Milner.MIT Press, 2000.

[6] G. Berry. The Constructive Semantics of Pure Esterel. Draft Book, available at: http://www.esterel-technologies.com 1996.

[7] Matthew Ginsberg. Multivalued logics: A uniform approach to inference in artificial intelligence. Computational Intelligence, 4:265–316,1988.

[8] Daniel Gaffe´ and Annie Ressouche. Algebraic Framework for Synchronous Language Semantics. In Laviana Ferariu and Alina Patelli, editors, 2013 Symposium on Theoretical Aspects of Sofware Engineering, pages 51–58, Birmingham, UK, July 2013. IEEE Computer Society.

[9] A. Tarski. A lattice-theoretical fixpoint theorem and its applications. Pacific Journal of Mathematics, 5(2):285–309,1955.

[10] Ines Sarray, Annie Ressouche, Sabine Moisan, Jean-Paul Rigault, and Daniel Gaffe´. Synchronous Automata For Activity Recognition. Research report, Inria Sophia Antipolis, April 2017.

[11] Annie Ressouche and Daniel Gaffe´. Compilation modulaire d'un langage synchrone. Revue des sciences et technologies de l'information, se´rie The´orie et Science Informatique, 4(30):441–471, June2011.

[12] Thomas Gazagnaire, Blaise Genest ,Lo¨ıc He´loue¨t, P .S. Thiagarajan, Shaofa Yang, and Vasco T. Vasconcelos. Causal Message Sequence Charts, pages 166–180. Springer Berlin Heidelberg, Berlin, Heidelberg,2007.

[13] Rajeev Alur and Mihalis Yannakakis. Model Checking of Message Sequence Charts, pages 114–129. Berlin, Heidelberg,1999.

[14] L. Li, H. Gao, and T. Shan. An executable model and testing for Web software based on Live Sequence Charts. In 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), pages 1–6, June2016.

[15] Haitao Dan, Robert M. Hierons, and Steve Counsell. A framework for pathologies of Mes- sage Sequence Charts. Inf. Softw. Technol., 54(11):1283–1295, nov2012.

[16] Rahul Kumar, Eric G. Mercer, and Annette Bunker. Improving translation of Live Sequence Charts to temporal logic. Electron. Notes Theor. Comput. Sci., 250(1):137–152, September 2009.

[17] G. Lavee, E. Rivlin, and M. Rudzsky. Understanding video events: A survey of methods for automatic interpretation of semantic occurrences in video. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 39(5):489–504, Sept 2009.

[18] Van-ThinhVu, Francois Bremond, and Monique Thonnat. Automatic video interpretation: A novel algorithm for temporal scenario recognition. In Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI'03, pages 1295–1300, San Francisco, CA, USA, 2003. Morgan Kaufmann Publishers Inc.

[19] George Okeyo, Liming Chen, and Hui Wang. Combining ontological and temporal formalisms for composite activity modelling and recognition in smart homes. Future Generation Computer Systems, 39:29 – 43, 2014. Special Issue on Ubiquitous Computing and Future Communication Systems.

INTENTIONAL BLANK

CRESUS: A TOOL TO SUPPORT COLLABORATIVE REQUIREMENTS ELICITATION THROUGH ENHANCING SHARED UNDERSTANDING AND SIMULATION

Paul Stynes[1], Owen Conlan[2] and Declan O'Sullivan[3]

[1]School of Computing, National College or Ireland, Dublin, Ireland
[2,3]School of Computer Science and Statistics,
Trinity College Dublin, Dublin 2, Ireland

ABSTRACT

Communicating an organisation's requirements in a semantically consistent and understandable manner and then reflecting the potential impact of those requirements on the IT infrastructure presents a major challenge among stakeholders. Initial research findings indicate a desire among business executives for a tool that allows them to communicate organisational changes using natural language and a simulation of the IT infrastructure that supports those changes. Building on a detailed analysis and evaluation of these findings, the innovative CRESUS tool was designed and implemented. The purpose of this research was to investigate to what extent CRESUS both aids communication in the development of a shared understanding and supports collaborative requirements elicitation to bring about organisational, and associated IT infrastructural, change. This paper presents promising results that show how such a tool can facilitate collaborative requirements elicitation through increased communication around organisational change and the IT infrastructure.

KEYWORDS

Collaborative requirements elicitation, Shared Understanding, and Semantically enabled web services.

1. INTRODUCTION

Successfully understanding an organisation's requirements in such a manner that their impact on the IT infrastructure can be analysed and discussed, presents a major challenge [1-3] between the business executive, IT architect and other stakeholders. The most significant problem arises in communicating the requirements desired in a semantically consistent and understandable manner and then reflecting the potential impact of those requirements on the IT infrastructure.

Research indicates that there are numerous tools [4-7] that support requirements elicitation. However, none of them incorporate simulation as an elicitation technique to study the impact requirements will have on the organisational and associated IT infrastructure changes. Simulation models play an important role in displaying significant features and characteristics of a dynamic system, which one wishes to study [8]. One of the approaches Chandrasekaran et al [9] proposes

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 47– 61, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80205

is the creation of simulation models from web services. This approach provides a high fidelity between the simulation and real world. Moreover, it provides an ability to plug real web services into simulated entities, thus creating simulations that utilise as much 'real world' data as possible. An initial evaluation of a scenario-based prototype was conducted with ten business executives in higher education. The purpose of the evaluation was to examine the desire for a tool that allows business executives communicate about organisational and IT infrastructural changes [10]. The analysed data produce results that indicate a desire for such a tool with open comments such as "The idea seems to be very good, especially the natural language interface, which I particularly like", and "Good to specify objectives through the use of goals". Quantitative results reinforce the comments. They indicate that the majority of this small sample group (6 executives) like the approach of specifying their goals and rules in natural language. In addition, the majority (7 executives) liked the approach where the system automatically identifies a business process that may solve the executive's goal, and applies rules to services in the process. While this was only a limited survey the results indicate promise. The executives intuitively understood the approach being proposed. That is to augment the requirements gathering process with semantics that drives the formulation of organisational changes using controlled natural language. In addition, they also liked the simulation of the IT infrastructure that supports those changes. Moreover, they strongly indicated that they believed such an approach was desirable.

Building on the evaluation, CRESUS was designed and implemented as a collaborative requirements elicitation tool. The tool allows stakeholders communicate requirements guided by an ontological domain model, then validate the requirements, and finally create a simulation of the IT infrastructure that supports those requirements.

This research investigates to what extent CRESUS both aids communication in the development of a shared understanding and supports collaborative requirements elicitation to bring about organisational, and associated IT infrastructural, change.

The remainder of this paper describes the related work in section 2, then an outline of research, followed by architecture and implementation. Section 5 describes the evaluation, followed by limitations, and future work.

2. RELATED WORK

The elicitation or communication of requirements is recognised as a critical activity of software development [11]. The goal of which is to reach a shared understanding between all parties involved in the communication process. An increased amount of communication effort [12] is often required to overcome the gap in communicating the requirements desired in a semantically consistent and understandable manner and then reflecting the potential impact of those requirements on the IT infrastructure.

Effective communication has been difficult to achieve and is a recurring problem in the elicitation of requirements [13]. Macaulay [14] identifies communication as a key factor to the design of successful systems. Christal et al. [15] categorise the problems of requirements elicitation that relate to scope, understanding and volatility where communication features strongly in gaining an understanding. Coughlan et al. [16] conducted an analysis of effective communication in requirements elicitation. They discuss the user-centred (problem finding viewpoint that advocates an emergent and collaborative nature to requirements that emerge as part of on-going interactions and negotiations between participants. This problem finding approach involves spending more time communicating and developing relationships which has been linked to greater success in the determination of requirements [17]. Coughlan et al. [16] conclude with four recommendations for effective communication for an organization and its stakeholders in attempting to integrate technology namely, include users in the design; select an adequate mix of IT and business users who then interact on a cooperative basis; the incorporation of communication activities that relate to knowledge acquisition, knowledge negotiation and user acceptance; the use of elicitation techniques for mediating communication for the requirements of a system such as prototyping, questionnaires, brainstorming, scenarios, etc.

This research proposes to use simulation as an elicitation technique for creating the IT infrastructure. In this context, simulation has the potential to evolve into a prototype.

Tools that support requirements elicitation can be categorised as non web-based and web-based. Non web-based tools are VORDTool [18], FAES [5], AMORE [6], and JPreview [7]. The web-based tools are TeamWave [19] and WRET [20]. Web based tools support distributed collaborative requirements elicitation.

Current research on web-based tools indicates that they can support a collaborative requirements process and in some cases support the creation of requirements documentations. However, none of them incorporate simulation as an elicitation technique to study the impact requirements will have on the organisational and associated IT infrastructure change. This research proposes to address the issue by incorporating this technique into a web-based collaborative requirement elicitation tool.

Rogers and Kincaid [21] convergence model describes communication as a dynamic process of idea and knowledge generation, which occurs over time through interaction with others and which leads to mutual understanding and collective action. The work of Lind and Zmud [3] shows that frequent communication helps create a mutual understanding of the role of IT infrastructure in supporting the business functions. Johnson and Lederer [2] extend this work to communication between the Chief Executive and Chief Information Officers. Their measures however are based on the subject's perception and they recommend that communication frequency would benefit from objective assessments. In addition, they mention that future research would benefit if they incorporate other dimensions to the frequency of communication such as the richness of the communication. The recommendation of using objective measures for frequency of communication and other dimensions that lead to a shared understanding form the basis of evaluating this research.

Requirements are often written in natural language (e.g., English) which is inherently ambiguous. This allows customers that are unfamiliar with modelling notations to still understand and validate requirement specifications [22]. A controlled natural language is a precisely defined subset of full English that can be used for communicating the organisations requirements in such a way that it can be automatically reasoned over by machines and thus removes the ambiguity issues of natural language. Attempto Controlled English (ACE) [23] demonstrates how a controlled natural language describes knowledge for the semantic web. This allows the language to be understandable by people and machines. A predictive editor [24] guides the stakeholders, word-by-word in the construction of a sentence that complies with ACE. The sentence can be converted to an ontological format using the ACE Parsing Engine (APE) [25]. The approach described with ACE shows the potential for a natural language interfaces to utilize semantic inference to refine and hone requirements which form part of the controlled natural language interface in CRESUS.

Simulation models play an important and inexpensive role in displaying significant features and characteristics of dynamic system, which one wishes to study, predict, modify, or control [8]. The true potential of simulation is in portraying the envisaged impact of certain decisions and changes on an operational system or process. Chandrasekaran et al [9] examines the synergy between web service technology and simulation. One of the approaches they propose is the creation of simulation models or components from web services in order to provide a high fidelity between the simulation and real world. It provides an ability to plug real web services into simulated entities, thus creating simulations that utilise as much 'real world' data as possible. The proposed approach of creating a simulation model from web services forms the basis of the simulation platform that represents the evolution of the IT infrastructure in CRESUS.

In summary, the communication of requirements is recognised as a critical activity to reaching a shared understanding between stakeholders involved in software development. Effective communication has been difficult to achieve and is a recurring problem in the elicitation of requirements. An increased amount of communication effort is often required to overcome the gap in communicating the requirements desired in a semantically consistent and understandable manner and then reflecting the potential impact of those requirements on the IT infrastructure.

The literature review demonstrates the potential for controlled natural language interfaces to utilise semantic inference to collaboratively communicate organisational requirements. A simulation represented by web services provides the stakeholders with the means to effectively communicate the potential impact of the requirements on the IT infrastructure. Thus, the use of a simulation model comprised of semantically enabled web services could potentially represent the evolution of the organisation's IT infrastructure. The literature also demonstrates the criteria for evaluating this research using objective measures for frequency of communication and other dimensions that lead to a shared understanding.

3. OUTLINE OF RESEARCH

3.1. Research background

This research explores to what extent CRESUS both aids communication in the development of a shared understanding and supports collaborative requirements elicitation to bring about organisational, and associated IT infrastructural, change. In addition, the experiment controls for title, role, level of service, and academic qualifications of employees at the National College of Ireland.

The experiment employed a pre-test-post-test control group design with matching participants in the experimental and control groups to evaluate the frequency of communication, quality of communication, participant's identification of organisational change issues, participant perceptions of attaining a shared representation of the IT infrastructure supporting organisational change, and, control variables that relate to title, role, level of service and academic qualifications.

In this context CRESUS is generally defined as a collaborative communication tool that allows stakeholders to communicate requirements from an ontological domain model, validate the requirements, and create a simulation of the IT infrastructure that supports those requirements. Communication will encompass three dimensions namely, frequency of communication, quality of communication and the identification of organisational change issues. Shared understanding is generally defined as the stakeholders' perception that the requirements represent the organisational and associated IT infrastructural change.

Organisational and associated IT infrastructural change will be generally defined as a series of IT systems denoted by web services that are semantically enabled and may access ontological data that represents the organisational changes as denoted by the stakeholders' requirements.

3.2. Method

The following steps were carried out in developing CRESUS namely a literature review, evaluation of a scenario-based prototype, implementation of CRESUS and evaluation of CRESUS.

The literature review focused on organisational communication using a controlled natural language, shared understanding of organisational and associated IT infrastructural changes, collaborative requirements elicitation to bring about organisational and associated IT infrastructure change, and simulation that comprises of semantically enabled web services representing the evolution of the organisational IT infrastructure.

The evaluation of a scenario-based prototype is described in section 1. The implementation of CRESUS is described in section 4. Finally, the evaluation is described in section 5.

4. ARCHITECTURE AND IMPLEMENTATION

The architecture for the CRESUS tool was built around a controlled natural language interface, ontological domain model and a language translator that creates a simulation platform consisting of web services representing the evolution of the IT infrastructure as shown in Figure 1.

The architecture revolves around collaborative requirements elicitation between the business executive, IT architect and other stakeholders using controlled natural language. The language is guided by a lexicon that consists of nouns, proper nouns and verbs from the ontological domain model. The Knowledge Engineer creates an ontological domain model based on concepts, relationships and actual data from the problem domain.

The business executive, IT architect and stakeholders may create or modify requirements through the controlled natural language interface [24] as shown in Figure 2. These requirements are parsed to ensure that the grammar is based on a subset of natural language, namely Attempto Controlled English (ACE) [23] and stored in the ontological domain model. This results in machine accessible semantics that are automatically process-able. The controlled natural language interface component represents the only 3rd party component in this architecture.

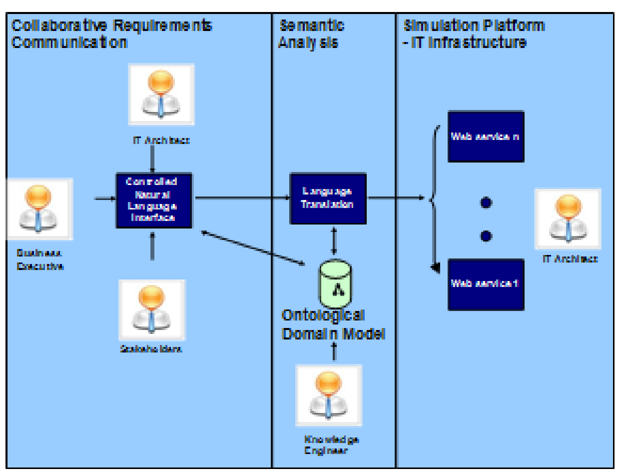


Figure 1. Architecture of CRESUS

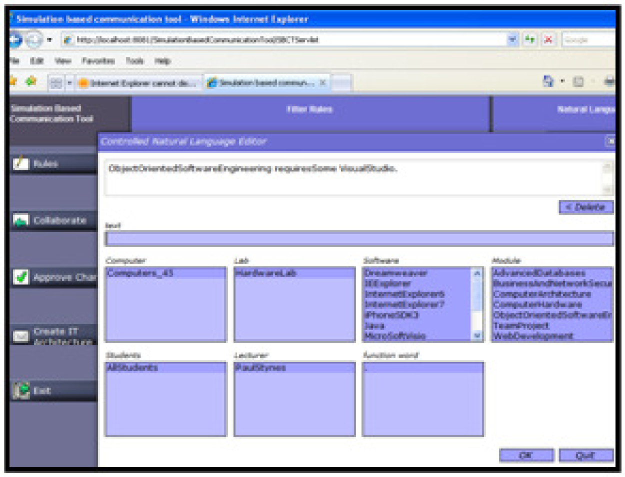


Figure 2. CRESUS – Controlled natural language interface (Preditor)

The language translator component parses the ontological domain model and applies a template rule to create the web services and simulation data.

Deployment of the web services results in the creation of the simulation platform that represents the evolution of the IT infrastructure. Staff can retrieve, create, delete and modify simulation data through the web service interfaces.

The innovation with CRESUS lies in the collaboration between the business executive, IT architect and stakeholders in creating a shared understanding of the evolving IT infrastructure that supports the requirements through the creation of semantically enabled web services and data structures during the simulation. The advantage of this approach is the high fidelity between the simulation and real world. It provides an ability to plug real web services into simulated entities, thus creating simulations that utilise as much empirical data as possible.

CRESUS is a web based application that is implemented in Java using the echo web framework [26]. The web based application is hosted on port 8081. Preditor [24] is a 3rd party product for the controlled natural language interface and is written in Java based on the echo web framework. The grammar that guides the controlled natural language interface is created using the Document Object Model (DOM) [27], The Attempto Controlled English Parsing Engine (APE) [25] server is a 3rd party product that is responsible for parsing the requirements from the business executive, IT architect and other stakeholders to ensure that they are a subset of ACE. In addition, APE converts the natural language to the Resource Description Framework (RDF) [28]. APE is set-up on port 8000. The ontological domain model is created using protégé [29]. The server side implementation uses Jena [30] which is a Java framework for building semantic web applications with RDF and the OWL Web Ontology Language [31]. The ontological domain model is incorporated into Jena. Jena uses Pellet [32] to reason over the ontological domain model. The server side implementation is also responsible for creating the simulation platform using the Java XML API and XSLT. XSLT is used for the language translation from the ontology domain model to the web services. The web services representing the simulation platform are implemented in Java and deployed on port 8083. Data for the web based application and the simulation platform is stored in an XML database, eXist [33] on port 8080. Interaction with the database involves the technologies XQuery, XUpdate and XPath.

5. EVALUATION

5.1. Background

A member of the IT department at the National College of Ireland was interviewed about the process for creating the software environment in the computer laboratories. This process involves gathering software and hardware requirements for each module from lecturers. Then creating and testing an image of the software environment. And finally rolling the image out to all the pc's in the computer laboratories. Several issues that arose related to getting the module requirements in a timely fashion and confusion over changes to the software requirements of modules. This confusion related to requests for different software versions. There was an indication that the current process may improve by modifying the IT infrastructure for identifying the software and hardware requirements of each module. The results of the interview formed the basis of the scenario that was given to the control and experimental groups.

5.2. Population

Twelve employees of the National College of Ireland took part in the experiment. Eight of the employees were academics and three were from the IT department. The academics consisted of four course directors (lecturer grade II), one lecturer grade I, one fulltime lecturer grade II from the school of business, one postdoctoral research fellow and one support tutor. All academics lectured in the school of computing. The IT personnel consisted of one senior IT administrator, one IT support specialist and two IT support personnel.

5.3. Operationalisation of variables

In this context, communication is operationalised by the frequency of communication, quality of communication, and the identification of organisational change issues. Frequency of communication indicates the degree to which messages and responses take place between the business executive, IT architect, and staff. Quality of communication indicates the degree to which messages that identify organisational changes are successfully approved by the appropriate authority. Organisational change issues indicate the degree to which messages that identify organisational change issues are discovered but there is no progress to resolve those issues by the participants.

Shared understanding indicates the degree to which the business executive, staff and IT architect perceive that the IT infrastructure supports the organisational change faithfully.

5.4. Methodology

Data was captured through online questionnaires, a workshop, logs, observations and a debriefing session. The level of service and qualification's online questionnaire is a pre-test that contains nine items in three major content sections. The content sections are demographic details, service details and qualification details. The majority of these questions asked for factual information such as the person's title, role and work experience in the organisation. Other questions contain selection buttons such as gender, age and the number of years' service with NCI. The control information that relates to the title, role, level of service and academic qualifications used in the analysis for matching participants is derived from this questionnaire's data. Thereafter one employee of each match pair is randomly assigned to the experimental group and the other to the control group. The workshop involves a brainstorming session with each group that involves creating a conceptual model of the problem domain which is converted into an ontological domain model. The ontological domain model is used to constrain the words in the controlled natural language interface and the naming conventions for the automatically generated web services. The matching participants experiment involves exposing all groups to email as the communication tool and exposing the experimental group to an additional treatment of communication through CRESUS. The experiment is biased towards the use of asynchronous communication where the control group use only one tool i.e. email and the experimental group use two tools i.e., email and CRESUS. During the experiment, a log of all messages is stored in a communal email for each group and in addition the CRESUS logs each message. The information that relates to the frequency and quality of communication, and organisational change issues used in the analysis is derived from the log's data. Observations are noted on an A4 writing pad. The asynchronous communication for decision making online questionnaire is a post-test that contains thirteen items in four major content sections. The content sections are demographics, frequency of communication, quality of communication and perception. The majority of these questions asked for factual information such as the person's name, list any positive aspects of the experiment, list any negative aspects of the experiment, and list any suggested improvements to the experiment. Another question on participant's perception is based on a 1-5 semantic differential scale from "strongly disagree" to "strongly agree". The information that relates to the participant's perception of attaining a shared representation of the IT infrastructure supporting organisational change, used in the analysis is derived from this questionnaire's data. The online questionnaire was chosen as a data collection procedure for the study because of the rapid turnaround in data collection and the proficiency of participants with technology. The experimental group are administered the simulation based communication tool online survey. The survey contains thirty-six items. It was created specifically for this research by this author and contains items from several instruments [34-36]. The major content sections of this survey are demographic details, perceived usefulness, usability heuristics, user interface satisfaction, screen layout, learning and project specific questions. The perceived usefulness identifies details on frequency and quality of communication, shared understanding of how IT can support organisational changes, and if the web service simulation of the IT architecture supports the organisational change faithfully. The project specific questions identify details on how understandable the scenario is and if components of the simulation-based communication tool are understandable. The final step involves a debriefing of the results and validation by the participants. The results are analysed based on a paired samples t-test using SPSS.

