An Ontology-Based Affective Computing Approach for Passenger Safety Engagement on Cruise Ships

Annarita Cinquepalmi
DEI - Politecnico di Bari
Via E. Orabona 4
Bari, Italy I-70125
Email: annarita.cinquepalmi@poliba.it

Umberto Straccia
Istituto di Scienza e Tecnologie dell’Informazione
Consiglio Nazionale delle Ricerche
ISTI-CNR, Pisa, Italy
Email: straccia@isti.cnr.it

Abstract—The safety of cruise passengers is a key element for a cruise company. Among various aspects, the ability to interpret and recognize cruisers’ emotions, so that he/she feel safe, plays a central role in human communication. Affective computing addresses the computational processing of emotions. Current automatic emotion recognizers basically use automated classification tools to label emotions without capturing relations between biosignals and observations measured by the various sensors. This work proposes instead an Ontology Web Language (OWL) ontology-based emotion recognition framework by (i) monitoring human body vital signals through wearable, non-invasive sensors; and then the (ii) emotion detection is based on an ontology-based matchmaking process via non-standard automated reasoning services. A key factor is the use of so-called vague/fuzzy concepts, which are intrinsic in the realm of emotions and their dynamic evolution. To this end, we exploit Fuzzy Description Logics (Fuzzy DLs), which are the logical foundation of fuzzy OWL ontologies, i.e., OWL ontologies extended with vague/fuzzy concepts. An early prototype has been implemented w.r.t. a reference dataset and a preliminary experiment has been carried out with the aim to monitor the emotions experienced by cruise passengers while viewing safety video instructions.

Keywords - Affective Computing; Semantic Matching; Fuzzy OWL 2.

I. INTRODUCTION

The number of people taking cruises across the world has increased year-on-year. The Cruise Lines International Association \cite{1} estimates that 24 million passengers are expected to cruise this year. Despite statistics, passengers opinion about cruise ship safety changed since Costa Concordia accident in January 2012 off the coast of Italy (32 people died). Among others, passenger safety training is crucial for cruise company. To this end, it appears to be useful to monitor and capture passengers emotions, when viewing safety video instructions, to improve the emotional condition or prevent harmful health states of the passengers. According to \cite{2}, Affective Computing (AC) aims to create computational systems, which are Emotionally Intelligent (Eml), i.e., capable to recognize, understand and express emotions in order to improve users’ well-being. Eml systems may establish empathy with the user, e.g., through an interactive automated agent, i.e., affective avatar designed to perceive user emotional experiences when engaged in specific activities.

Physiological signals have been used increasingly in AC thanks to technological improvements in low-cost miniaturized unobtrusive wearable biosensors for continuous monitoring. Recently, manufacturers have been developing increasingly robust and cost-effective biosensors for fast and sensitive analysis of human body vital signals. Over the last decade, EmI systems have gained momentum for a wide number of applications in several important companies, e.g., NeuroFocus \cite{3} utilized electroencephalography (EEG), eye tracking, and biometrics to capture the non-conscious aspects, emotions, and preferences of consumer decision-making; and EmSense \cite{4} developed the proprietary unobtrusive EmBand\textsuperscript{TM} hardware for measuring positive/negative emotional response and cognitive engagement to advertising.

Emotion-aware systems identify specific outcomes from biosensors and respond by triggering appropriate actions within a given context. Additional examples include: (i) monitoring the elderly to recognize signs of health issues \cite{5}, such as sadness bouts as a symptom in depressed patients, and alerting healthcare providers; (ii) increasing safety of drivers by observing their emotions \cite{6} and, suggesting a relaxation technique if a state of anger or frustration is detected (biofeedback); and (iii) improving user satisfaction in smart home environments \cite{7}, by controlling domotic devices to favor comfort and resting.

Nevertheless, most existing approaches are still quite intrusive. Furthermore, studies are typically carried out in controlled laboratory conditions, hardly transferable to real scenarios. Basically, they rely on conventional computing architectures running procedures for signal processing and features extraction, which have high computational costs, affecting the performance in real time applications.