5.5. Results and data analysis

The eight academics and four IT personnel that took part in the experiment completed the pre-test questionnaire with results described in Table 1. The table describes the participants' role, title, number of years of service with the National College of Ireland and their level of educational qualification attained based on the National Qualification Framework. Academic's 2, 3, 5 and 8 with a title of lecturer grade II and where their role indicates that they are programme directors are assigned the role of business executive. The academics 7, 11, 10 and 12 whose titles are lecturer grade I, lecturer grade II from the school of business, postdoctoral research fellow and support tutor, and where their role indicates that they lecture in the school of computing are assigned to the role of staff. The IT personnel 1, 4, 6 and 9 with the titles senior IT administrator, one IT support specialist and two IT support personnel are assigned the role of IT architect. The academics and IT personnel are pair matched based on their number of years of service with the National College of Ireland and their level of educational qualification attained based on the national qualification framework. The matched participants are 2 and 3; 7 and 11; 4 and 6; 5 and 8; 10 and 12; and finally 1 and 9. Participants were randomly allocated to the experimental and control groups as follows, participants 2, 7 and 1 to experimental group one; participants 3, 11 and 9 to control group two; participants 8, 10 and 4 to experimental group three; and participants 5, 12 and 6 to control group four.

Table 1. Matching participants based on level of service and academic qualification

Role	Name	Title	Number of years of service with NCI						The level of education attained based on the National Qualification Framework		
			1-2 years	3-4 years	5-6 years	7-8 years	9-10 years	>10 years	Honours Degree	Masters	PhD
Programme Director	2	Lecturer II				X				X	
	3	Lecturer II				X				X	
	5	Lecturer II			X						X
	8	Lecturer II					X			X	
Lecture in School of Computing	7	Lecturer I	X						X		
	10	Computing Support Tutor	X							X	
	11	Post-Doctoral Research Fellow	X						X		
	12	Lecturer II						X		X	
IT personnel	1	Senior IT Administrator				X				X	
	4	IT support		X					X		
	6	IT support		X					X		
	9	IT Support Specialist				X			X		

Legend	
(yellow)	Experimental group 1
(green)	Control group 2
(orange)	Experimental group 3
(blue)	Control group 4

The results from the brainstorming session define the problem domain model and the approval process for any requirements that will be identified. The problem domain model is converted into an ontology domain model which is used to constrain the grammar in the controlled natural language interface and the naming conventions in the automatically generated web services. A sample of the conceptual model and approval process that one of the group's identified, related to the concepts of Module and Software. An example of a rule was "Module requires Software" with instance data that represents an organisational change such as "Software Development requires Netbeans". In one group the approval process was for the executive to approve any requirements that the staff identifies. If approval is granted, then the rule goes to the IT architect for further approval before becoming part of the ontological domain model.

The results from CRESUS logs and the email tool are displayed from Figure 3 to Figure 5. A comparative analysis of the data from Figure 3 indicates that the frequency of communication is significantly higher with the experimental group than the control group, where the experimental group use the CRESUS tool and email to communicate. The control group uses only email.

Figure 4 indicates that there are no significant differences in the quality of communication between the experimental and control groups. However, a deeper analysis of the results clearly shows that staff in the experimental groups (participants 7 and 10) identifies the majority of organisational changes that are successfully approved by the appropriate authority compared to the control groups. This result may be worth further exploration through additional experiments. Another observation was that in general the IT architect did not contribute to defining the organisational changes. This is as expected as the IT architect is responsible for identifying the IT infrastructure that provides support to the organisational change rather than forcing the organisational change in any particular direction. There was one exception to this observation in the fourth control group where the IT architect identifies one organisational change. The result in this instance is biased as the IT architect is actually part of the real process for rolling out the IT environment to the pc's and had knowledge that the other IT architects did not have.

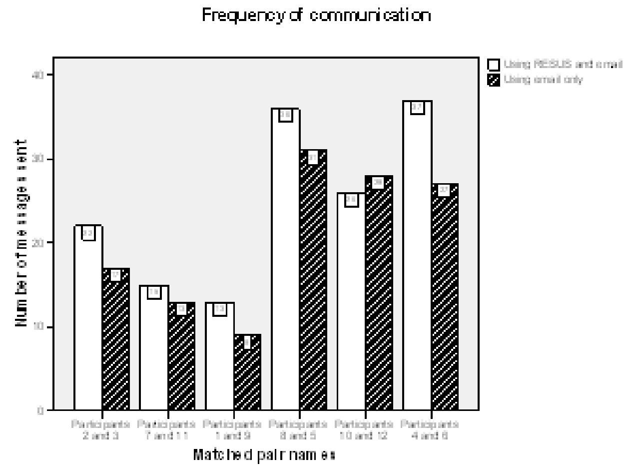


Figure 3. Frequency of communication

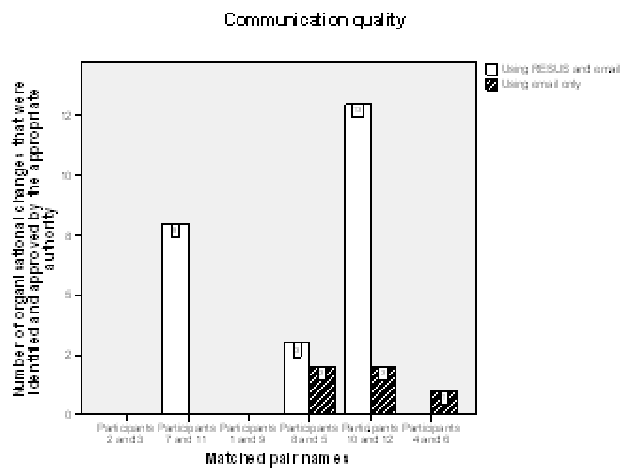


Figure 4. Communication Quality

The comparison in Figure 5 indicates that the identification of organisational change issues is significantly higher among participants in the control group that use email only, compared to the experimental group. Though initially unexpected, upon reflection the control group's communication was not constrained by identifying organisational changes only, and so the control group's communication was wider in scope which led to the identification of more issues. Other results from the post-test online questionnaire indicate there is no difference between the experimental and control groups for the participant's perception of attaining a shared representation of the IT infrastructure supporting organisational change. On reflection one of the goals of all four groups was to identify an IT system that could support the organisational changes. As the operational metric was based on perception then it stands to reason that there will be no difference between the groups in identifying the IT infrastructure that supports the organisational changes. The qualitative feedback from the experimental groups indicates that CRESUS facilitates collaboration and communication around the IT architecture with comments such as "Very useful tool for defining requirements and collaboration" and "Provided a new means for communication around IT infrastructure".

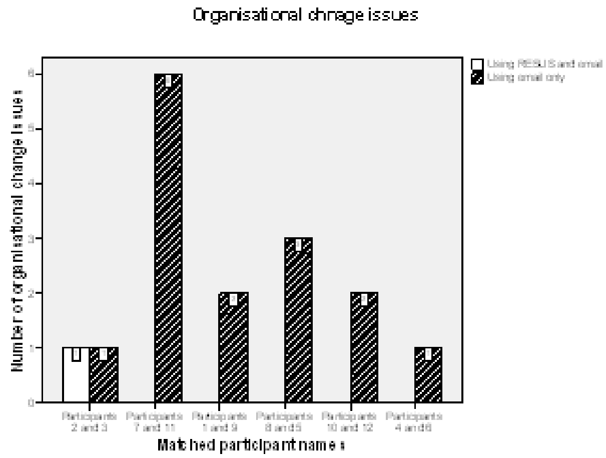


Figure 5. Organisational change issues

The design of CRESUS ensures that participants focus the communication on identifying requirements that relate to the organisational changes and approving those changes. Followed by, an automatic creation of the IT architecture to support those changes. This was as observed with the experimental group's one and three. An observation of control group two by the experimenter and validated by the group indicates that they spent a substantial amount of work defining the IT infrastructure. This group had two suggested solutions, but had not made a decision on which solution to adopt. One of the participants of that group made the following comment "Maybe the contributed should be told not to focus on specific technologies too much". This comment highlights an advantage of using CRESUS in that it allows the participants to focus on communicating about the business and not the technology.

6. LIMITATIONS

This was the first experiment with the tool and so it was important to get early feedback and direction on further development and experimentation. As such the experiment took place at the National College of Ireland. One of the goals of the experiment was to demonstrate that the architecture has the possibility to scale up to the full expressivity of the controlled natural language and so for this experiment sentences were constrained to simple "noun verb noun".
Maturation was seen as a threat to the matched participants' research design used in this experiment. Maturation is where participants mature or change during the experiment. The experimenter attempted to select participants that would mature at the same rate. In group one and group two, participant 2 and participant 3 are both studying for a PhD. During the experiment, participant 2 completed the viva. Also in group one and group two, participant 7 and participant 11 both currently have an honours degree and are both completing a PhD. During the course of the experiment, participant 11 completed a viva. Survey items that capture potential maturation should be incorporated into a pre-test.

Metrics used for the participants' perception of attaining a shared representation of the IT systems that supports the organisational change were not suitable for a matching participant's research design. As each group is communicating collaboratively to make decisions about organisational changes and creating the IT system that supports those changes. Thus it stands to reason that there will be no significant differences in perception. Objective metrics instead of participants' perceptions should be used in further experimentation.

In group four, participant 6 is involved with the roll out of the IT environment and this extra knowledge would have biased the results in particular when identifying an IT system and identifying organisational change solutions, issues and constraints. Survey items that capture the participants' role in relation to the scenario should be incorporated into a pre-test and controlled for in future experiments.

A weakness in the experimental design was that the controlled experiment was conducted in a one hour setting which would not be representative of actual communication behaviours among the employees. Further experimentation should be conducted over a longer period of time and incorporated into their actual job so that it is representative of their actual communication patterns. The limitations will be addressed in further experimentation.

7. CONCLUSION

The goal of requirements elicitation is to reach a shared understanding between all parties involved in the communication process which often involves an increased amount of communication effort to overcome the gap in communicating the requirements desired in a semantically consistent and understandable manner and then reflecting the potential impact of those requirements on the IT infrastructure.

An initial study conducted among ten business executives in higher education indicates a desire by the majority of this small group for a tool that allows them to communicate organisational changes using natural language where these changes are automatically translated into the IT infrastructure that supports a business process.

Building on this research, CRESUS was implemented as a collaborative communication tool that allows stakeholders to communicate requirements from an ontological domain model, validate the requirements, and create a simulation of the IT infrastructure that supports those requirements. The tool was evaluated at the National College of Ireland.

Results show that the tool significantly increases the frequency of communication which is a predictor of reaching a shared understanding between all stakeholders during requirements elicitation. In addition, evidence from the experiment suggests that the tool increases the quality of communication initiated by staff. Qualitative feedback from participants that use CRESUS indicates that the tool facilitates collaboration and communication around the IT architecture with comments such as "good for storing knowledge and facilitating collaboration" and "Provided a new means for communication around the IT infrastructure".

In conclusion CRESUS shows promise as a tool for collaborative requirements elicitation through increased communication and understanding around organisational and IT infrastructural change.

8. FUTURE WORK

Web services can be orchestrated into an executable business process using the business process execution language. As such future work will involve investigating CRESUS's role in collaborative requirements elicitation for the creation of the IT infrastructure that supports a business process.

Further evaluation is required to investigate the significance of the quality of communication among staff, objective metrics for attaining a shared understanding of the IT infrastructure that supports the organisational change with the provision that the metrics are objective and not based on the participant's perception. The experiment will be conducted in different organisations over a longer period of time such that it replicates their actual communication patterns. Questionnaire items that capture maturation and participants' role that relates to the organisational change scenario will be included in the pre-test.

REFERENCES

[1] Preston, D. S., Karahanna, E., & Rowe, F. (2006). Development of shared understanding between the chief information officer and top management team in U.S. and French Organizations: A cross-cultural comparison. IEEE Transactions on Engineering Management, 53(2), 191-206. doi: 10.1109/TEM.2006.872244.

[2] Johnson, A.M., & Lederer, A.L., (2005). The effect of communication frequency and channel richness on the convergence between chief executive and chief information officers. Journal of Management Information Systems / Fall 2005, Vol. 22, No. 2, pp. 227-252.

[3] Lind, M. R., & Zmud, R. W. (1991). The Influence of a Convergence in Understanding between Technology Providers and Users on Information Technology Innovativeness. Organization Science, 2(2), pp. 195-217.

[4] Sommerville I., Software Engineering, 6h Edition, Addison-Wesley,Harlow, England. 2001.

[5] Gilvaz A. P.,and J. C. S.P. Leite, .FAES : A Case Tool for Information Acquisition., Case .95, Proceedings of the Seventh Intl. Work on Computer Aided Software engineering. IEEE Computer Society Press, July 1995, pp. 260-269.

[6] Christel, Micheal G, et al. (1993) .AMORE: The Advanced Multimedia Organizer for Requirements Elicitation., Technical Report CMU/SEI-93-TR-12.

[7] Sawyer P., Sommerville, I., and Viller, S. (1996). PREview: Tackling the Real Concerns of Requirements Engineering., CSEG Technical Report, Computing Department, Lancaster University.

[8] Kellner, Madachy, and Raffo, (1999), "Software Process Modeling and Simulation: Why, What, How," Journal of Systems and Software, Vol. 46, No. 2/3.

[9] Chandrasekaran, S., Silver, G., Miller, J., Cardoso, J. & Sheth, A. (2002) XML-based modeling and simulation: Web service technologies and their synergy with simulation. IN: Proceedings of the 34th Winter Simulation Conference: exploring new frontiers (WSC 2002), 8-11, December, San Diego, California, USA. Winter Simulation Conference. pp 606-615.

[10] Stynes, P., Conlan, O., O'Sullivan, D. (2008) Towards a Simulation-based Communication Tool to Support Semantic Business Process Management. IN: Proceedings of the Fourth International Workshop on Semantic Business Process Management in Proceedings of Workshops held at the Fifth European Semantic Web Conference, (ESWC08), 2nd June, Tenerife, Canary Islands, Spain.

[11] Gottesdeiner, E., Requirements by Collaboration, Addison-Wesley, 2002.

[12] Bostrum, R.P. (1989). Successful application of communication techniques to improve the systems development process. Inform Manage 1989; 16: 279-295.

[13] Saiedian H, Dale R. Requirements engineering: making the connection between the software developer and the customer. Inform Software Tech 2000; 42 (6): 419-428.

[14] Macaulay LA. Requirements Engineering. Springer-Verlag, London, 1996.

[15] Christel MG, Kang KC. (1992). Issues in requirements elicitation. Technical Report CMU/SEI-92-TR-12 ESC-TR-92-012.

[16] Coughlan, J., & Macredie, R. D. (2002). Effective communication in requirements elicitation: A comparison of methodologies. Requirements Engineering, 7(2), 47–60.

[17] Marakas, G.M., Elam, J.J., (1998). Semantic structuring in analyst and representation of facts in requirements analysis. Inform Syst Research; 9 (1): 37-63.

[18] VordTool, available at URL http://www.filewatcher.com/b/ftp/ftp.gmu.edu/drine/cs_se2/CourseMaterials/SupplementaryProjectMaterials/Supplements_TextBookChapter7_Technology-0.html.

[19] TeamWave Software Ltd., Availabale at URL http://www.teamwave.com/, 2001.

[20] Hassan, S., & Salim, S. S. (2004). A Tool to Support Collaborative Requirements Elicitation using Viewpoint Approach'. In 1st International Conference on Informatics, Izmir, Turkey, September 2004, (pp. 1–4).

[21] Rogers, E.M., and Kincaid, D.L. Communication Networks. New York: Free Press, 1981.

[22] Institute of Electrical and Electronic Engineers, (1998), IEEE Recommended Practice for Software Requirements Specifications. IEEE Standard 830-1998, Institute of Electrical and Electronic Engineers, New York.

[23] Norbert E. Fuchs, Kaarel Kaljurand, and Tobias Kuhn. (2008). Attempto Controlled English for Knowledge Representation. In Cristina Baroglio, Piero A. Bonatti, Jan Maluszynski, Massimo Marchiori, Axel Polleres, and Sebastian Schaffert, editors, Reasoning Web, Fourth International Summer School 2008, number 5224 in Lecture Notes in Computer Science, pages 104–124. Springer.

[24] Kuhn, T., (2008). Combining Semantic Wikis and Controlled Natural Language. Proceedings of the Poster and Demonstration Session at the 7th International Semantic Web Conference (ISWC2008), CEUR Workshop Proceedings, 2008.

[25] Kaarel Kaljurand and Norbert E. Fuchs. (2006) Bidirectional mapping between OWL DL and Attempto Controlled English. In Fourth Workshop on Principles and Practice of Semantic Web Reasoning, Budva, Montenegro.

[26] Echo2 Web framework Available from http://echo.nextapp.com/site/ [last accessed on 11th February, 2010.

[27] Document Object Model (DOM). Available from http://www.w3.org/DOM/ [last accessed on 11th February, 2010]

[28] Resource Description Framework (RDF). Available from http://www.w3.org/RDF/ [last accessed on 11th February, 2010]

[29] Protégé. Available from http://protege.stanford.edu/ [last accessed on 11th February, 2010]

[30] Jena – A semantic web framework for Java. Available from http://jena.sourceforge.net/ [last accessed on 11th February, 2010]

[31] OWL Web Ontology Language. Available from http://www.w3.org/TR/owl-features/ [last accessed on 11th February, 2010]

[32] Pellet: OWL 2 Reasoner for Java. Available from http://clarkparsia.com/pellet [last accessed on 11th February, 2010]

[33] XML Database, eXist. Available from http://exist.sourceforge.net/ [last accessed on 11th February, 2010]

[34] Nielsen, J. (1993) Usability Engineering. Academic Press. Chapter 5, p. 115. http://hcibib.org/perlman/question.cgi?form=NHE

[35] Chin, J.P., Diehl, V.A., Norman, K.L. (1988) Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface. ACM CHI'88 Proceedings, 213-218. ©1988 ACM. http://hcibib.org/perlman/question.cgi?form=QUIS.

[36] Davis, F. D. (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technolog. http://hcibib.org/perlman/question.cgi?form=PUEU

AUTHORS

Dr Paul Stynes

Dr Paul Stynes is Vice Dean of Academic Programmes and Research at the National College of Ireland. He completed a PhD in 2015 at Trinity College Dublin. His research interests are in the area of Collaborative requirements elicitation and Intelligent Systems specifically semantic web, and ontologies.

Prof. Owen Conlan

Prof. Conlan is an internationally recognised research leader in User Modelling, Adaptation, Personalisation and Visualisation research with over 160 publications in those fields. Owen is a Fellow of Trinity College Dublin and leads the Personalisation research in the ADAPT Centre (www.adaptcentre.ie).

Prof. Declan O'Sullivan

Prof. Declan O'Sullivan is Director of Research and Head of Discipline for Intelligent Systems at Trinity College Dublin. Declan was awarded a B.A. (Mod) in Computer Science from TCD in 1985, an M.Sc. in Computer Science from TCD in 1988, and a Ph.D. in Computer Science from TCD in 2006. Declan has substantial research experience in academia and industry, having worked for IONA Technologies and Broadcom Eireann Research.

INTENTIONAL BLANK

INCREASING THE ARCHITECTURES DESIGN QUALITY FOR MAS: AN APPROACH TO MINIMIZE THE EFFECTS OF COMPLEXITY

Howayda Abdallah Ali Elmarzaki and Tawfig M. Abdelaziz

Department of Software Engineering, Benghazi University, Benghazi, Libya

ABSTRACT

The efficiency of multi agent system design mainly relies on the quality of a conceptual architecture of such systems. Hence, quality issues should be considered at an early stage in the software development process. Large systems such as multi agents systems (MAS) require many communications and interactions to fulfil their tasks, and this leads to complexity of architecture design (AD) which have crucial influence on architecture design quality. This work attempts to introduce approach to increase the architecture design quality of MAS by minimizing the effect of complexity.

KEYWORDS

Multi agent system (MAS), a general architectures, Quality attributes, Recommendations systems (RS).

1. INTRODUCTION

MAS belong to the field of Artificial Intelligence, the study addressing the approaches of construction of complex systems using a large number of entities, which alter their behavior in order to accommodate with a particular problem[1], [2]. An intelligent agent can be reactive and proactive,[3] because it responses to the actions and alteration which appears in the working environment, can tack the initiative to establish the goals and interacts with other agents[4], [1],[5]. Most literatures refer that the complexity emerges clearly in architecture design of multi agent systems that assigned many and different tasks[6], [7], [8].The research work introduces an approach to increase the architecture design quality of MAS by minimizing the effect of complexity. The solution mainly presents a set of guidelines including the influential factors on the complexity of architecture design. These factors are extracted from several sides of AD. Several factors and guidelines are presented to decrease the complexity in architectures of multi agent systems. Each FG is established based on developer's previous practice or experimental methods. The FG is extracted from concepts which related to software architecture and they are presented as symbols used in application phase. For example, depending on FGM1 the hierarchical decomposition approach can be applied on books recommendation system to demonstrate the main components in visual manner to increase the understandability. The

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 63– 74, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80206

modularity has a major role in decreasing the complexity in software design since the interaction among agents to accomplish their tasks can lead to system complexity. Thus, this approach increases the architecture design quality of MAS by minimizing the effect of complexity. The reduction of complexity from AD, eventually reinforces the reusability concept.

2. Proposed Solution Approach

- The proposed solution is to achieve the desired goals of this research work. It mainly presents a set of guidelines including the influential factors on the complexity of architecture design. These factors are extracted from several sides of AD which should be taken into consideration at the early stages of developing the architecture.

- The sides represent concepts (Abstraction, Modularity and Modeling) which be able applying in both analyses and design phases. Figure1 illustrating the approached conceptsin FG4Complexity approach.

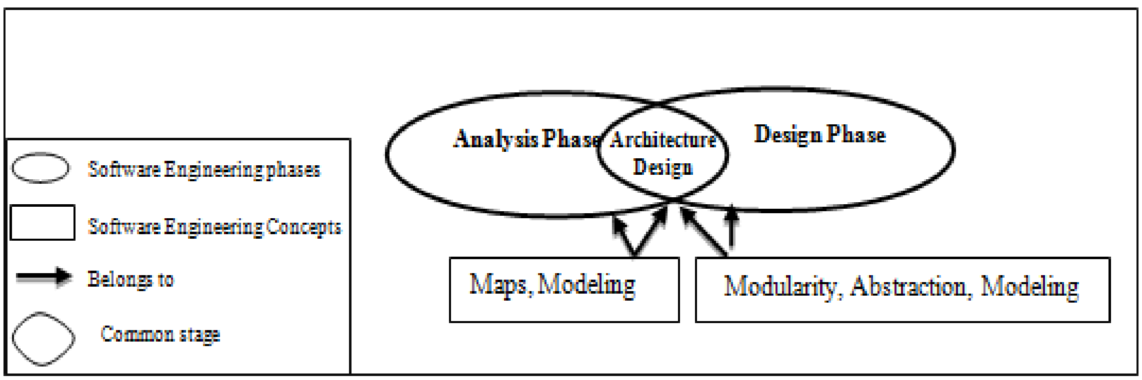


Figure 2: The concepts of analyzing and design which were addressed in FG4 Complexity approach.

- To label the proposed solution approach we suggested that "FG4Complexity". Thereby, "F" liter means Factors, "G" liter means Guidelines, and the "number 4" means for. The next figure shows the proposed approach mechanism.

- The work will be applied via some models used in methodologies related to agents systems such as HLIM[9], MASD [10].

2.1 Factors and Guidelines (FG)

In this section several factors and guidelines are presented to decrease the complexity in architectures of multi agent systems. Each FG is established based on developer's previous practice or experimental methods. The FG is extracted from concepts which related to software architecture and they are presented as symbols used in application phase. For example, the FG is related to modeling concept and represented by FGMOD symbol. The FG is related to abstraction concept and represented by FGA symbol and the FG is also related to modularity concept and represented by FGM symbol. Also, each FG should be numbered for example, FGA4 means the factor and guideline number4 in abstraction concept section, FGMOD2 means the factor and guideline number2 in modeling concept section as illustrated in the table below.

Table1: The symbols interpretation of architecture concepts

Instances	Symbols Interpretation	symbols	Architecture Concept
FGA1….i where I is Integer number	Factors and Guideline of Abstraction	FGA	Abstraction
FGM1….i where I is Integer number	Factors and Guideline of Modularity	FGM	Modularity
FGMOD1….i where I is Integer number	Factors and Guideline of Modeling	FGMOD	Modeling

Factors and Guidelines for Abstraction (FGA)

FGA1. Developers should use Simplifying Abstraction type if they want to decrease the dynamic complexity type. [11]

FGA2. Choosing the appropriate level of abstraction.[12]

FGA3. Avoid to adopting the concept of (gold plating).[13]

Factors and Guidelines for Modularity (FGM)

FGM1. Using Hierarchical Decomposition Approach (HDA) which considers a major method of handling complexity in conventional software analysis and design. [6], [14], [15]

FGM2. It is useful to establish the software modularity based on roles or measurements such as Cohesion Communication Measurment (CCM). [16]

Factors and Guidelines for Modeling (FGMOD)

FGMOD1. Using Use Case Maps (UCM) to clarify the most relevant, interesting, and critical tasks of MAS system. [17]

FGMOD2. Using simple notations is very important to enhance understandability and decrease complexities in AD such as arrows, components, domains…etc. [18]

3. CASE STUDY APPLICATION STEPS AND DISPLAY THE RESULTS

The case study is a "books recommendations system" based on MAS to help users select books. The system can switch to three recommendation approaches Content-based filtering approach (CBF) [19], [20] Collaborative Filtering approach (CF) [21], [22] and knowledge based approach (KBA).[23], [24] The agents within the system can exchange the messages among each other via one of agent communication languages. In this case study, the messages exchanged will be via Knowledge Query and Manipulation Language (KQML).The work will be applied via some models used in methodologies related to agents systems such as HLIM[9], MASD [10].

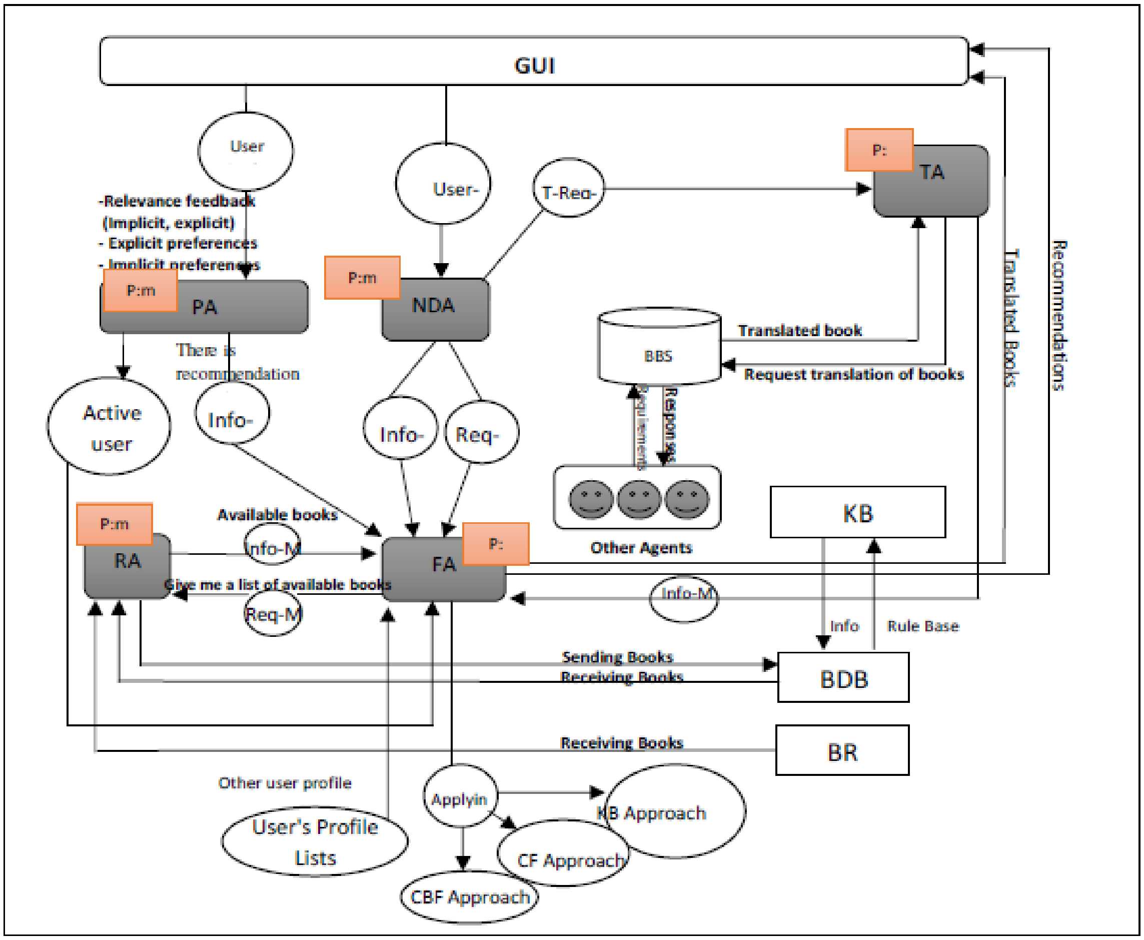
3.1 Agents and Their Tasks

A brief summary of agents and their tasks in the next table:

Table 2: The agents and their tasks

Agents	**Roles (Tasks)**
Profiling agent	• Gathering the user's preferences, gathering the relevance feedback, and building and updating the active user profile
NDA	Gathering the user current needs
Filtering agent	• Producing the recommendations, removing the books that are not currently offered from the recommendation list, and transferring the recommendation to the GUI
Retrieval agent	• Retrieving the books that are currently offered from the books database and storing the available books in the recommender system database
Translation agent	• Producing books translation service for users

3.2 Conceptual Overview of Books Recommendations System Architecture Design



3.3 The FG4 Complexity Approach Application Strategy

As we have earlier pointed out that all the previous FG will be within 4 steps to correspond to the current case study as the next figure shows:

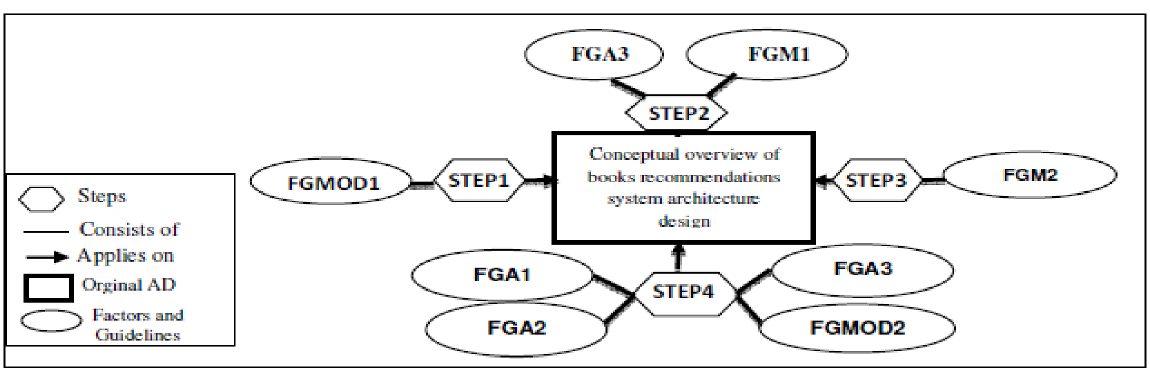


Figure 3: Illustrating of the applied steps on AD

Step1. Initially, this step is based on applying UCM represented in FGMOD1 of FG4Complexity approach which used in between analysis and design phases. These maps give high view of system specifically the responsibilities (Tasks) and interactions in a simple way, reinforce system understanding and overcome some situations of complexity such as intercommunication among agents. The following figure illustrate example to use the use case maps in analysing agents, tasks, scenarios and the most significant interactions among agents in books recommendations system. [25], [26]

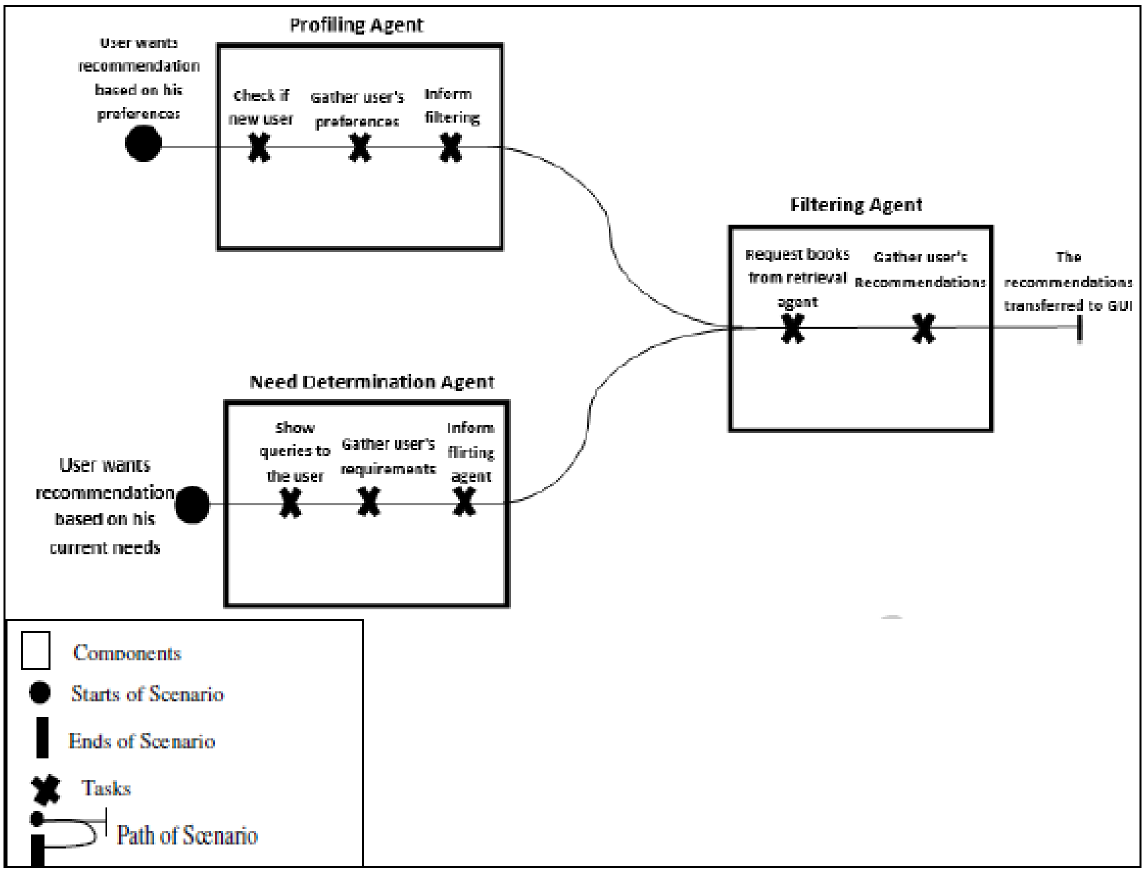


Figure 4: The UCM of translating book mechanism.

Step 2. If the system requirement specifications (SRS) [27] of a system do not have a translation function; then, this function is considered as Gold Plating concept; therefore, we should apply the FGA3 which avoid the part of gold plating represented in translation agent (TA) and all components connected from AD as illustrated in the figure below.

Depending on FGM1 the hierarchical decomposition approach (HDA) could be applied on books recommendation system to demonstrate the main components in visual manner to increase the understandability. Next table shows the main components and their connected components in books recommendations system.

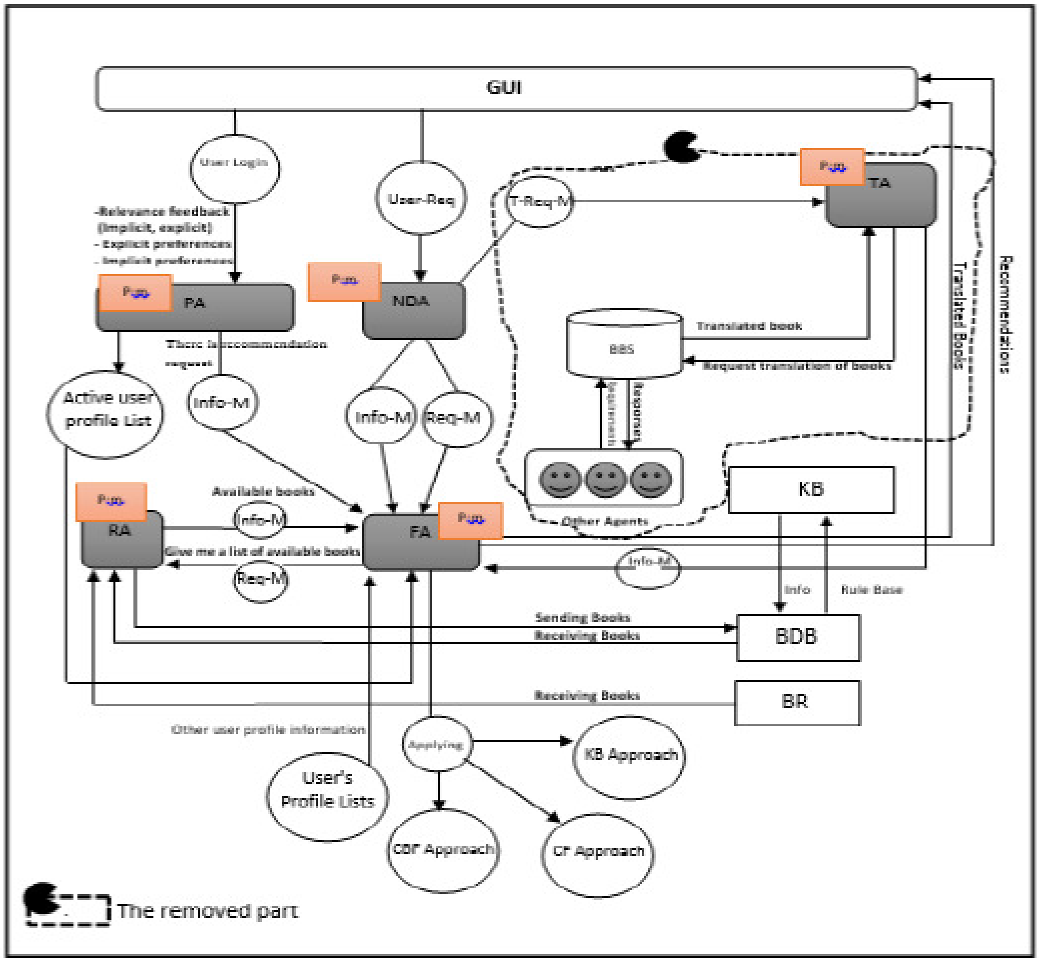


Figure 5: Omitting the part representing the gold plating

Table 3: The main components and their connected components in books recommendations

Main Components	Connected component(1)	Connected component(2)	Connected component(3)
Retrieval Agent	Book Data Base	Filtering Agent	Book Resource
Filtering Agent	Knowledge Base	GUI	Retrieval Agent
Profiling Agent	GUI	-	-
Need determination Agent	GUI	-	-
Book Data Base	Retrieval Agent	-	-
Book Resource	Retrieval Agent	-	-
Knowledge Base	Filtering Agent	-	-
GUI	Profiling Agent	NDA	Filtering Agent

Next figure demonstrates the majeure components in case study by applying HDA.

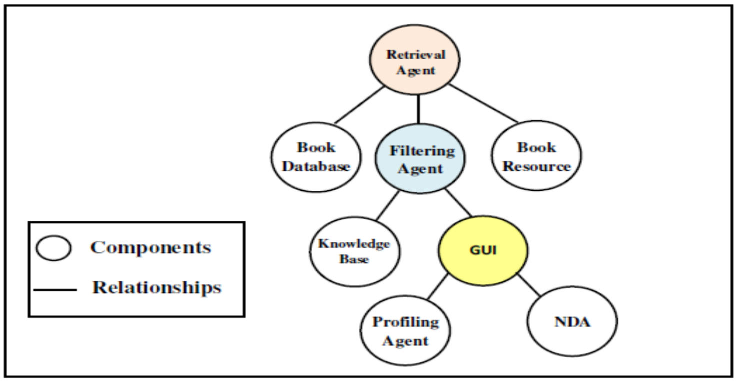


Figure 6: Conceptual system after applying HAD

Step 3. As we have pointed out, the modularity has a major role in decreasing the complexity in software design since the interaction among agents to accomplish their tasks can lead to system complexity. This step totally relies on cohesion measurement principle which uses the Communication Cohesion Measurement (CCM). This measurement works as a testing tool. This enables us to discover which agent needs more decompositions. In this research work, we have four agents described in the case study: filtering agent, profiling agent, need determination agent, and retrieval agent in respect that the translation agent has been omitted in the last step. The formulation of communication cohesive measurement is .The next illustration shows how.

Based on the architecture design of book recommendation system, the filtering agent has 4 internal relationships and 2 external relationships, profiling agent has just one internal relationship and 4 external relationships, need determination agent has one internal relationship and 2 external relationships and retrieval agent has 4 internal relationships and 3 external relationships as shown in the following:

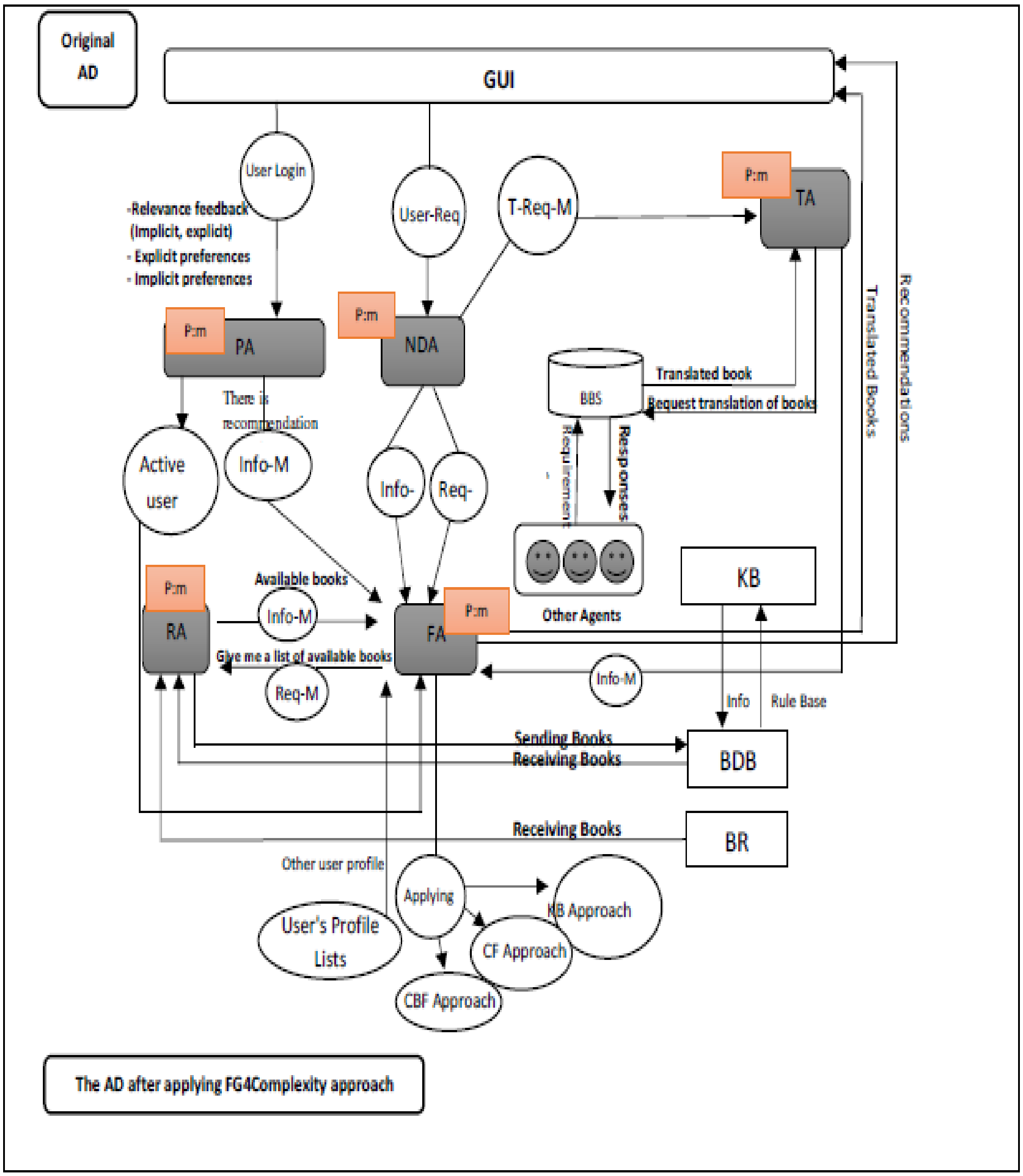
$$\mathbf{CCM\ (Ai)} = \frac{\mathbf{R}\text{ internal}}{\mathbf{R}\text{ internal+ }\mathbf{R}\text{ External}}$$

Table 4: The calculating by using CCM technique

Profiling agent		**Filtering agent**		**NDA**		**Retrieval agent**	
R internal	1	**R** internal	7	**R** internal	1	**R** internal	3
R external	2	**R** external	4	**R** external	3	**R** external	2
CCM(**PA**)	1/3	CCM(**FA**)	7/11	CCM(**NDA**)	1/4	CCM(**RA**)	3/5
Assessment		**Assessment**		**Assessment**		**Assessment**	
CCM(**PA**) = 0.3		CCM(**FA**) = 0.6		CCM(**NDA**) = 0.3		CCM(**RA**) = 0.6	

So, the results are:CCM (FA)<0.91, CCM(NDA) <0.91, CCM(RA) <0.91, and CCM(PA)<0.91. It is worth noticing that all results less than 0.91 by this, they do not need more decomposition.

Step 4. Applying a group of FG on the architecture design. This group consist of FGA1, FGA2, FGA3 and FGMOD2 which influence the architecture directly and the changes can clearly be observed.Next figures show the architectural design before and after applied **FG4Comlexity** approach.



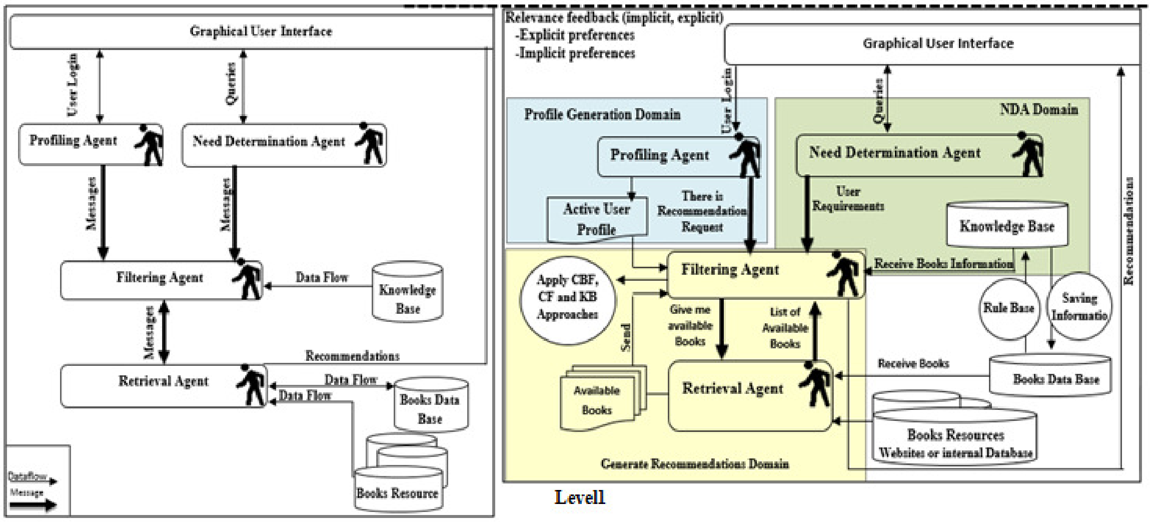


Figure 7: Displaying the AD before and after applying the FG4Complexity approach

4. CONCLUSION

The Research work approached the complexity of architectures design (AD) in systems based on multi agents (MAS) by a proposed solution method represented in a set of guidelines. These guidelines were introduced by extracting the factors affecting the complexity from three major sides of AD represented in abstraction, modularity and modeling thus, the approach labeled as "FG4complexity". It discussed the decrease of coupling which usually occurs during the interactions among agents and supporting the understandability of MAS architectures. The FG4complexity approach is useful for large systems such as recommendation systems that are based on MAS to avoid the complexity problems found in the most existing architectures. Thus, it enhances the quality standards, the reduction of complexity from (AD), and eventually reinforces the reusability concept.

FUTURE WORK

For future work, other aspects of architecture design will be addressed to attempt to make the proposed approach more effective. Those aspects may be are represented in the style, design patterns, documentation and so on. ALSO, we hope to apply the FG4complexity approach on other larger and more complex systems.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank ALLAH, without ALLAH this work would never have been finished.

I would like to express my sincere thanks to my supervisor Dr. Twfig Eltwel for his invaluable guidance and advice. I would like to thank my beloved husband Fathi El faitouri for his unlimited and faithful support as well as his patience and unconditional love. Also, The last but not least, I am profoundly grateful to my kind friend Asya Sohaim for her fruitful collaboration and advice. Finally, I thank everyone who encouraged me.

REFERENCES

[1] I. a. S. Markic, Maja and Maras, Josip, "Intelligent Multi Agent Systems for Decision Support in Insurance Industry," in Information and Communication Technology, Electronics and Microelectronics (MIPRO), ed: IEEE, 2014, pp. 1118--1123.

[2] M. Oprea, "Applications of multi-agent systems," in Information Technology, ed: Springer, 2004, pp. 239-270.

[3] Z. S. Ahmed Taki, "Formal Specification of Multi-Agent System Architecture," presented at the International Conference on Advanced Aspects of Software Engineering, ICAASE,, 2014.

[4] K. O. Chin, K. S. Gan, R. Alfred, P. Anthony, and D. Lukose, "Agent Architecture: An Overviews," Transactions on science and technology, vol. 1, pp. 18-35, 2014.

[5] S. A. D. Mark F. Wood, "An Overview of the Multiagent Systems Engineering Methodology," Springer, vol. 1957, 2001.

[6] B. H. Far, "Software Agents: Quality, Complexity and Uncertainty Issues," IEEE, 2002.

[7] B. R. Sinha, P. P. Dey, M. Amin, and H. Badkoobehi, "Software complexity measurement using multiple criteria," Journal of Computing Sciences in Colleges, vol. 28, pp. 155-162, 2013.

[8] D. N. M. Ghazal Keshavarz, Dr. Mirmohsen Pedram "Metric for Early Measurement of Software Complexity," International Journal on Computer Science and Engineering (IJCSE) vol. 3, 2011.

[9] M. Elammari and W. Lalonde, "An agent-oriented methodology: High-level and intermediate models," in Proc. of the 1st Int. Workshop. on Agent-Oriented Information Systems, 1999, pp. 1-16.

[10] M. E. T. Abdelaziz1, R. Unland3, C. Branki4, "MASD: Multi-Agent Systems Development Methodology," Multiagent and Grid Systems Journal,, 2010.

[11] S. a. D. Wagner, Florian, "Abstractness, Specificity, and Complexity in Software Design," ACM, pp. 35--42, 2011.

[12] F. Tsui, A. Gharaat, S. Duggins, and E. Jung, "Measuring Levels of Abstraction in Software Development," in SEKE, 2011, pp. 466-469.

[13] B. J. Kirandeep Kaur, Rekha Rani, "Analysis of Gold Plating: A Software Development Risk," International Journal of Computer Science and Communication Engineering, vol. 2, 2013.

[14] F. Medeiro, B. Pérez-Verdú, and A. Rodríguez-Vázquez, Top-down design of high-performance sigma-delta modulators vol. 480: Springer Science & Business Media, 2013.

[15] H. a. v. V. De Bruin, Hans, "Quality-driven software architecture composition," Journal of Systems and Software, Elsevier, vol. 66, pp. 269--284, 2003.

[16] S. Misra, "An approach for the empirical validation of software complexity measures," Acta Polytechnica Hungarica, vol. 8, pp. 141-160, 2011.

[17] A. Lawgali, "TRACEABILITY OF UNIFIED MODELING LANGUAGE DIAGRAMS FROM USE CASE MAPS," International Journal of Software Engineering & Applications (IJSEA), 2017.

[18] A. Zalewski, "Modelling and evaluation of software architectures," Prace Naukowe Politechniki Warszawskiej. Elektronika, 2013.

[19] B. L. M. Montaner, and J. De La, "A Taxonomy of Recommender Agents on the Internet," Artificial Intelligence Review, vol. 19, 2003.

[20] H. Castillo, "Hybrid Content-Based Collaborative-Filtering Music Recommendations," Department of Computer Science, Information System Engineering (ISE), Netherlands, 2007.

[21] J. Itmazi, "Flexible Learning Management System To Support Learning In The Traditional And Open Universities," PhD Thesis , university of Granada, 2005.

[22] J. Obando, "Methodology to obtain the user's Human Values Scale from Smart User Models," PhD Thesis, Department of Electronics, Computer Science and Automatic Control, University of Girona, 2008.

[23] R. Burke, "Hybrid Recommender Systems:Survey and Experiments," User Modeling and User-Adapted Interaction, vol. 12, 2002.

[24] T. T. a. R. Cohen, "Hybrid Recommender Systems for Electronic Commerce," in the 17th National Conference on Artificial Intelligence AAAI, 2000.

[25] R. N. A. A. E. frerjani, "Towards A General Architecture for Building Intelligent, Flexible, and Adaptable Recommender System Based on MAS Technology," post graduation, compluter science, benghazi IEEE journal, 2010.

[26] E. M. Saleh, "Architecture for Design Pattern Selection based on Multi-Agent System," post graduation, benghazi university, 2014.

[27] P. Thitisathienkul and N. Prompoon, "Quality assessment method for software requirements specifications based on document characteristics and its structure," in Trustworthy Systems and Their Applications (TSA), 2015 Second International Conference on, 2015, pp. 51-60.

MULTIPLEXING VOIP PACKETS OVER INTERNET TELEPHONY TRANSPORT PROTOCOL (ITTP)

Mosleh M. Abualhaj[1], Sumaya N. Al-khatib[2] and Mahmoud Baklizi[3]

[1]Department of Networks and Information Security,
Al-Ahliyya Amman University, Amman, Jordan
[2]Department of Software Engineering,
Al-Ahliyya Amman University, Amman, Jordan
[3]Department of Computer Networks Systems,
The World Islamic Sciences & Education University W.I.S.E, Amman, Jordan

ABSTRACT

Voice over IP (VoIP) is a technology of making phone calls over IP network. A considerable number of VoIP applications have emerged in the last decade, to make calls over the Internet. Accordingly, a huge number of VoIP packets are running over the Internet, which consume a considerable share of bandwidth. VoIP applications produce packets with small payload, which cause a considerable header overhead. Whereas, the header overhead is between around 46.4% to 72.2%, depends of the payload size, when using Internet Telephony Transport Protocol (ITTP)/IP protocol. VoIP packets header multiplexing is one of the most common techniques that used to reduce header overhead. In this paper, we propose a new multiplexing technique, called Delta-Multiplexing over ITTP (D-Mux-ITTP). The D-Mux-ITTP is aiming to reduce the wasted bandwidth resulting from the 26-byte ITTP/IP VoIP packet header, by combining the packets destined to the same destination into a single ITTP/IP header.

KEYWORDS

Packets Multiplexing; Bandwidth Utilization; VoIP; D-Mux-ITTP; ITTP.

1. INTRODUCTION

Voice over IP (VoIP) applications have propagated widely in the last decade [1]. In 2017, the size of VoIP packets, running over the internet, are expected to exceed 158 petabytes per month [2]. This noticeable propagation is driven by the fact that the VoIP calls are very cheap or free. In addition, the VoIP calls can be made from anywhere using PC, mobile, tab, and …etc [3,4]. Notwithstanding the foregoing, VoIP still provide calls with less quality than the traditional Public Switched Telephone Network (PSTN). In addition, VoIP waste network bandwidth due to the big packet header size. Whereas, a normal VoIP packet made up of a 40-byte RTP/UDP/IP header, while, the normal VoIP packet payload size is between 10 bytes and 30 bytes [5,6]. To consider this issue, a new protocol called, Internet Telephony Transport Protocol (ITTP), was created to carry VoIP calls. The 6-byte ITTP replaces the 20-byte RTP/UDP header [7,8]. However, the 26-byte ITTP/IP is still considerable header in comparison to the 10 to 30 bytes payload. In this paper, we propose a new VoIP packets multiplexing technique to reduce the wasted bandwidth resulting from the 26-byte ITTP/IP VoIP packet header. In this paper, we

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 75– 81, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80207

propose a new VoIP packets multiplexing technique to reduce the wasted bandwidth resulting from the 26-byte ITTP/IP VoIP packet header.

The rest of this paper is organized as follows: Section 2 discusses the current multiplexing methods of VoIP packets. Section 3 demonstrates the proposed multiplexing method. Section 4 provides a discussion of the expected bandwidth utilization of the proposed method. Finally, Section 5 elaborates the conclusions.

2. RELATED WORKS

VoIP calls are increasingly used over the Internet. A massive number of packets are transmitted over the Internet, which consume a considerable share of bandwidth. There is outstanding effort from researchers to save the bandwidth consumption resulting from VoIP packets. One of the most common techniques is VoIP packet multiplexing. This section discusses some of existing techniques.

In 2001, Sze et al. proposed a VoIP packets multiplexing method over RTP/UDP/IP protocols. The proposed technique combines several VoIP packets transmitted to the same destination together in a single UDP/IP header, while the original RTP header is remained in each packet. However, a compression technique was applied on the RTP header. Implementation of the proposed technique showed that combining both packets multiplexing and RTP header compression succeeded in improving bandwidth utilization, whereas, bandwidth employment improved by 72% [9].

Another multiplexing method that multiplex RTP/UDP/IP packets was proposed by Abualhaj et al. in 2010. The proposed multiplexing method called Delta-multiplexing. In Delta-multiplexing, the packets that are destined to the same destination are assembled together in one UDP/IP header, while the RTP header remains unchanged. In addition, Delta-multiplexing succeeded to compress VoIP packets payload through transmitting the difference of the successive packet payloads. Combining both packet multiplexing and payload compression have shown highly efficient bandwidth exploitation. When implemented, Delta-multiplexing method succeeded to save the bandwidth between 68% and 72%, depending on the VoIP packet payload size [10].
As we can see, the previous two methods were proposed to work with RTP/UDP/IP header. In 2016, M.M.Abualhaj proposed a VoIP packets multiplexing method to work with ITTP/IP protocols. The proposed multiplexing method called ITTP-Mux. In ITTP-Mux, the packets that are destined to the same destination are assembled together in one IP header, while the ITTP header remains unchanged. Implementation of ITTP-Mux method showed high reduction in packet overhead, especially when compared with the traditional ITTP protocol (without multiplexing), which eventually save the bandwidth. The simulation result showed that bandwidth exploitation improved by up to 29.1% in the tested cases [11].

In this paper, we propose a new multiplexing method called Delta-Multiplexing over ITTP (D-Mux-ITTP). D-Mux-ITTP combines between both packet multiplexing as in [11] and packet payload compression as in [11]. The following section discusses the architecture of the proposed method, namely; D-Mux-ITTP method.

3. D-MUX-ITTP ARCHITECHTURE

This section discusses the design of the proposed method, namely; D-Mux-ITTP. The D-Mux-ITTP consists of two components. First is the Sender-Delta-Multiplexing (SDM) component, which reside at the sender side. Second is the Receiver-Delta-DeMultiplexing (RDD) component, which reside at the receiver side. The SDM component works as follows: i) it extract the payload of the packets transmitted to the same destination, ii) the extracted payloads are subtracted from each other, which produces a smaller payload called s-payload iii) the SDM attaches a mini-header to each s-payload, to distinguish these payloads, which produces a small packet called mini-packet, and iv) these mini-packets are combined together in one ITTP/IP packet. Figure 1 demonstrates the SDM component process.

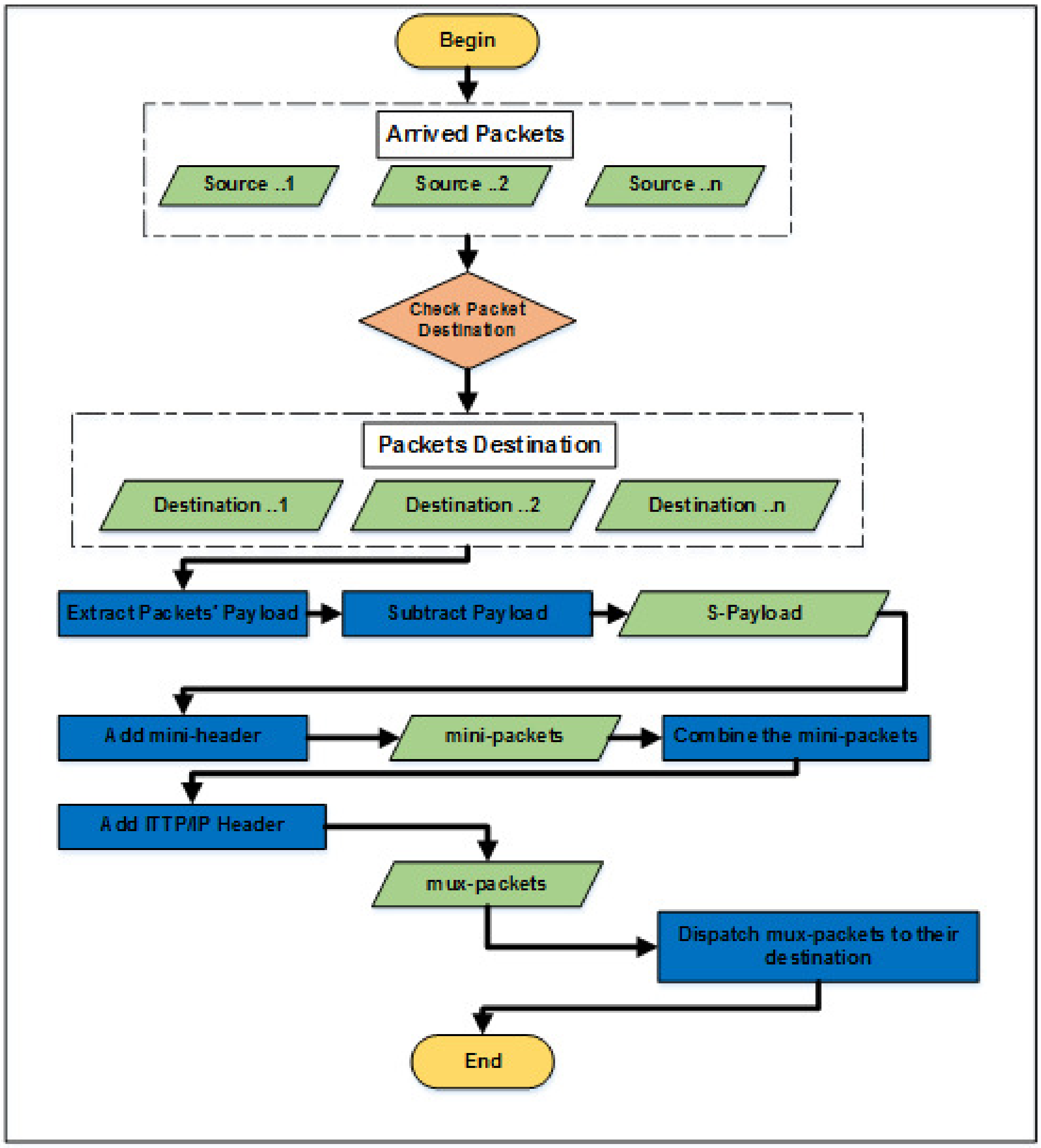


Figure 1. SDM Component Process

The RDD component works as follows: i) it separates the received ITTP/IP packet into mini-packets by inspecting the mini-header ii) it then removes the mini-header from the mini-packet iii) after that it restores the original size of the payload, iv) and finally, the RDM attach the ITTP/IP packet to the restored payload, which re-construct the original packet. Figure 2 demonstrates the RDD component process. Figure 3 demonstrates the D-Mux-ITTP method architecture.

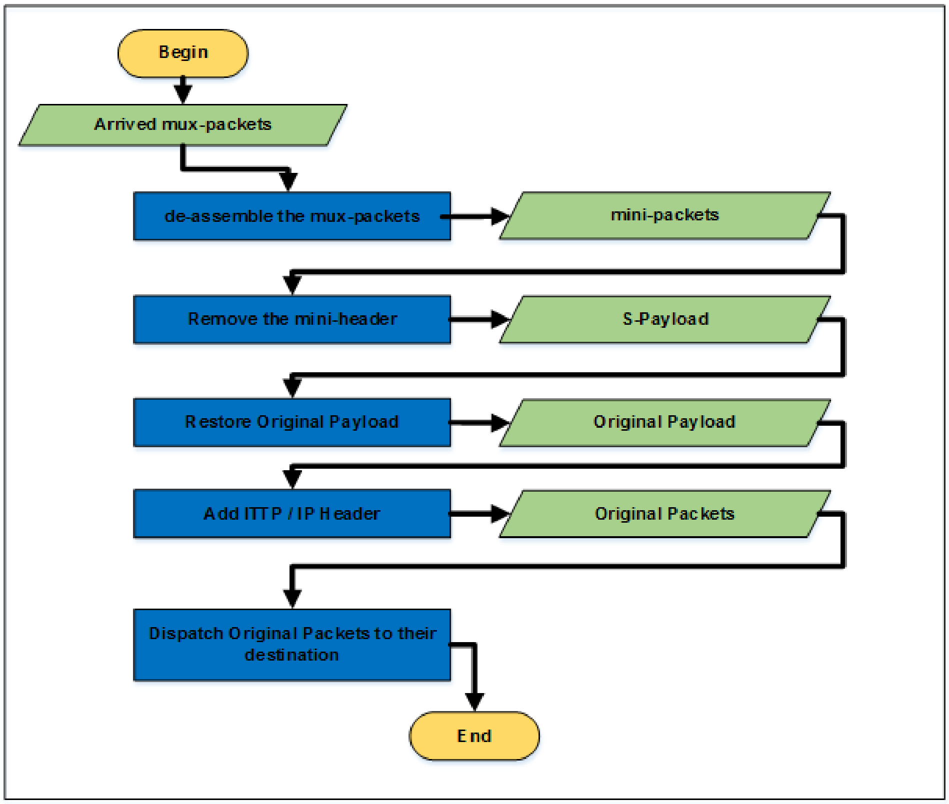


Figure 2. RDD Component Process

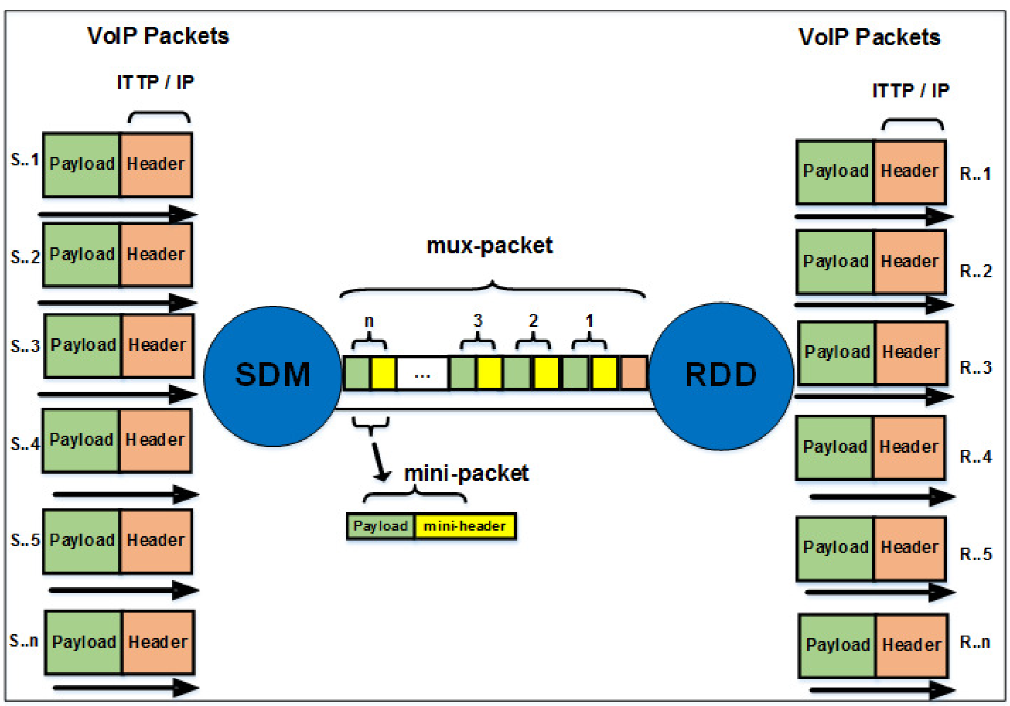


Figure 3. D-Mux-ITTP Method Architecture

The D-Mux-ITTP achieves bandwidth utilization through two phase. In phase1, several VoIP packets from different sources are combined together in one ITTP/IP header. This phase reduces the header overhead resulting from attaching a separate header to each packet. In phase2, the VoIP packets payload, before get combined together, are assumed to be integer number and subtracted from each other. This phase reduces the payload size through transmit the difference between the payloads instead of the full payloads size.

4. D-MUX-ITTP BRANDWIDTH UTILIZATION EFFICIENCY

This section discusses the D-Mux-ITTP bandwidth utilization efficiency in comparison to the traditional method (no multiplexing). The traditional VoIP packet payload size between 10 and 30 bytes [5, 6]. Accordingly, the header overhead resulting from the 26-bytes ITTP/IP header, which is the relative ratio between the header size and the packet size, is between around 46.4% to 72.2%. Figure 4 demonstrates the header overhead ratio with different payload sizes.

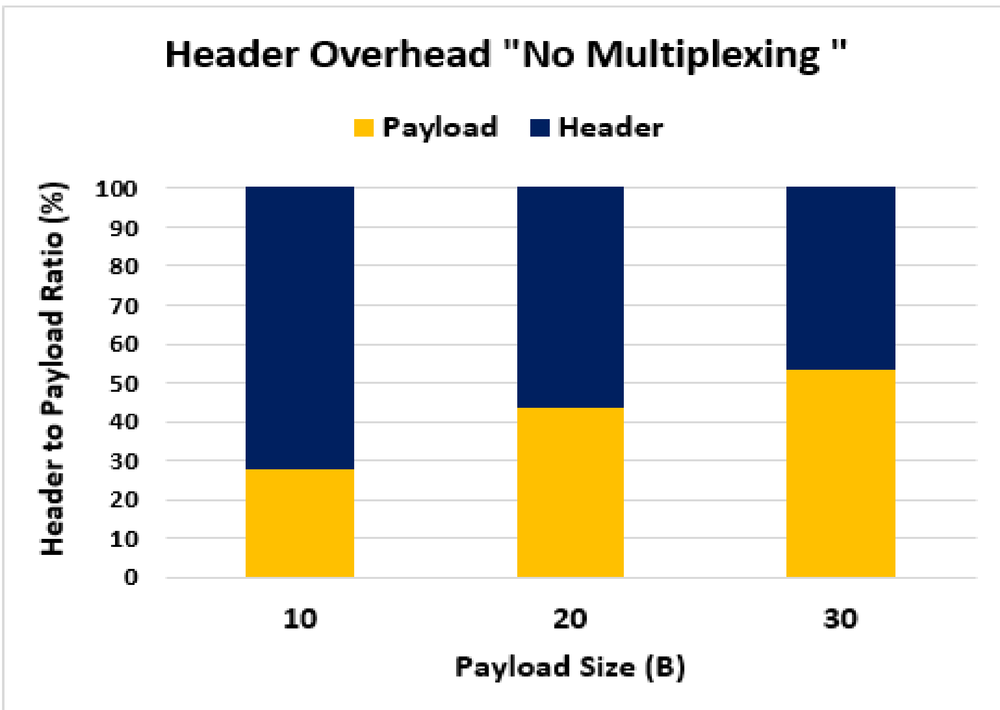


Figure 4. Header Overhead Ratio (No Multiplexing)

The aim of the D-Mux-ITTP method is to reduce the wasted bandwidth resulting from the considerable ITTP/IP header overhead. D-Mux-ITTP method accomplish this by multiplexing several VoIP packets payload in a single ITTP/IP header, which reduces the ITTP/IP header overhead. This is efficiently improve bandwidth utilization, depending on the number of multiplexed packets. Figure 5 demonstrates the header overhead ratio when multiplexing several packets in single ITTP/IP header, assuming packet's payload size is 20-byte. Note that Figure is only for demonstration purpose and the numbers are approximated, based on the relative ratio between the header size and the multiplexed packets size. In future work, we will implement the D-Mux-ITTP method and provide an accurate values of the bandwidth utilization efficiency when using it.

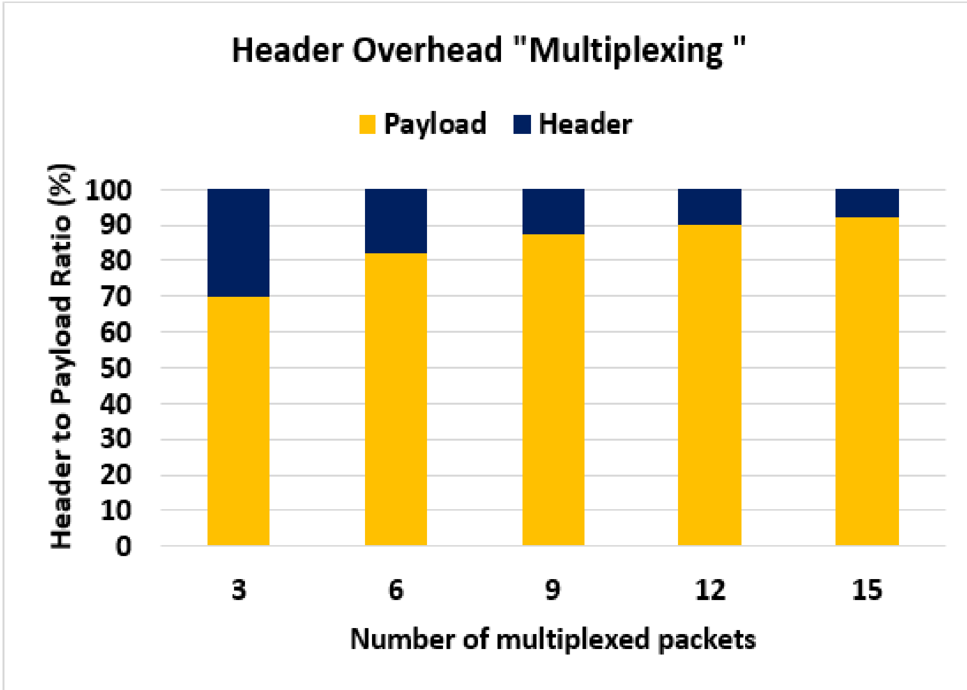


Figure 5. Header Overhead Ratio (Multiplexing)

5. CONCLUSIONS

In this paper, we proposed a new multiplexing method called D-Mux-ITTP. The aim of the D-Mux-ITTP method is to efficiently improve bandwidth exploitation, by reducing the wasted bandwidth resulting from the considerable ITTP/IP header overhead. The D-Mux-ITTP achieves this by i) combining several VoIP packets payloads in a single ITTP/IP header and ii) by transmit the difference between the payloads instead of the full payloads size. The D-Mux-ITTP consists of SDM and RDD entities. SDM performs packets multiplexing and the sender side, while, RDD performs packets de-multiplexing and the receiver side. Future work will implement the D-Mux-ITTP method and provide a detailed discussion and analysis of the D-Mux-ITTP bandwidth utilization, in terms of header overhead, saved bandwidth, call capacity, and goodput. In addition, the impact of the D-Mux-ITTP on the network performance and call quality will be provided. Network performance and call quality will be measured through delay, overload, congestion, and …etc.

REFERENCES

[1] MM Abualhaj, M Kolhar, K Qaddoum, AA Abu-Shareha, (2016) “Multiplexing VoIP Packets over Wireless Mesh Networks: A Survey, KSII Transactions on Internet & Information Systems”, Vol. 10, No.8.

[2] http://www.statista.com/statistics/267183/forecast-for-the-worldwide-voip-traffic/, November-2017.

[3] AbuAlhaj, Mosleh, Manjur S. Kolhar, M. Halaiyqah, O. Abouabdalla and R. Sureswaran, (2009) ”Multiplexing SIP applications voice packets between SWVG gateways”, in Proc. of International Conference on Computer Engineering and Applications (ICCEA 2009).

[4] Manjur, S. K., Mosleh Abu-Alhaj, Omar Abouabdalla, Tat-Chee Wan, Rahmat Budiarto and Ahmed M. Manasrah, (2011), "Conference gateway for heterogeneous clients: Real time switching clients and interasterisk exchange clients", International Journal of Innovative Computing Information and Control 7, No. 1, 395-406.

[5] MM Abu-Alhaj, A Manasrah, M Baklizi, N Abdullah, (2011) "Transport layer protocols taxonomy from Voice over IP perspective", Advanced Computing: An International Journal (ACIJ), Vol.2, No.4.

[6] M Abu-Alhaj, MS Kolhar, M Halaiyqah, O Abouabdalla, (2009) " MuxComp-A New Architecture to Improve VoIP Bandwidth Utilization", International Conference on Future Networks.

[7] Abu-Alhaj, Mosleh M., S. K. Manjur, R. Sureswaran, Tat-Chee Wan, Imad J. Mohamad and Ahmed M. Manasrah, (2012) "ITTP: A New Transport Protocol for VoIP Applications", International Journal of Innovative Computing, Information and Control (IJICIC), Vol. 8, No. 3(A), pp. 1879–1895.

[8] A Kayed, MM Abualhaj, M Alharibat, (2013),"PERFORMANCE COMPARISON OF IAX AND ITTP VoIP PROTOCOLS", International Journal of Academic Research, Vol. 5, No. 2.

[9] Sze, H. P., Soung C. Liew, Jack YB Lee and Danny Yip, (2002), "A multiplexing scheme for H. 323 voice-over-IP applications", Selected Areas in Communications, IEEE Journal, No. 7, pp. 1360-1368.

[10] Abu-Alhaj, Mosleh M., (2015)," ITTP-MUX:An efficient multiplexing mechanism to improve voip applications bandwidth utilization", International Journal of Innovative Computing, Information and Control (IJICIC), Vol. 11, No. 6, pp. 1–1503-0010.

[11] Abu-Alhaj, Mosleh M., Manjur S. Kolhar, Lingeswari V. Chandra, O. Abouabdalla and Ahmed M. Manasrah, (2010), "Delta-Multiplexing: A Novel Technique to Improve VoIP".

INTENTIONAL BLANK

APPLICATION OF DYNAMIC CLUSTERING ALGORITHM IN MEDICAL SURVEILLANCE

Zhuohui Ren and Cong Wang

Department of Software Engineering, Beijing University of Posts and Telecommunications, BeiJing City, China

ABSTRACT

The traditional medical analysis is based on the static data, the medical data is about to be analysis after the collection of these data sets is completed, but this is far from satisfying the actual demand. Large amounts of medical data are generated in real time, so that real-time analysis can yield more value. This paper introduces the design of the Sentinel which can realize the real-time analysis system based on the clustering algorithm. Sentinel can realize clustering analysis of real-time data based on the clustering algorithm and issue an early alert.

KEYWORDS

Algorithms, Data Mining , Cluster, Data stream, Medical

1. INTRODUCTION

With the arrival of big data era, Medical big data has gradually entered the people's vision, Medical big data refers to all the big data related to medical and life health. According to the source of medical big data can be broadly divided into biological big data, clinical big data and health big data[1].This The potential value of medical data is enormous. For example, public health departments can conduct comprehensive disease surveillance in the monitoring of infectious diseases through a nationwide electronic medical record database, and analysis the characteristics of the spread of illness through data mining.

In the field of health care, most of the data can be seen as streaming data, such as out-patient records, electronic medical records and so on. These data increase by the of time and the numbers of people, It has the characteristic of continuity. Because of its real-time nature, it plays an important role in disease surveillance. For example, mining of outpatient records can dynamically detect diseases that increase in a large amount over a certain period of time, for example, sudden infectious diseases or collective poisoning. Unlike traditional databases that contain static data, data stream are inherently continuous and unconstrained, and there are many problems when working with such data. In addition, the result of data analyse is very unstable and constantly generating new patterns. Static pattern mining techniques proved inefficient when working with data streams. With the deepening of information technology in the medical field, the ability of generating data is rapidly increasing. Mining useful information from these streams has become an inevitable task.

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 83– 86, 2018. DOI : 10.5121/csit.2018.80208

2. RELATED WORK

Sudipto Guha proposed a clustering algorithm based on stream data [2].In his algorithm, the idea of divide and conquer is adopted, the data flow is divided into multiple segments, and the segments is separately clustered to obtain the first cluster center. When the first cluster center reaches a certain number, the second segment of data is introduced to cluster to get the second cluster center. As data continues to flow in, this process will continue. At each time point, the system only needs to maintain m i-th layer center points. This division of mind is very efficient for the analysis of streaming data. Since only a limited number of data needs to be saved at each time point in the system, the storage and memory shortage due to the large incoming stream data are avoided. Because the stream data analysis is a dynamic process, Most of the algorithms are based on the needs of the application to choose the time as a standard, select a period of time to analysis. According to the selected timing range can be divided into snapshot model, landmark model And sliding window model. Landmark model and sliding window model are more used.

As an important algorithm in data mining, the main goal of clustering is to classify the internal relations between data into a large category and distinguish each category as much as possible, which is an extension in taxonomy. According to the different basic principles of clustering can be divided into, division clustering, hierarchical clustering ,density-based clustering, model-based clustering and grid-based clustering [3].

With the extension time or space will produce a wide range of data, and data mining is to extract valuable information from these complex types of data. These complex types of data can be divided into spatial data, timing data, web data, text data [4]. From its process, dynamic data mining can be divided into several stages such as dynamic data collection, data processing, data mining, and mining evaluation [5]. In general, data mining and mining evaluation are closely integrated. Dynamic data mining needs better handling of real-time data and the impact of real-time data on analysis results. The main problems of k-mean algorithm in dealing with dynamic data mining are as follows: Since the initial value of k is fixed means that it can not be changed after it is selected, that makes k-means algorithm unsuitable for mining of dynamic data. Therefore, in the k-means algorithm for dynamic data clustering algorithm, the improvement mainly focuses on the selection and dynamic adjustment of the k value [6], which can be mainly divided into two directions:1, in the process of dynamic data acquisition of data preprocessing, according to the predetermined strategy to adjust the size of k;2 In the data mining process according to the data mining results and predetermined criteria, the data results are dynamically adjusted, and then update k value. The difference between the two methods is that the former is adjusted in the data processing stage and the latter is adjusted in the data mining stage.

The algorithms based on the first idea are: K-means clustering algorithm based on KD tree[6]. The KD tree represents a k-dimensional storage structure that stores data separately at each node in a well-spaced space. Since the initial cluster centers in the k-means algorithm are randomly selected, they can not reflect the true distribution of the data. In order to distribute the actual reaction data as much as possible, it is better to distribute the initial center points more evenly. The basic idea of clustering using KD tree is as follows: Firstly, the advantages of KD tree are used to divide the spatial extent of data set and the data of the corresponding interval is stored. This will effectively improve the effect of the initial center point selection. Using KD tree to divide the space and preprocess the data, we can know the distribution of the data truly. Then according to the partitioned interval, the initial center point is chosen directionally. Finally, the clustering operation is carried out. The algorithm can better find the k-value and the clustering center point, but the computational cost is larger when the clustering operation is re-performed. Compared with the first method of dynamically adjusting the k value, the second method is based on The computation overhead caused by the local dynamic adjustment of the clustering result and the result evaluation index will be significantly reduced. For dynamic data sets, it is obviously inefficient to re-execute the clustering algorithm on the updated new data set to update the clustering results accordingly, so it is very important to adopt incremental clustering algorithm effectively .

Among them, the algorithm based on the second idea has a two-point k-means clustering algorithm [7]. The main idea is to adjust the clustering result locally instead of the global adjustment according to the threshold in the process of data clustering, which can effectively improve the efficiency of the algorithm and does not affect the final result of the clustering.

3. PROBLEM SETUP

The main problem in the processing of streaming data is that streaming data is a sequence of data sequences that is massively and continuously arriving[8]. When the clustering algorithm is applied to streaming data, it is mainly necessary to consider the real-time performance and the scale unpredictability of the streaming data.
Sentinel's main process is as following steps:

Step 1. Monitoring data cache, if the cache data to meet the conditions to step two.

Step 2. Cached data submitted to the data analysis module, analysis module is used for data analysis, and based on the results, update corresponding parameters.

Step 3. Data early warning module to update the data to determine, greater than the predetermined value issued a alert.

Step 4. Return to step one.

Data caching is mainly for real-time streaming data processing, according to Sudipto Guha' algorithms in the treatment of streaming data, the idea of data segmentation processing, application cache technology can be very good to achieve this. The process of caching is to segment the data base on time line. Data analysis moudle is mainly based on dynamic clustering algorithm. In the data analysis of a block need to use the relevant information in the database, the information is the system needs long-term maintenance. The content of this information includes the number of clusters, the center of each cluster, and the data set that belongs to each cluster. The data processing flow train of thought is as following steps:

Step 1 According to the cluster center stored in the system, the data will be assigned to the corresponding cluster.

Step 2 pairs of clusters of data are calculated and compared with the threshold, according to the comparison result,then adjust of clusters.

Step 3 According to the results of the adjustment, update the relevant records in the database.

In the local adjustment of the cluster, the main reference is the intra-cluster similarity and inter-cluster similarity. The inter-cluster similarity is defined as the mean of the data in the cluster. The similarity between clusters is defined as the distance between the centers of two adjacent clusters. If the similarity between clusters in a cluster is greater than the threshold, the k-means algorithm for k = 2 is performed on the cluster[9]. If the cluster similarity between two clusters is greater than the threshold, the two clusters are merged.

4. CONCLUSION AND FUTURE WORK

There are some places in the system design that deserve further study, mainly for setting the threshold of division and consolidation. The setting of the threshold determines the quality of the splitting and merging[10]. At the same time, the setting of the threshold has a great relationship with the selection of data types. How to find out a suitable model to adapt the model to more types of The data set will be very necessary. Relevant researchers can conduct in-depth research based on different subjects in medical data, and build a better model to make the algorithm better adapt to various data mining.

REFERENCES

[1] Meng Qun, Bi Dan, Zhang Yiming et al .Chinese Journal of Health Information Management, 2016, 13 (6): 547-552..

[2] Sudipto Guha. Asymmetric k-center is log^*n-hard to Approximate[J]. Journal of the Acm, 2013, 52(4):538-551.

[3] Wang Juan.Study on Evolutionary Clustering Algorithm in Dynamic Data Mining [D]. Nanjing University of Aeronautics and Astronautics, 2012.

[4] Zhang Yufeng, Zeng Yitang, Hao Yan.Study on Intelligent Strategy of Logistics Information Based on Dynamic Data Mining [J] .Library Science, 2016 (5): 46-49.

[5] Wang Lunwen, Feng Yanqing, Zhang Ling.A Review of Constructive Learning Methods for Dynamic Data Mining [J] .Microsoft Microcomputer Systems, 2016, 37 (9): 1953-1958.

[6] Tang C, Ling C X, Zhou X, et al. Proceedings of the 4th international conference on Advanced Data Mining and Applications[C]// International Conference on Advanced Data Mining and Applications. Springer-Verlag, 2008:15-15.

[7] Lunwen Wang, Yanqing Feng, Ling Zhang.A Review of Constructive Learning Methods for Dynamic Data Mining [J] .Microsoft Microcomputer Systems, 2016, 37 (9): 1953-1958.

[8] Xiujin Shi, Yanling Hu,et al.Privacy Protection of Dynamic Set-valued Data Publishing Based on Classification Tree [J] .Computer Science, 2017, 44 (5): 120-124.

[9] Guangcong Liu,Tingting Huang,Haiin Chen,et al.An improved dichotomous K-means clustering algorithm [J].Computer Applications and Software, 2015 (2): 261-263.

[10] Zhu Y T, Wang F Z, Shan X H, et al. K-medoids clustering based on MapReduce and optimal search of medoids[C]// International Conference on Computer Science & Education. IEEE, 2014:573-577.

AUTHORS

Ren Zhuohui, Male, 1989, renzh@bupt.edu.cn, Master's degree of Beijing University of Posts and Telecommunications. Main research areas include privacy protection and data mining

Wang Cong, Female, 1958, Professor and doctoral tutor at Beijing University of Posts and Telecommunications. Main research areas include intelligent control and Wisdom-Web information security.

AN INVESTIGATION OF WATERMARKING MEDICAL IMAGES

Majdi Al-qdah[1] and Anas Alqudah[2]

[1]Department of Computer Engineering, University of Tabuk, Tabuk City, KSA
[2]Department of Finance, College of Business, University of Abu Dhabi, UAE

ABSTRACT

This paper presents the results of watermarking selected various medical cover images with simple string of letters image (patients' medical data) using a combination of the Discrete Wavelet Transform (DWT) Discrete Cosine Transform (DCT) and singular value decomposition (SVD). The visual quality of the watermarked images (before and after attacks) was analyzed using PSNR and four visual quality metrics (WSNR, MSSIM, PSNR-HVS-M, and PSNR-HVS). The PSNR, PSNR-HVS-M, PSNR-HVS, and WSNR average values of the watermarked medical images before attacks were about the 32 db, 35 db, and 42 db, 40 db respectively; while the MSSM index indicated a similarity of more than 97% between the original and watermarked images. The metric values decreased significantly after attacking the images with various operations but the watermark image could be retrieved after almost all attacks. Thus, the initial results indicate that watermarking medical images with the patients' data does not significantly affect their visual quality and they still can be utilized for their medical purpose.

KEYWORDS

Watermarking, medical images, DWT, DCT, SVD, visual metrics

1. INTRODUCTION

There is an increasing need to store and transfer medical images over computer networks for sharing among doctors. Data hiding has increasingly become an important tool in authentication of images and protection of owners copyright. Image watermarking, which hide important details inside images techniques can be divided into two broad domains: spatial domain and frequency domain [1, 2]. Three of the most important frequency watermarking methods are the discrete cosine transform (DCT), discrete wavelet transform (DWT) and Singular Value Decomposition (SVD). Various medical images based watermarking schemes have been proposed in literature [3,4,5]. Many researchers have used a hybrid of two or more transforms in order to compensate for the shortcomings of various transforms.

2. METHODOLOGY

2.1. Watermarking algorithms

This paper uses a combined approach of the discrete wavelet transform (DWT), the discrete cosine transform (DCT), and the singular value decomposition (SVD) watermarking. The DWT decomposes an image into frequency channels of constant bandwidth on a logarithmic scale by separating an image into a set of four non-overlapping multi-resolution sub bands denoted as

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 87– 95, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80209

lower resolution approximation image (LL), horizontal (HL), vertical (LH) and diagonal (HH) with the availability of multiple scale wavelet decomposition. The watermark is usually embedded into the high frequency detail sub-bands (HL, LH and HH sub-band) because the human visual system (HVS) is sensitive to the low-frequency LL part of the image. In general, sensitive data such as medical information are embedded in higher level sub-bands since the detail levels carry most of the energy of the image [6]. Wavelet transform methods achieve higher robustness since they have the characteristics of space frequency localization, multi-resolution representation, multi-scale analysis, adaptability and linear complexity [7].

The DCT has a very good energy compaction property. It works by separating the image into different low, high, and middle frequency coefficients [8]. The watermark is embedded in the middle frequency band that gives additional resistance to the lossy compression techniques with less modification of the cover image. The DCT coefficients D(i, j) matrix of an image (N x M) with pixel intensity I(x, y) are obtained as follows:

$$D(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} I(x,y) \cos\left(\frac{(2x+1)i\pi}{2N}\right) \cos\left(\frac{(2y+1)i\pi}{2M}\right) \quad \text{.......................(Eq.1)}$$

$$C(i), C(j) = \frac{1}{\sqrt{N}}, \frac{1}{\sqrt{M}} \text{ for } i,j = 0 \text{ and } C(i), C(j) = \sqrt{\frac{2}{N}}, \sqrt{\frac{2}{M}} \text{ for } i,j = 1,2,\ldots\ldots N-1 \text{ or } M-1 \quad \text{.........(Eq.2)}$$

The (SVD) of a rectangular matrix R_m is a decomposition of the form

$$R_m = USV^T \quad \text{...(Eq.3)}$$

Where R_m is a M x N matrix, U and V are orthonormal matrices, and S is a diagonal matrix comprised of singular values of R_m. The singular values $S_1 \geq S_2 \geq S_3 \geq \ldots\ldots\ldots\ldots S_{n-1} \geq S_n \geq 0$ are unique values that appear in descending order along the main diagonal of S. They are obtained by taking the square root of the Eigen values of $R_m R_m^T$ and $R_m^T R_m$. The U, V are not unique. In the Singular Value Decomposition, the slight variations of singular values do not affect the visual perception of the cover image, which achieves better quality of the watermarked image and better robustness against attacks. Also, singular values represent the intrinsic algebraic image properties [8].

Various medical images based watermarking schemes have been proposed in literature [9.10,11]. Many researchers have used a hybrid of two or more transforms in order to compensate for the shortcomings of various transforms; i.e. in image compression [12], image denoising [13], or image coding [14], and watermarking [15]. In this work, a combined approach of the three transforms is used for watermarking: DWT, DCT, and SVD. The combination of the three transforms increases the robustness and imperceptibility of the watermarked images [16]. Figure 1 shows the approach taken in embedding a watermark (patients data) into a cover image (medical image); The singular values of the watermark (after DCT transformed) are embedded in the singular values of the cover image (after DWT transformed). Figure 2 shows the extraction approach of the patient's image data from the watermarked image. The watermarked images is DWT and DCT transformed then SVD is applied to the DCT coefficients; the watermark is extracted from the LL sub band of DWT. For an added security, the watermark image can be encrypted before embedding it in the cover image, which is not utilized in this paper.

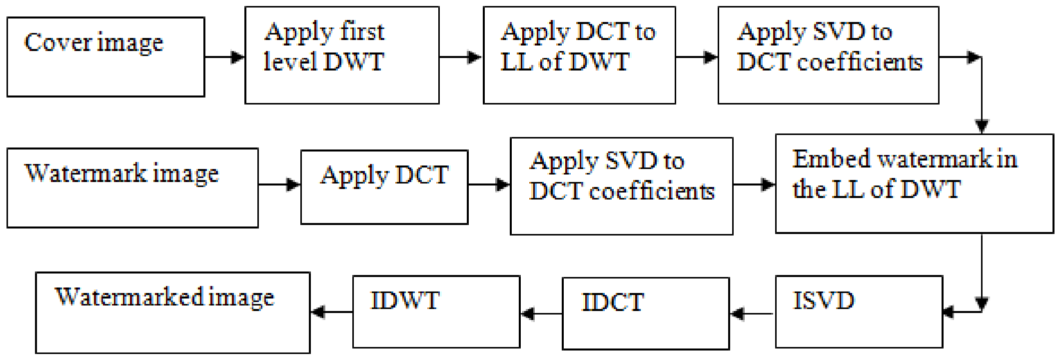


Figure1. Embedding process

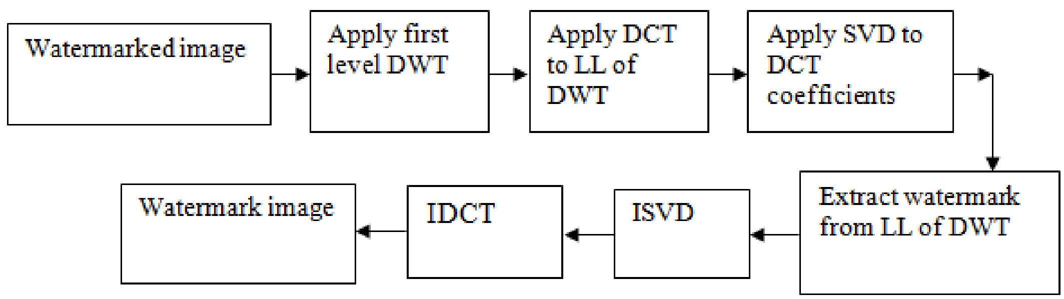


Figure 2. Extraction process

2.2. Performance Measures

This work utilizes the visual metrics (WSNR, MSSIM, PSNR-HVS-M, and PSNR-HVS) described by Ponomarenko et. al. [17] for comparing the watermarked images with their originals. Traditionally, the efficiency of an image processing operation ; i.e. lossy compression is usually analyzed in terms of rate-distortion curves. These curves represent dependencies of PSNR (or MSE) on bits per pixel (bpp) or compression ratio (CR) where PSNR and MSE are calculated for some original image and the corresponding processed image.

$$PSNR = 10\log\left(\frac{255^2}{MSE}\right) \qquad \text{(Eq.4)}$$

$$MSE = \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{(I_{ij}^{d}-I_{ij}^{n})^2}{NM} \qquad \text{(Eq.5)}$$

where I_{ij}^{d}, I_{ij}^{n} denote the values of the original and processed pixels and N, M denote an image size [18,17]. In order to obtain a high imperceptibility of the watermarked image, it is desirable to have a high value of PSNR; meaning a lesser value of MSE.

Also, usually the similarity and differences between an original image and a processed image is measured by the Normalized Correlation (NC). Its value is generally 0 to 1. Ideally it should be 1 but a value 0.7 or higher is usually acceptable [13].

$$NC = \sum_{i=1}^{X}\sum_{j=1}^{Y}(I_{ij}^{d} \times I_{ij}^{n}) \,/\, \sum_{i=1}^{X}\sum_{j=1}^{Y}(I_{ij}^{d})^2 \qquad \text{(Eq6)[19]}$$

where I_{ij}^{d}, I_{ij}^{n} denote the values of the original and processed pixels and X, Y denote an image size.

On the other hand, in evaluating the quality of the watermarked images, it is well known that conventional quality metrics, such as MSE, SNR and PSNR do not always correlate with image visual quality [20,21]. In other words, two different distorted images with the same value of PSNR with respect to the same original image, may give significantly different visual impact. Therefore, the choice of a proper visual quality metric for analysis and comparisons is always a problem and can be argued since the human visual system (HVS) is nonlinear and it is very sensitive to contrast changes and to noise [22]. Many studies have confirmed that the HVS is more sensitive to low frequency distortions rather than high frequency components. The best performance was achieved by the metrics PSNR-HVS-M, PSNR-HVS, and WSNR [18]especially if there is noise or the images are to be compressed. HVS-based models are the result of trade-off between computational feasibility and accuracy of the model. HVS-based models can be classified into two categories: neurobiological models and models based on the psychophysical properties of human vision. Psychophysical HVS-based models are implemented in a sequential process that includes luminance masking, colour perception analysis, frequency selection, and contrast sensitivity [22].

More recent ways to evaluate processing of images is by using perceptual image quality assessment methods, which attempt to simulate the functionality of the relevant early human visual system (HVS) components. These methods usually involve a pre-processing process that may include image alignment, point-wise nonlinear transform, low-pass filtering that simulates eye optics, and color space transformation, a channel decomposition process that transforms the image signals into different spatial frequency as well as orientation selective subbands, an error normalization process that weights the error signal in each subband by incorporating the variation of visual sensitivity in different subbands, and the variation of visual error sensitivity caused by intra- or inter-channel neighbouring transform coefficients, and an error pooling process that combines the error signals in different subbands into a single quality/distortion value [23].

PSNR-HVS takes into account the HVS properties such as sensitivity to contrast change and sensitivity to low frequency distortions; while the PSNR-HVSM takes into account the contrast sensitivity function (CSF). Similar to PSNR and MSE, the visual quality metrics PSNR-HVS and PSNR-HVSM can be determined:

$$\mathrm{PSNR} - \mathrm{HVS} = 10 \log \frac{255^2}{\mathrm{MSE}_{nf}^{HVS}} \quad \text{(Eq.7)}$$

$$\mathrm{MSE}_{nf}^{HVS} = K \sum_{i=1}^{I-7} \sum_{j=1}^{J-7} \sum_{m=1}^{8} \sum_{n=1}^{8} \left((X[m,n]_{i,j} - X[m,n]_{ij}^{e}) T_c[m,n] \right) \quad \text{(Eq.8)[24]}$$

where I,J denote image size, K=1 [(I-7)(J-7)64] , $X[m,n]_{i,j}$ are DCT coefficients of 8x8 image block for which the coordinates of its left upper corner are equal to i and j, Xij e are the DCT coefficients of the corresponding block in the original image, and $T_c[m,n]$ is the matrix of correcting factors [24].

The Weighted Signal to Noise Ratio (WSNR) is a noise metric where the difference (residual) between the original and the processed images must be noise. (WSNR) uses a Contrast Sensitivity Function (CSF) given by the following:

$$\mathrm{CSF} = 2.6(0.02 + 0.1 f_a) e^{-(0.1 f_a)^{1.1}} \quad \text{(Eq.9)}$$

where f_a is a radial angular frequency

The WSNR between an original image (x) and a processed image (y) is:

$$\mathrm{WSNR} = 10\log_{10}\left(\frac{\sum |\mathrm{CSF} \times X(u,v)|^2}{\sum |\mathrm{CSF} \times X(u,v) - Y(u,v))|^2}\right) \quad \text{(Eq.10)}$$

The structural similarity index (SSIM) measures the similarity between two images [19]. SSIM compares two images using information about luminous, contrast and structure. SSIM metric is calculated on various windows of an image. The measure between two windows x and y of common size N×N is given as follows:

$$SSIM(x,y) = \frac{\{(2u_x u_y + C_1)(2\sigma_{xy} + C_2)\}}{\{(u_x^2 + u_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)\}} \quad \text{(Eq.11)}$$

SSIM takes values between 1 and -1; u_x is an average of x, u_y is an average of y, σ_x, σ_y are the standard deviations between the original and watermarked image pixels; while C_1, C_2 are positive constant chosen to avoid the instability of SSIM measure.

MSSIM (Multi-Scale Structural Similarity) is a multi-scale extension of a SSIM metric. MSSIM [28] is introduced to incorporate the variations of viewing conditions to the previous single-scale SSIM measure. MSSIM is known as mean structural similarity index metric [25] and it is given by:

$$MSSIM(x,y) = \frac{1}{M}\sum_{i=1}^{M} SSIM(x_i, y_i) \quad \text{(Eq.12)}$$

where M is the correlation between two images x, y

Correlation is a similarity measure between two functions. The correlation measure between two functions x(x,y) and s(x,y) in discrete form is defined as:

$$M = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} r(x,y)[s(x,y)]^* \quad \text{(Eq.13)}$$

Where $[]^*$ is the complex conjugate, x=0, 1,…….., M-1 and y=0, 1,……, N-1

3. RESULTS AND DISCUSSIONS

Five medical cover images of size [512×512] and a watermark image of size [256×256] are selected for analysis shown in Figure 3. The medical cover images contain medical information based on the characteristics of each image and the purpose of its capture. The medical images reveal characteristics of the bones, tissues, vessels, nerves....etc. Thus, embedding a watermark image inside a medical cover image should preserve the existing medical information in the cover medical image: the unique pattern of the fingerprint, vessels and optical nerves inside the retina, bone fracture in the wrist, size and development signs of the fetus, shape, and sliced layers and soft tissue of the human skull. The patients' personal details can be embedded in the captured medical image in textual or image format and saved in one file. The personal details (watermark) are embedded by a combined method of DWT, DCT, and SVD transforms; while the imperceptivity of the watermarked images is evaluated using PSNR, P-HVS, P-HVS-M, WSNR, and MSSIM.

 Retina	 Broken wrist	 Fingerprint

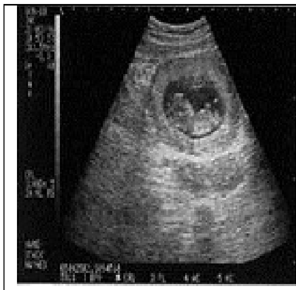
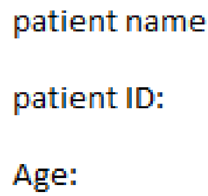
 Ultrasound	 MRIofHead	 Watermark

Figure 3. Images and Watermark

Table 1 shows the values of PSNR, P-HVS, P-HVS-M, WSNR, and MSSIM metrics of the watermarked images before any attacks. The PSNR average value is about 32 db, P-HVS average value is around 35 db, P-HVS-M average value is about 42 db, and the WSNR average value varies from 35 db to 47 db. The MSSIM metric shows that the watermarked images are highly visually similar to the original images with similarity values of more than 0.97% between the original and the watermarked images. Also, it can be observed that there is no significant difference between the average metric values among the various images; only the WSNR value of the MRIofHead image varies from one image to another with approximately 15 db difference between the Fingerprints image and the MRIofHead image; that is due to the characteristics of the two images.

Table 1. Metric values of the watermarked images

Image	**PSNR**	**P-HVS**	**P-HVS-M**	**WSNR**	**MSSIM**
Fingerprints	32.7049	34.8745	46.2079	47.0602	0.9920
Retina	32.9101	34.8738	40.4924	38.0317	0.9740
Broken Wrist	32.7310	34.9020	40.7815	43.3029	0.9734
Ultrasound	33.2059	34.8428	41.3834	37.8052	0.9850
MRIofHead	33.3870	35.1103	40.0242	34.3916	0.9770

To test the imperceptivity of the watermarked image, they were attacked with various types of attacks. Tables 2 shows the average values of the same metrics for all the image after the watermarked images are attacked with various operations (Gaussian noise, Salt & Pepper noise, 2D FIR filter, Cropping, Rotation & Cropping, Weiner filter, Intensity adjustment, Gaussian filter, and Sharpening). It is shown that the numerical values decrease after an attack operation is performed on the images. Thus, there is a degradation in the quality of the attacked images. The drop in the numerical values is not significant after the Gaussian Noise, Salt & Pepper Noise, and 2D FIR filter attacks. The values of PSNR, P-HVS, P-HVS-M, and WSNR stay above the value of 20 db and the MSSIM metric values remain above 0.82%. On the other hand, there is a significant decrease in the values after the Cropping, Rotation & Cropping, Weiner Filter, Intensity adjustment, Gaussian filter, and Sharpening image attack operations. The numerical values of PSNR, P-HVS, P-HVS-M, and WSNR drop to less than 6 db while the MSSIM similarity index drops to 10% approximately. On the other hand, there is no correlation between the drop in the metric values and the recovery of the watermark; for example, the P-HVS, P-HVS-M, and the WSNR values drop greatly after the sharpening attack but the watermark is recovered perfectly.

Table 2. Average metric values of all five watermarked images after some attacks

Attack	PSNR	P-HVS	P-HVS-M	WSNR	MSSIM
No attack	32.9878	34.9207	41.7779	40.1183	0.9803
Gaussian Noise	19.9103	19.9790	22.6101	27.0916	0.8212
Salt&Pepper Noise	24.6345	24.8935	27.9674	32.1470	0.9304
2D FIR filter	25.3646	26.6690	30.0951	35.1960	0.9618
Cropping	13.7011	9.5336	9.5680	8.1108	0.7391
Rotation&Cropping	5.9136	1.7654	1.7872	0.2728	0.0982
Weiner Filter	5.9212	1.7732	1.7950	0.2801	0.1029
Intensity adjustment	5.9411	1.7932	1.8150	0.3001	0.1113
Gaussian filter	5.9212	1.7733	1.7950	0.2801	0.1030
Sharpening	5.9214	1.7733	1.7951	0.2801	0.1031

.

Finally, this research cannot determine how much of medical information is lost after watermarking medical images or even after attacking the images with a watermark image. Only medical doctors can decide the important segments of a medical image that are affected by watermarking or by attacks; and the effects can vary from one image to another. Also, recovering the watermark after some attacks does not necessarily indicate that all medical information is preserved in the cover image.

4. CONCLUSIONS

The initial results show that watermarking medical images with the patients' personal details does not significantly affect their visual appearance and they can be used by medical staff for their medical purpose if the watermarked images are not attacked; it was experimentally demonstrated that the watermarked medical images appeared similar to their originals and the Human Visual System (HVS) metrics proved a high quality watermarked images. Also, choosing the appropriate watermarking algorithm is essential to obtain the robustness, imperceptivity and security needed to protect the patients' personal data inside a medical image. There are many transform domain algorithms that are available and can be utilized to preserve the characteristics of the original images. A future direction of this research will involve artificial intelligence methods to watermark the images.

ACKNOWLEDGMENT

The authors would like to acknowledge financial support of this work from the Deanship of Scientific Research (DSR), University of Tabuk, Tabuk, Saudi Arabia, under grant no. S/0180/1438

REFERENCES

[1] Ashourian (2006), A new mixed spatial domain watermarking of three dimensional triangle mesh, proceeding of the 4th international conference on computer graphics and interactive techniques in Australia and Southeast Asia

[2] Ahmed (2008), Intelligent watermark recovery using spatial domain extension, International conference on intelligent information hiding and multimedia signal processing, IIHMSP' 08

[3] Lai, C.C., Tsai, C.C. (2010): Digital Image Watermarking Using Discrete Wavelet Transform and Singular Value Decomposition. IEEE Transactions on Instrumentation and Measurement 59(11), 3060-3063

[4] Soliman MM, Hassanien AE, Ghali NI, Onsi HM, (2012) "An Adaptive Watermarking Approach for Medical Imaging using Swarm Intelligence", Int Journal Smart Home 6:37-50

[5] Zain J, Clarke M, (2011) Security in Telemedicine: Issue in Watermarking Medical Images, International Conference: Science of Electronic, Technologies of Information and Telecommunications

[6] Giakoumaki A, Pavlopoulos S, KoutsourisD (2006) Secure and efficient health data management through multiple watermarking on medical images. Med Biol Eng Comput 44: 619–631

[7] Lin W-H, Wang Y-R, Horng S-J, Kao T-W, Pan Y (2009) A blind watermarking method using maximum wavelet coefficient quantization. Expert Syst Appl 36(9) : 11509–11516

[8] Liu, R., Tan, T. (2002): An SVD-based watermarking scheme for protecting rightful ownership, IEEE Transactions on Multimedia 4(1), 121-128

[9] Lai, C.C., Tsai, C.C (2010).: Digital Image Watermarking Using Discrete Wavelet Transform and Singular Value Decomposition. IEEE Transactions on Instrumentation and Measurement 59(11), 3060-3063

[10] Soliman MM, Hassanien AE, Ghali NI, Onsi HM, (2012) "An Adaptive Watermarking Approach for Medical Imaging using Swarm Intelligence", Int Journal Smart Home 6:37-50

[11] Zain J, Clarke M, Security in Telemedicine: Issue in Watermarking Medical Images, International Conference: Science of Electronic, Technologies of Information and Telecommunications

[12] Paul (2006), A novel VLSI architecture for image compression, Eighth IEEE international symposium on multimedia

[13] Shaick (2000), A hybrid transform method for image denoising. 10th European. Signal Processing Conference,

[14] Chappelie (2004)r, image coding with iterated contourlet and wavelet transforms, international conference on image processing,

[15] Ahire, V.K., Kshirsagar, V. (2011): A Digital Watermarking Scheme Based Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) for Copyright Protection of Digital Images. International Journal of Computer Science and Network Security 11(8), 208-213

[16] Singh AK, Dave M, Mohan A (2014) Hybrid technique for robust and imperceptible image watermarking in DWT- DCT-SVD domain. Natl Acad Sci Lett 37(4):351–358

[17] N. Ponomarenko, V. Lukin, M. Zriakhov, K. Egiazarian, and J. Astola (2006), Estimation of accessible quality in noise image compression, in Proceedings of European Signal Processing Conference (EUSIPCO ’06), pp. 1–4, Florence, Italy.

[18] S. G. Chang, B. Yu, and M. Vetterli, (2000) Adaptive wavelet thresholding for image denoising and compression, IEEE Transactions on Image Processing, vol. 9, no. 9, pp. 1532–1546.

[19] Z. Wang and A. C. Bovik (2006). Modern Image Quality Assessment. Morgan and Claypool Publishing Company, New York

[20] Z. Wang, A. C. Bovik, H. R. Sheikh (2004), and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612

[21] Z. Wang and A. C. Bovik (2009), Mean squared error: love it or leave it? A new look at signal fidelity measures, IEEE Signal Processing Magazine, vol. 26, no. 1, pp. 98–117.

[22] N. Ponomarenko, F. Battisti, K. Egiazarian, J. Astola, and V. Lukin (2009), Metrics performance comparison for color image database, in Proceedings of the 4th International Workshop on Video Processing and Quality Metrics, pp. 1–6, Scottsdale, Ariz, USA, CD-ROM.

[23] Zhou Wang1, Eero P. Simoncelli1 and Alan C (2003). Bovik multi-scale structural similarity for image quality assessment. Proceeding of the 37th IEEE Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 9-12, 2003.

[24] N. Nill, (1985) A visual model weighted cosine transform for image compression and quality assessment, IEEE Transactions on Communications COM-33, pp. 551-557.

[25] R. F. Zampolo, R. Seara, (2003) A Measure for Perceptual Image Quality Assessment”, in Proc. of Int. Conf. on Image Proc., Barcelona, Spain, pp: 433-436, Sept.

INTENTIONAL BLANK

MULTIPLE SCLEROSIS DIAGNOSIS WITH FUZZY C-MEANS

Saba Heidari Gheshlaghi[1], Abolfazl Madani[2], AmirAbolfazl Suratgar[3] and Fardin Faraji[4]

[1,3]Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
[2]Department of Control Engineering, South Tehran Branch Islamic Azad University (IAU) Tehran, Iran
[4]Neurology Department, Arak University of Medical Sciences, Arak, Iran

ABSTRACT

Magnetic resonance imaging (MRI) can support and substitute clinical information in the diagnosis of multiple sclerosis (MS) by presenting lesion. In this paper, we present an algorithm for MS lesion segmentation. We revisit the modification of properties of fuzzy c means algorithms and the canny edge detection. Using reformulated fuzzy c means algorithms, apply canny contraction principle, and establish a relationship between MS lesions and edge detection. For the special case of FCM, we derive a sufficient condition for fixed lesions, allowing identification of them as (local) minima of the objective function.

KEYWORDS

Multiple Sclerosis, MRI, T2, fuzzy c-means (FCM), Canny.

1. INTRODUCTION

Over time, the importance of correct diagnosing for diseases are increased and it is an essential to find an optimal method for diagnosing disease. Multiple sclerosis, also known as MS, is a chronic disease that attacks the central nervous system(CNS) and affects white matter by the person's own immune system so that MS is known as an auto-immune disease. As we know, Nerve fiber is surrounded by myelin, which protects the nerve and helps it to conduct electrical Signals (Impulse). In MS patients, the myelin is disappeared. Because of this, the MS patients could not move their body properly. MS is a largely unknown disease these days and diagnosis it correctly and early, has a significant impact on disease progression [1]

Brain lesions detection plays an important role in MS studies, as it is used to evaluate patient disease and its future evolution. Currently, lesions are detected manually or semi-automatic segmentation methods, which are very time to consuming and show a high inter and intra-raters variability [2]

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 97– 104, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80210

Magnetic resonance imaging (MRI) were officially included in the diagnostic workup of patients presenting with a clinically isolated syndrome suggestive of multiple sclerosis in 2001 by an international panel of experts [3]. Diagnosing multiple sclerosis depends on evidence of disease dissemination in space and time and exclusion of other syndromes that can imitate multiple sclerosis by their clinical and laboratory Specifications. MRI can support and substitute clinical information for multiple sclerosis diagnoses, allowing an early and accurate diagnosis. [4] The clinical presentation of MS includes a wide range of physical disorders and cognitive symptoms. Cognitive Impairment decrease quality of life and treatment amongst MS patients [5-8]. MS is an inflammatory demyelinating and degenerative disease of the CNS, distinctive pathologically by a different part of brain inflammation, demyelination, axonal loss, often causing motor, sensorial, vision, correlation, and cognitive impairment [9].

Also, MS is the recurring, neurological disease capable of causing disability in young adults. In these days, getting MS disease has been increased, and geographical areas play an important role in getting MS. Moreover, Multiple sclerosis is between two and three-times more average in females than in males, but males have a propensity for later disease [10]. As mentioned, one of the most famous facilities for diagnosis MS disease is Magnetic Resonance Imaging (MRI) techniques [9], T2-weighted (T2-w) and gadolinium-enhanced T1-weighted (T1-w), are highly aware in detecting MS plaques. MRI-derived metrics have become the most important medical tool for diagnosing MS disease.

Both sharp and persistent MS plaques appear as focal high-signal intensity areas on T2-w sequences, illustrate their increased tissue water content. The total T2 lesion volume of the brain increases by almost (5–10%) each year in the relapsing forms of MS [11]. Gadolinium-enhanced T1-w imaging is highly fragile in detecting inflammatory activity. CNS atrophy, which involves both gray matter and white matter, is a progressive phenomenon that becomes worse with argument disease length and progresses at a rate of between 0.6% and 1.2% of brain loss per year in this disease [12]. It should consider that different method of segmentation might have an important impression on the small volume of the lesion detection.

In this paper, Section 2 reviews of the pre-processing steps needed when automatically segmenting MS lesions. Section 3 approaches segmentation methods and also reviews algorithms which are used. In Section 4, conclusion and results are showing and give some ideas for future works.

In this research, MRIs for patients with MS are from different sources. Mostly, our high gratitude should be directed to C.P. Loizoua and his colleagues: The Laboratory of Health at the University of Cyprus (http://www.medinfo.cs.ucy.ac.cy/)

Figure 1: Samples of MR images for MS patient

2. PRE-PROCESSING STEPS

Accurate identification of brain in MRI images is a critical first step in many neuroimaging studies. There are many technics for skull extraction [19, 20]. Although this step is a significant part, using an efficient method is important. In our research, we used the automated brain extraction tool (BET) (Smith, 2002) [13-15]. Also, image binarization is done in this step Eq. (1). th is described as a threshold.

$$Bin(x,y)=\begin{Bmatrix} 1 & if & f(x,y)>th \\ 0 & if & Ow \end{Bmatrix} \quad (1)$$

3. SEGMENTATION METHODS AND ALGORITHM

Varieties of methods for automatic analysis and segmentation have been developed that the image can be segmented into its basic elements.

Edge detection is an image processing technique for finding the boundaries of objects by detecting brightness discontinuities; it is also used for image segmentation and data extraction methods such as image processing, computer vision, and machine vision.

There are various classification methods for edge detection, but they are usually divided into two main groups:

a) Gradient method: The gradient method looks for maximum and minimum in the first derivative of the image for detecting the edges [16].

b) Laplacian method: It searches for zero crossings in the second derivative of the image to find edges. An edge has the one-dimensional shape of a ramp and calculating the derivative of the image can highlight its location. The zero-crossing of the differentiated signal as edge points [16]. The two-dimensional Gaussian operator G(x, y) is given by:

$$G(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2})} \quad (1)$$

With condition; i.e. $\sigma_x=\sigma_y$ and $x^2+y^2=r^2$ the equation simplify to:

$$G(r)=\frac{1}{2\pi\sigma^2}e^{\frac{r^2}{2\sigma^2}} \quad (2)$$

For detecting edges, searching for zero-crossing in 2nd order derivative of the image needed which f(x,y) is the image [2].

$$g(x,y)=\nabla^2 G(r)*f(x,y) \quad (3)$$

$$\nabla^2 G(r) = \frac{-1}{4\pi\sigma^2}[1-\frac{r^2}{2\sigma}]e^{\frac{r^2}{2\sigma^2}} \quad (4)$$

Common edge detection algorithms consist Sobel, Canny, Prewitt, Marr-Hildreth. Those methods are determined in this research. The main difference between these methods is that canny edge detection method is in the spatial domain while others are in the frequency domain [9].

Below, the differences between Canny, Sobel, and Marr-Hildreth edge detection is shown in figure2. As shown in figure 2, canny edge detection has a better result in comparison with other methods.

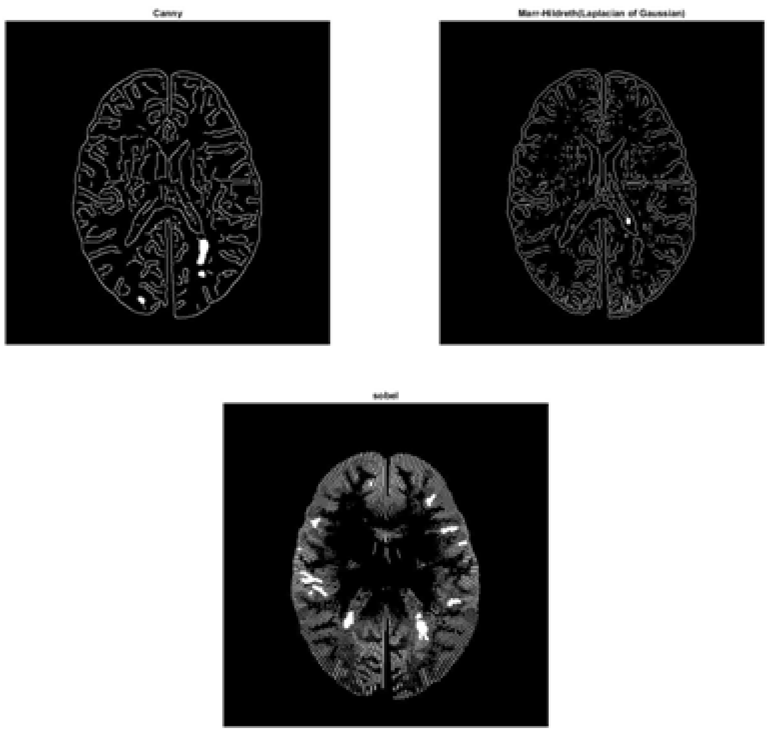


Figure 2: differences between Canny, Sobel and Marr-Hildreth edge detection

4. FUZZY C-MEANS

Clustering of numerical data forms of many classification and system modeling methods. The suggestion of clustering is to identify natural groupings of data from a large dataset to create a concise representation of a system's behavior. Fuzzy c-means (FCM) is a data clustering algorithm in which a data set is grouped into n clusters with any data point in the data set belonging to every cluster to the main degrees [17]. The FCM algorithm assumes that each point belongs to more than one cluster with given a dataset $X = \{x_1, x_2, x_3 ..., x_N\}$ clustering goals to discrete the dataset into varying subsets each with a cluster midpoint. FCM algorithm apportions fuzzy memberships u_{ij} any pixels x_j $(j = 1, 2, ..., N)$ in each variety c by decreasing the following cost function:

$$J = \sum_{j=1}^{N}\sum_{i=1}^{C} u_{ij}^{m} \| x_j - v_i \|^2 \tag{5}$$

Where m is the weighting fuzziness parameter. v_i Illustrate the i-th cluster midpoint. $\|.\|$ Explain the Euclidean interval. u_{ij} Means a partition matrix that is subject $u_{ij} \in [0,1] \quad and \quad \sum_{i=1}^{c} u_{ij} = 1$. Using the Lagrangian method, partition matrix, and cluster centers by calculating as follows:

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} x_j}{\sum_{j=1}^{N} u_{ij}^{m}} \tag{6}$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\| x_j - v_i \|^2}{\| x_j - v_k \|^2} \right)^{\frac{1}{(m-1)}}} \tag{7}$$

In this study, considering the MR image is often with spatial inhomogeneity due to the defect in the magnetic field, the intensity x_j is modeled by $x_j = y_i - \gamma_j$ here, y_i and γ_j, and mean the measured absorption and the consistent bias field. We propose a novel algorithm as an addition of the traditional FCM clustering by modifying the cost function in Eq. (9) as [18, 19]:

$$J = \sum_{j=1}^{N}\sum_{i=1}^{c} u_{ij}^{m} \| x_j - v_i \|^2 + \alpha S + \beta R \tag{8}$$

Where S is a normalizer indicating the neighborhood influence during segmentation. R is the regularization term on bias field. α and β are constants that control the effect of the two regularization term correspondingly definitely, S, R are established as [14]:

$$S = \sum_{j=1}^{N}\sum_{i=1}^{C} u_{ij}^{m} \left(\frac{1}{N_\varepsilon} \sum_{x \in N_\varepsilon} \| x - v_i \|^2 \right)$$
$$and \tag{9}$$
$$R = \sum_{j=1}^{N}\sum_{i=1}^{C} u_{ij}^{m} \left(\| \gamma_j \|^2 \right)$$

Where N_ε represent neighborhood positioning at x_j.

Also to the traditional FCM algorithm, by differentiating the cost function respect to, V_i and γ_j respectively and setting the result to zero, we can obtain the updating u_{ij}^* , v_i^* and γ_j^* as:

$$u_{ij}^* = \frac{1}{\sum_{k=1}^{c}\left(\frac{\| x_j - v_i \|^2 + \alpha . \frac{1}{N_\varepsilon} \sum_{x \in N_\varepsilon} \| x - v_i \|^2 + \beta . \gamma_j}{\| x_j - v_k \|^2 + \alpha . \frac{1}{N_\varepsilon} \sum_{x \in N_\varepsilon} \| x - v_k \|^2 + \beta . \gamma_j}\right)^{\frac{1}{(m-1)}}} \tag{10}$$

$$v_i^* = \frac{\sum_{j=1}^{N} u_{ij}^m (x_j + \alpha . \frac{1}{N_\varepsilon} \sum_{x \in N_\varepsilon} (x - v_i))}{(1+\alpha) \sum_{j=1}^{N} u_{ij}^m} \tag{11}$$

$$\gamma_j^* = y_j - \frac{\sum_{i=1}^{c} u_{ij}^m v_i}{(1+\beta) \sum_{i=1}^{c} u_{ij}^m} \tag{12}$$

The result is shown below (fig4). For achieving this result, first, we use the pre-processing method as described in section 2, then applying canny edge detection and finally, for increasing accuracy, FCM is applied.

Figure 3: Result of adding FCM after canny edge detection

5. CONCLUSIONS

Multiple sclerosis caused when the immune system attacks myelin, which covered around human nerve fibers to protect them and help them to send the messages efficiently.[1] Without this shell, nerves became damaged. Scar tissue may form MS, is a long-lasting disease that can affect the brain, spinal cord, and the optic nerves in eyes. By this damage, transferring messages between brain and body will be affected. The message will slow down or blocked during transferring and leading to the symptoms of MS. Doctors use medical history, physical exam, neurological exam, especially MRI to diagnose it. MR images still is a valuable tool for identification multiple sclerosis. There is no cure for MS, but medicines may slow it down and help control the disease progress.[7] Physical and occupational therapy may also be helpful. Because of the sensitivity of MR images, determine the disease clinically is hard and take a long time so having an automatic method for diagnosing the disease fast and accurate is critical.

In our research, different image processing methods are used, such as edge detection and segmentation. Also, we use some famous pre-processing techniques for having the best result, like Brain extraction tool [13-15] and binarization. In addition, we presented a modification method with Fuzzy C Means for better lesion segmentation and using canny for better edge detection. also show that, canny edge detection has a better result. In our method, we first used edge detection method for finding edges and then applying FCM for increasing diagnostic accuracy. In addition, we change clustering parameters to get different results for better lesion detection, Also as it was mentioned before, in MS lesion detection, effectiveness is the most important part. In this paper, lesion diagnosis accuracy is improved in comparison with others works [20].

REFERENCES

[1] Van Waesberghe, J. H. T. M., et al. "Axonal loss in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability." Annals of neurology 46.5 (1999): 747-754.

[2] Beaumont, J., Commowick, O., & Barillot, C. Multiple Sclerosis lesion segmentation using an automated multimodal Graph Cut. In Proceedings of the 1st MICCAI Challenge on Multiple Sclerosis Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure–MICCAI-MSSEG (2016, October): 1-8.

[3] Golkov, Vladimir, Tim Sprenger, Jonathan Sperl, Marion Menzel, Michael Czisch, Philipp Samann, and Daniel Cremers. "Model-free novelty-based diffusion MRI." In Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on , (2016) :1233-1236.

[4] Filippi, Massimo, et al. "MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines." The Lancet Neurology 15.3 (2016): 292-303

[5] Uher, T., et al. "Identification of multiple sclerosis patients at highest risk of cognitive impairment using an integrated brain magnetic resonance imaging assessment approach." European journal of neurology 24.2 (2017): 292-301.

[6] Kjølhede, Tue, et al. "Can resistance training impact MRI outcomes in relapsing-remitting multiple sclerosis?." Multiple Sclerosis Journal (2017): 1-10.

[7] Preziosa, Paolo, et al. "Comparison of MRI Criteria for the Diagnosis of Multiple Sclerosis: Role of Cortical Lesions (S45. 003)." Neurology 86.16 Supplement (2016): S45-003.

[8] Favaretto, Alice, et al. "MRI-detectable cortical lesions in the cerebellum and their clinical relevance in multiple sclerosis." Multiple Sclerosis Journal 22.4 (2016): 494-501.

[9] Denitto, Matteo, et al. “A biclustering approach based on factor graphs and max-sum algorithm.” Pattern Recognition 62(2017): 114-124

[10] A. Compston, A. Coles, Multiple sclerosis, Lancet 359 (9313) (2006) 1221–1231.

[11] S. Saha, S. Bandyopadhyay, A new point symmetry based fuzzy genetic clustering technique for automatic evolution of clusters, Inf. Sci. 179 (9) (2009): 3230–3246.

[12] H. McFarland, L. Stone, P. Calabresi, H. Maloni, C. Bash, J. Frank, MRI studies of multiple sclerosis: implications for the natural history of the disease and for monitoring effectiveness of experimental therapies, Mult. Scler. 2 (4) (1996): 198–205.

[13] Popescu, Valeriu, et al. "Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis." NeuroImage 61.4 (2012): 1484-1494.

[14] Eskildsen, Simon F., et al. "BEaST: brain extraction based on nonlocal segmentation technique." NeuroImage 59.3 (2012): 2362-2373.

[15] Rex, David E., et al. "A meta-algorithm for brain extraction in MRI." NeuroImage 23.2 (2004): 625-637.

[16] Walker, Stephen G. "A Laplace transform inversion method for probability distribution functions." Statistics and Computing 27.2 (2017): 439-448.

[17] Meena Prakash, R., and R. Shantha Selva Kumari. "Fuzzy C means integrated with spatial information and contrast enhancement for segmentation of MR brain images." International Journal of Imaging Systems and Technology 26.2 (2016): 116-123.

[18] Horsfield, Mark A., et al. "Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: application in multiple sclerosis." Neuroimage 50.2 (2010): 446-455.

[19] L.P. Clarke, R.P. Velthuizen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, M.L. Silbiger, MRI segmentation: methods and applications, Magn. Reson. Imag. 13 (3) (1995): 343–368.

[20] Ali, S. M., and Asmaa Maher. "Identifying multiple sclerosis lesions in MR images using image processing techniques." Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), Al-Sadeq International Conference on. IEEE, (2016):1-4

INFORMATIZED CAPTION ENHANCEMENT BASED ON IBM WATSON API AND SPEAKER PRONUNCIATION TIME-DB

Yong-Sik Choi, YunSik Son and Jin-Woo Jung

Department of Computer Science and Engineering,
Dongguk University, Seoul, Korea

ABSTRACT

This paper aims to improve the inaccuracy problem of the existing informatized caption in the noisy environment by using the additional caption information. The IBM Watson API can automatically generate the informatized caption including the timing information and the speaker ID information from the voice information input. In this IBM Watson API, when there is noise in the voice signal, the recognition results are not good, causing the informatized caption error. Especially, it is more easily found in movies such as background music and special sound. Specifically, to reduce caption error, additional captions and voice information are entered at the same time, and the result of the informatized caption of voice information from IBM Watson API is compared with the original text to automatically detect and modify the error part. Based on the database containing the average pronunciation time, each word for each speaker is changed into the informatized caption in this process. In this way, more precise informatized captions could be generated based on the IBM Watson API.

KEYWORDS

Informatized caption, Speaker Pronunciation Time, IBM Watson API, Speech to Text Translation

1. INTRODUCTION

Recently, artificial intelligence technology is being researched and developed in various fields. Artificial intelligence refers to the intelligence created by a machine, and is the intelligence that a computer program behaves and calculates, such as human thinking. However, since artificial intelligence that does not understand human language is useless, the most important thing in artificial intelligence technology is natural language processing technology and speech recognition technology. Typical speech recognition technologies include speech to text conversion. Among captions in which speech is converted into characters, captions including timing information and speaker ID information [1] are referred to as informatized captions [2]. Such an informatized caption can be generated using the IBM Watson API or the like [3]. However, the IBM Watson API is more susceptible to clipping errors due to poor recognition results if there is noise in the audio signal, especially in movies such as movies where background music and special sounds are used. In order to solve this problem, there has been proposed a method of predicting the timing information of the informatized caption based on a linear estimation [2] formula proportional to the number of alphabets. In this paper, we use the IBM

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 105– 110, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80211

Watson API, which provides basic functions of informatized caption including timing information, speaker ID information, and so on, to generate a word information list based on the proposed method.

2. SPEAKER PRONUNCIATION TIME-DB (SPT-DB)

2.1. Structure

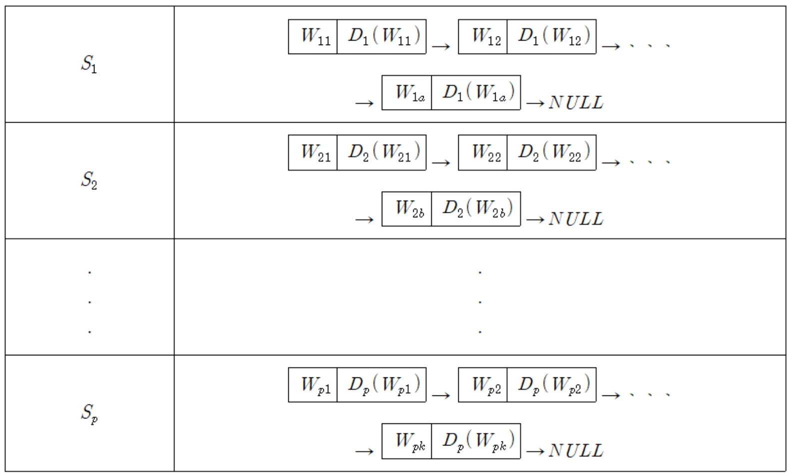


Figure 1. Structure of SPT-DB

SPT-DB consists of each node for each speaker(S_p) as shown in Fig. 1.The nodes consist of the average pronunciation times(D_p) of each word(W_{pk}).The nodes of the speaker are arranged in ascending order based on the average pronunciation time, and are connected to each other, and a null value is present at the end. When SPT-DB searches for a word spoken by the speaker, it searches based on the pronunciation time.

2.2. Assumption

Before proceeding with the study, the following assumptions are based on SPT-DB. [Assumption] SPT-DB is already configured for each speaker.

3. PROPOSED ALGORITHM

3.1. Algorithm modifying incorrectly recognized word based on SPT-DB

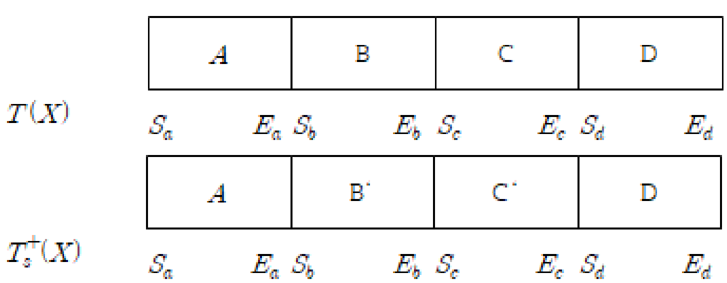


Figure 2. Original caption T(X) and informatized caption $T_s^+(X)$

Basically, original caption,T(X), and informatized caption from speech recognition result, $T_s^+(X)$, are input together.

Here, S_x and E_x mean the start time and end time of pronunciation for the word X, respectively.

[Step 1] Judge whether there is an incorrectly recognized word by comparing T (X) with $T_s^+(X)$.If there is no incorrectly recognized word, it terminates. If there is an incorrectly recognized word, go to the next step.

[Step2] Judge whether there are several consecutive words in the sequence, and pass the parameter to the case.

[Step3] Modify the words in the SPT-DB based on the start and end points of the cases.

[Step4] If there is an incorrectly recognized word in the following word, repeat steps 1 to 3 and terminate if there is no incorrectly recognized word.

3.2. Case 1: There is only one incorrectly recognized word.

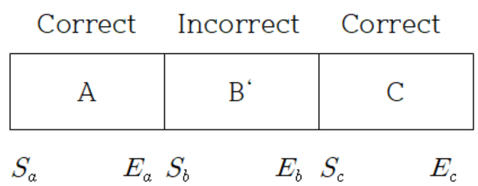


Figure3. There is one incorrectly recognized word

[Step1] Find the point at which the signal of a specific volume(dB) T or more starts for E_a to S_c and determine S_b.

[Step2] If there is a minimum time t' in S_b to S_c at which the signal intensity falls below a certain volume T and then remains below T until S_c, E_b = t' is determined. If there is no t' satisfying the above condition, $E_b = S_c$.

[Step3] Returns the start time and end time.

3.2. Case 2: There are more than two incorrectly recognized word.

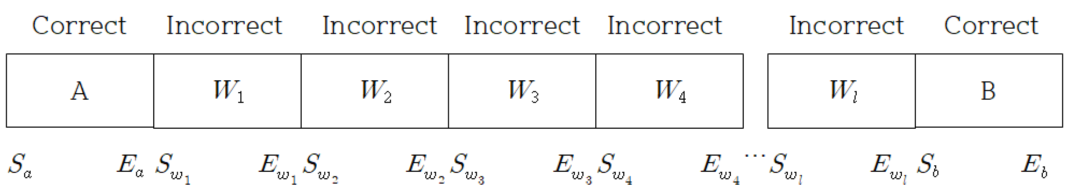


Figure 4. More than three incorrectly recognized word

[Step1] Find the point at which the signal of a specific volume(dB) T or more starts for E_a to S_{w2} and determine S_{w1}.

[Step2] If there is a minimum time t' in $S_{w1}$to S_{w2} at which the signal intensity falls below a certain volume T and then remains below T until S_{w2}, E_{w1} = t' is determined. If there is no t' satisfying the above condition,$E_{w1} = S_{w2}$.

[Step3] The ending point of the current word is obtained by multiplying the start time of the current word by the ratio of the pronunciation time of the incorrectly recognized words to the average pronunciation time of the current word. The following are summarized as follows.

$$E_{w_i} = \left(E_{w_l} - S_{w_1}\right) \times \frac{D(W_i)}{\sum_{i=1}^{l} D(W_i)}$$

[Step4] Returns the start time and end time.

4. CASE STUDY

The case was tested based on English listening assessment data. Fig.5 shows a problem of the English listening evaluation for university entrance examination. In a noisy environment like Fig.6, the accuracy dropped significantly. For reference, the original voice source was synthesized with raining sound using Adobe Audition CC 2017 to create a noisy environment.If we improve the proposed algorithm with noise, we can obtain the same result as Table1. The accuracy of speech recognition is 100% by the help of original caption and each word includes its own start time and end time.

W: Dad, I want to send this book to Grandma. Do you have a box?
M: Yeah. I've got this one to put photo albums in, but it's a bit small.
W: The box looks big enough for the book. Can I use it?

Figure 5. Original caption

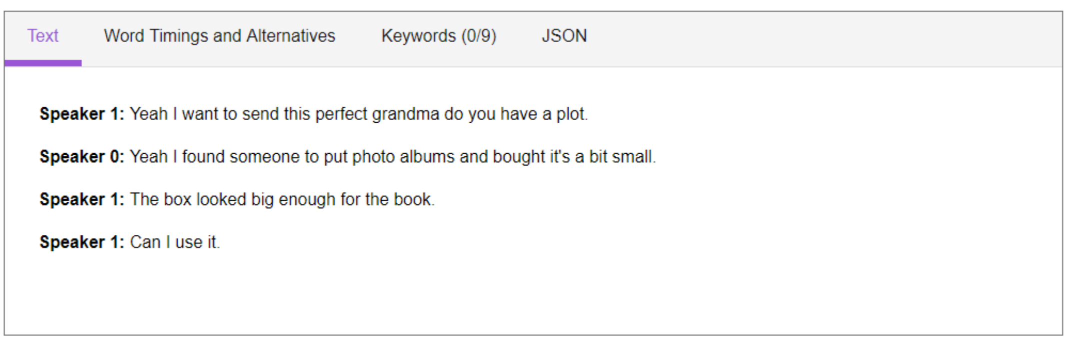


Figure 6. Recognition of mixed voice with rain noise by IBM Watson system

Table1.Informatized caption modified by the proposed algorithm

Sentence \ Word		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	Speaker 0	Dad	I	want	to	send	this	book	to	Grandma	Do	you	have	a	box	
		0.03-0.58	0.74-0.87	0.87-1.19	1.19-1.35	1.35-1.66	1.66-1.83	1.83-2.1	2.1-2.24	2.24-2.89	3.21-3.45	3.45-3.67	3.67-3.95	3.95-4.03	4.03-4.75	
B	Speaker 1	Yeah	I've	got	this	one	to	put	photo	albums	in	but	it's	a	bit	small
		5.22-5.7	6.01-6.27	6.27-6.62	6.62-6.86	6.86-7.15	7.15-7.26	7.26-7.48	7.48-7.88	7.88-8.29	8.29-8.59	8.69-9.1	9.28-9.48	9.48-9.55	9.55-9.81	9.81-10.51
C	Speaker 0	The	box	looks	big	enough	for	the	book	Can	I	use	it			
		10.86-10.99	10.99-11.41	11.41-11.67	11.67-11.96	11.96-12.26	12.26-12.47	12.47-12.6	12.6-13.16	13.46-13.71	13.71-13.79	13.79-14.12	14.12-14.42			

5. CONCLUSIONS

In this paper, we propose an algorithm to find and modify incorrectly recognized words based on the SPT-DB, which stores the average pronunciation times and appearance frequencies of the corresponding words in the speaker to correct the errors in the informatized caption obtained through the IBM Watson API. However, the proposed algorithm has a limitation that SPT-DB should be created first because it is assumed that the information of the corresponding words already exists in SPT-DB. Future research will be conducted to modify incorrectly recognized words while performing speech recognition and to update the SPT-DB in real time.

ACKNOWLEDGEMENTS

Following are results of a study on the "Leaders in INdustry-university Cooperation+" Project, supported by the Ministry of Education and National Research Foundation of Korea, and partially supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2015R1D1A1A09061368),and also partially supported by this work was supported by the Korea Institute for Advencement of Technology(KIAT) grant funded by the Korean government (Motie: Ministry of Trade, Industry &Energy, HRD Program for Embedded Software R&D) (No. N001884).

REFERENCES

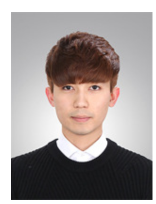
[1] CheonSun Kim, "Introduction to IBM Watson with case studies." Broadcasting and Media Magazine, Vol.22, No. 1,pp24-32.

[2] Yong-Sik Choi, Hyun-Min Park, Yun-Sik Son and Jin-Woo Jung, "Informatized Caption Enhancement based on IBM Watson API," Proceedings of KIIS Autumn Conference 2017, Vol. 27, No. 2, pp105-106.

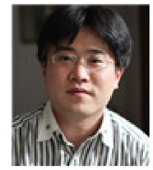
[3] IBM Watson Developer's Page, https://www.ibm.com/watson/developer

AUTHORS

Yong-Sik Choi has been under M.S. candidate course at Dongguk university, Korea, since 2017. His current research interests include machine learning and intelligent human-robot interaction.



YunSik Son received the B.S. degree from the Dept. of Computer Science and Engineering, Dongguk University, Seoul, Korea, in 2004, and M.S. and Ph.D. degrees from the Dept. of Computer Science and Engineering, Dongguk University, Seoul, Korea in 2006 and 2009, respectively. He was a research professor of Dept. of Brain and Cognitive Engineering, Korea University, Seoul, Korea from 2015-2016. Currently, he is an assistant professor of the Dept. of Computer Science and Engineering, Dongguk University, Seoul, Korea. Also, His research areas include secure software, programming languages, compiler construction, and mobile/embedded systems.



Jin-Woo Jung received the B.S. and M.S. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1997 and 1999, respectively and received the Ph.D. degree in electrical engineering and computer science from KAIST, Korea in 2004. Since 2006, he has been with the Department of Computer Science and Engineering at Dongguk University, Korea, where he is currently a Professor. During 2001~2002, he worked as visiting researcher at the Department of Mechano-Informatics, University of Tokyo, Japan. During 2004~2006, he worked as researcher in Human-friendly Welfare Robot System Research Center at KAIST, Korea. During 2014, he worked as visiting scholar at the Department of Computer and Information Technology, Purdue University, USA. His current research interests include human behaviour recognition, multiple robot cooperation and intelligent human-robot interaction.

MAMMOGRAPHY LESION DETECTION USING FASTER R-CNN DETECTOR

Reza Reiazi[1], Reza Paydar[2], Ali Abbasian Ardakani[3], Maryam Etedadialiabadi[4]

[1,3]Department of Medical Physics, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[2]Department of Diagnostic Radiology, School of Allied science, Iran University of Medical Sciences, Tehran, Iran
[4]Department of Radiology, Milad Hospital, Tehran, Iran

ABSTRACT

Recently availability of large scale mammography databases enable researchers to evaluates advanced tumor detections applying deep convolution networks (DCN) to mammography images which is one of the common used imaging modalities for early breast cancer. With the recent advance of deep learning, the performance of tumor detection has been developed by a great extent, especially using R-CNNs or Region convolution neural networks. This study evaluates the performance of a simple faster R-CNN detector for mammography lesion detection using a MIAS databases.

KEYWORDS

Mammography, Convolution Neural Network, R-CNN, lesion

1. INTRODUCTION

Breast cancer is one of the most common causes of death for women in many countries. In general, early detection is very important for cancer treatment and can decrease the mortality rate in a grate extent. One of the common used imaging modalities for early breast cancer detection is mammography in which an abnormality can be categorized as either normal, non-cancerous (benign) or cancerous (malignant).Studies have shown that Mammography can reduce30% of mortality in women who were undergoing mammography screening (1,2). Unfortunately the limited sensitivity and specificity of the screening mammography needs to improve diagnostic accuracy of mammography. R-CNN has been coined by Ross Girshick and his group at UC Berkeley as one of the impactful development in techniques convolutional neural networks. The purpose of R-CNNs is to draw bounding boxes over all of the objects in a given images following two step training schedules, first classification and then region detection step. Unfortunately R-CNN and its descendent fast R-CNN were computationally expensive, and extremely slow. Recently faster R-CNN has been introduced to overcome the somewhat complex training process that both R-CNN and Fast R-CNN encountered by implementing a classification network using CNN as the region proposal mechanism for training and prediction steps. This network reduces the overall training pipeline to just check the final convolutional feature map and produce region

Natarajan Meghanathan et al. (Eds) : NLP, JSE, CST, SIP, ARIA - 2018
pp. 111– 115, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80212

proposals from that. From that stage, the same pipeline as R-CNN is used (ROI pooling, FC, and then classification and regression heads).In this study, we conduct an experiment using faster R-CNN to detect breast abnormalities as normal, benign and malignant.

2. MATERIAL & METHODS

The proposed methods has been comprised in following steps:(1) Data preparation and preprocessing of input images(2) Design and train a faster R-CNN detector,(3) Evaluate Detector Using Test Set. Overall process of this work has been illustrated in figure 1.

2.1 DATA PREPARATION AND PRE-PROCESSING OF INPUT IMAGES

This study used mammography images from a public dataset of the mini Mammogram Image Analysis Society (MIAS) (3). This database includes of 332 mammograms in 1024 x 1024 sizes. It we just used 102 cases which was abnormality (benign or malignant).For the region detector training we have used Matlab Training Image Labeler app to produce required dataset. All cases were labeled in malignant, benign, calcification benign and background (figure 1).

1 imageFilename	2 Benign	3 Malignant	4 Calcification_Benign
'G:\mamography Faster R-CNN\Miasdatabase\mdb063.pgm'	[498,526,101,87]	[]	[]
'G:\mamography Faster R-CNN\Miasdatabase\mdb069.pgm'	[414,579,109,90]	[]	[]
'G:\mamography Faster R-CNN\Miasdatabase\mdb072.pgm'	[]	[261,512,38,58]	[]
'G:\mamography Faster R-CNN\Miasdatabase\mdb074.pgm'	[]	[]	[352,541,40,46]
'G:\mamography Faster R-CNN\Miasdatabase\mdb075.pgm'	[]	[445,272,49,74]	[477,523,50,46;567,656,93,58]
'G:\mamography Faster R-CNN\Miasdatabase\mdb080.pgm'	[394,844,71,60]	[]	[]
'G:\mamography Faster R-CNN\Miasdatabase\mdb081.pgm'	[352,436,228,218]	[]	[]
'G:\mamography Faster R-CNN\Miasdatabase\mdb083.pgm'	[481,774,144,101]	[]	[]
'G:\mamography Faster R-CNN\Miasdatabase\mdb090.pgm'	[]	[490,459,64,53]	[]

Figure 1.Sample labeling data prepared for train faster R-CNN network

Preprocessing step includes region extraction and contrast enhancement. At the next step, the images contrast was enhanced applying contrast limited adaptive histogram equalization (CLAHE) techniques. This improves the image contrast especially at over internal breast structures. Furthermore we have tried to remove redundant information from training images using texture filters and morphology operators. The effect of preprocessing has been shown on figure 2. Finally dataset were randomly has been divided into training, testing (respectively 80% and 20% of the total dataset).

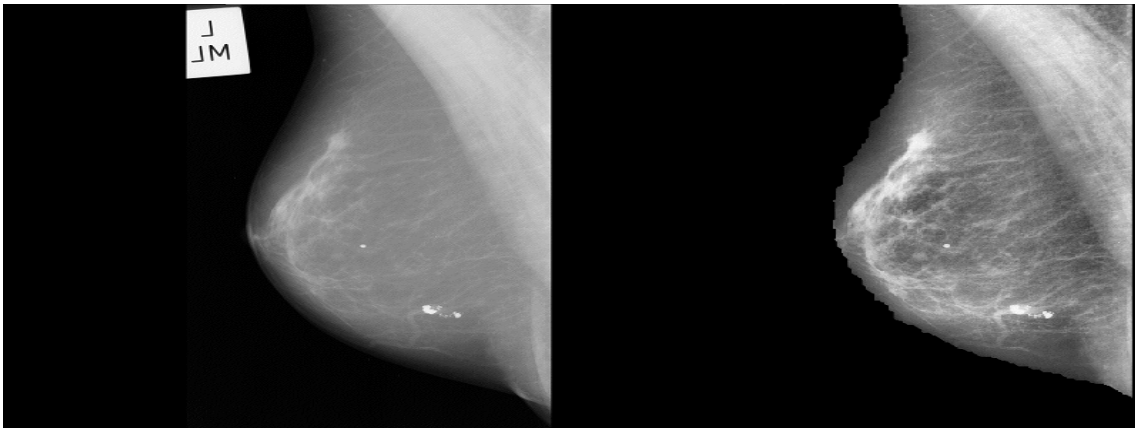


Figure2: the effect of image preprocessing. Contrast of original image (left) has been enhanced using CLAHE algorithm. Also any redundant information has been removed.

2.2 Design And Train Faster R-CNN Object Detector

The basis of the Faster R-CNN detector is a CNN so called region proposal network. We have design a CNN layer by using Matlab Neural Network Toolbox™ functionality. The type and size of the training images has been defined by the image Input Layer function. The input size is a trade off between spatial detail that a CNN needs to resolve and processing time which would be very high. For detection tasks, the CNN needs to analyze smaller sections of the image, so the input size must be similar in size to the smallest object in the data set. In this data set an input size of [75 75] has been selected. Building blocks of any CNN is a repeated basic layers of convolutional, ReLU (rectified linear units), and pooling layers follow by fully connected layers, a softmax loss layer and a classification. Repeating a consecutive blocks of convolutional and ReLU with or without pooling layer impose deepness of a CNN. The first convolution layer has a series of 10 filters of size 7x7x3. One pixel padding has been used to avoid discards of image border information too early in the network. A ReLu layer and max pooling layer (size of 7x7and a stride of 3 pixels) followed the Convolution Layer. The building block of our network has been shown in table 1.

Table1: Building blocks of faster R-CNN network

#	Layer Name	Layer Type	Layer Details
1	'imageinput'	Image Input	75x75x3 images with 'zerocenter' normalization
2	'conv_1'	Convolution	10 7x7x3 convolutions with stride [1 1] and padding [0 0]
3	'relu_1'	ReLU	ReLU
4	'maxpool_1'	Max Pooling	7x7 max pooling with stride [3 3] and padding [0 0]
5	'conv_2'	Convolution	10 7x7x10 convolutions with stride [1 1] and padding [0 0]
6	'relu_2'	ReLU	ReLU
7	'roi pooling layer'	ROI Max Pooling	ROI Max Pooling with grid size [3 3]
8	'fc_1'	Fully Connected	64 fully connected layer
9	'relu_3'	ReLU	ReLU
10	'fc_2'	Fully Connected	4 fully connected layer
11	'softmax'	Softmax	softmax
12	'classoutput'	Classification Output	crossentropyex with 'Benign', 'Malignant', and 2 other classes

2.3 EVALUATE DETECTOR USING TEST SET

Matlab Computer Vision System Toolbox™ provides functions for object detector evaluation which calculate average precision. The average precision is a single number that shows the ability of the detector to make correct classifications and the ability of the detector to find all relevant objects (recall).This will be done by running the detector on the test set.

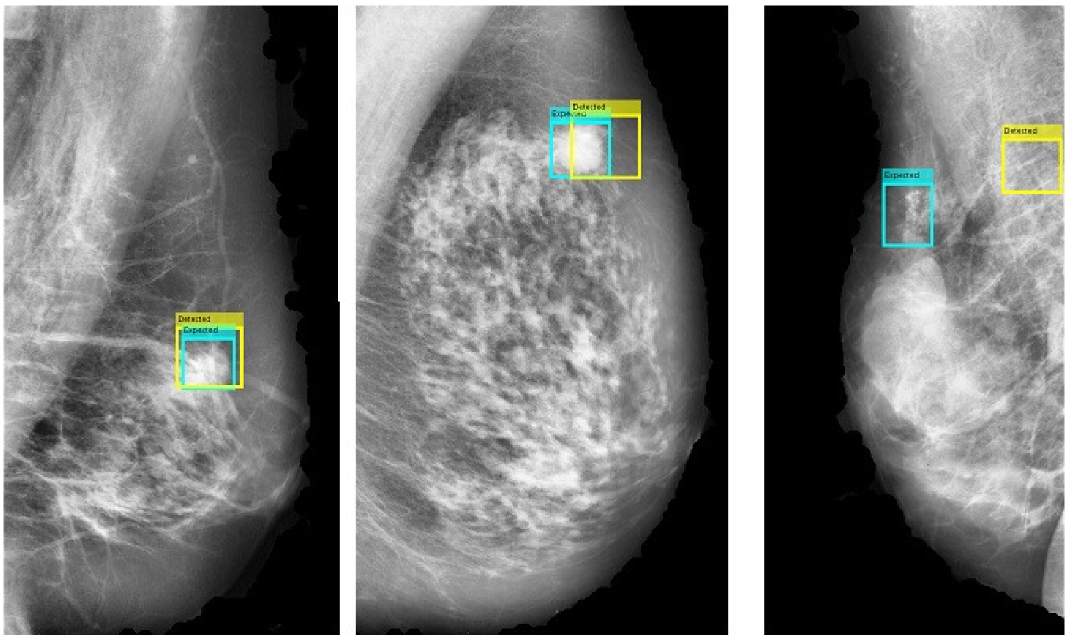


Figure3: Superimposing expected (blue) and detected (yellow) regions on a two malignant case with correct (right), semi correct (middle) and wrong (left) detection

3. RESULTS & DISCUSSION

We have run the trained detector over the test samples to evaluate the precision and recall quantities. Since more than one region with different labels has been detected, the closest box to the original ROI has been selected for evaluation process. If there were more than one closed detected region with different label, we have used that region which has the same label as the test data. Figure 3 has shown the resulting images from tree malignant cases. Ideally, the precision would be 1 at all recall levels. In this study, the average precision is 0.2. The use of more training data can improve the average precision, at the cost of longer training time.

In some cases despite wrong position detection, there were at least one or more detected malignant regions in 50 percent of malignant (15 from 16 malignant cases) and no detected malignant regions in 94 percent of benign cases (15 from 16 benign cases).This can be due to the similar features in malignant cases and benign cases. The precision/recall (PR) curve highlights how precise a detector is at varying levels of recall. Ideally, the precision would be 1 at all recall levels. The use of additional layers in the network can help improve the average precision. But our results from using deeper network using extra series convolution layer did not improve detection results since the network was not able to detect any region which may impose extra layers should be carefully designed and added. We will evaluate this in future applying different designs other that series networks.

4. CONCLUSION

Deep convolution neural networks have demonstrated the ability to outperform skilled humans in certain observational tasks (4,5). There are many studies showing reasonable performance of convolution neural networks on mammography lesion detection but none of them uses faster R-CNN detectors (6, 7, and 8). General evaluation studies have shown that an CNN network is able to outperform human performance (9). This study was performed based on Faster R-CNN detector applying a simple series convolution neural network. Trained detector had reasonable detection precision with very small training data (102 images) which can be concluded as

REFERENCES

[1] Duffy SW, et al, (2002), "The impact of organized mammography service screening on breast carcinoma mortality in seven swedish counties, Cancer, Vol. No. 95(3), pp 458–69.

[2] DB, et al, (2002), "Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality", Cancer, Vol. 94(2), pp 580

[3] Suckling J, et al, (1994), "The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica", International Congress Series, Vol. 1069, pp 375-378.

[4] Giusti R, Gambardella L, Schmidhuber J, (2012), "Neural networks for segmenting neuronal structures in EM stacks", ISBI Segmentation Challenge Competition.

[5] HeK, et al. (2015), "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification", arXiv, 1502.01852,.

[6] Krzysztof J, et al, (2017), "High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks", arXiv,1703.07047v2

[7] Araujo T, et al. (2017), "Classification of breast cancer histology images using Convolution Neural Networks", PLoS One, Vol. 12(6), pp e0177544.

[8] Zhang Q, et al, (2017), "Class Mammogram Classification Based on Descriptive CNN Features". Biomed Res Int", pp 3640901.

[9] Jane H, Anne C, (2017), "Applying Faster R-CNN for Object Detection on Malaria Images", CVPR workshop. pp 808-813.

AUTHOR INDEX