This paper proposes a quasi-real-time computing framework, which only leverages off-the-shelf technology for biosignal monitoring and analysis, attempting to go beyond current simple emotion classification by exploiting fuzzy OWL Ontologies \cite{8}\cite{9}. Biosignals and features are described through semantic annotations based on a reference ontology. In particular, fuzzy OWL enables the description and manipulation of vague concepts, such as emotions. Semantic-based processing of raw sensor data makes them machine-understandable and allows ontology-based knowledge to be
processed efficiently, even in mobile and pervasive contexts with severe resource limitations in memory, storage and energy consumption. For this purpose, the optimized Mini-ME embedded reasoning engine [10] is adopted.

Our framework works in three fundamental stages: (i) detect most relevant biosignals features; (ii) build an annotated description of the emotional dimensional model in terms of valence (V) and arousal (A) [11] via the FuzzyDL-Learner [12][13][14][15]; and (iii) use matchmaking to recognize the emotion from its dimensional features. Non-standard inference services [16] are exploited to compare annotations of the valence/arousal (VA) space, discovering the most emotion(s) experienced by the user.

The ultimate goal of the framework is to provide helpful and timely user feedback and/or provide customized services. The free public DECAF data set [17] was used for the initial implementation and experimentation of the proposed framework, in order to build a Fuzzy OWL ontology of emotions correlated with biosignals and the continuous VA model. Experiments on the proposed framework are still ongoing; this paper provides a proof of concept on the feasibility of our approach.

The remainder of the paper is organized as follows. In Section II, a literature analysis is given. Section III describes the framework in detail, while a representative case study for passenger safety engagement on cruise ships is described in Section IV. The conducted experiment is illustrated in Section V. The paper closes with concluding remarks and future perspectives.

II. RELATED WORK

Biosignals are multichannel time-varying recordings of parameters of the central and/or the autonomic nervous system. They are known to convey information that can be used for emotion assessment [18]. Using biosignals has some advantages over other methods: (i) they are relatively robust to voluntary control and manipulation, because they are governed by the human autonomous nervous system; (ii) using wireless wearable sensors they can be collected anytime and anywhere without active user input; (iii) they can be easily correlated with external channels like facial expression.

The literature about emotion recognition through biosensors shows a standardized procedure to build EmI systems, summarized in the four following stages: (i) signal acquisition; (ii) preprocessing (iii) feature extraction; and (iv) machine learning based classification [2][19]. They exploit conventional fixed computer architectures, thus preventing many realistic application scenarios.

Recent developments in Body Area Network (BAN) allow data gathered from wearable sensors routed through multi-hop wireless links toward a portable computing device (e.g., a smartphone) [20]. In mobile real-time emotion detection systems, performance of the processing pipeline is critical in terms of both computational efficiency and classification accuracy. Furthermore, classification yields trivial labels, without a formally structured description about the characteristics of the elicited user emotion.

The Semantic Web initiative generated standard logic-based languages for the representation of ontologies to enable machine-understandable characterization of knowledge domains. Emotion recognition approaches exploiting Semantic Web technologies exist in literature, although they are a minority. Zhang et al. [21] proposed a system based on reasoning rules applying a Decision Tree, but mining was exploited only to map data to a single class. Furthermore, a rule-based system is useful only if there is an exact match between rules and the data: this is quite rare in complex domains like emotion characterization.

The complexity of emotions, the non-linearity of biosignals, the impossibility to find a single model to represent emotions can be faced by adopting fuzzy systems [22]. They are generally robust and have the ability to process inaccurate and vague data.

Let us recap that, although a large amount of work has been carried out about fuzzy logic-based machine learning [23], fuzzy ontology-based machine learning techniques have been scarcely investigated in general [8] and not at all in the context of EmI. Together with the adoption of non-standard inference services, supporting approximate matches, this appears to be an interesting ingredient to improve EmI system performance in terms of fine-grained emotion categorization, flexible classification and logic-based explanation of the outcomes.

III. EMOTIONALLY INTELLIGENT SYSTEM: FRAMEWORK AND APPROACH

Our framework extends the standard emotion recognition works discussed in Section II. Workflow steps are basically the same, but semantic-based enhancements change the way each step is performed. The main peculiarities of the proposed approach are: (i) a real-time emotional patterns detection based on Fuzzy DLs [8]. Fuzzy DLs are the logical foundation of fuzzy OWL ontologies, i.e., OWL ontologies extended with vague/fuzzy concepts [24]; (ii) a semantic-based matchmaking process to recognize the most likely emotion, and (iii) a feedback action to improve the user’s emotional state.

The overall architecture is depicted in Fig. 1. Autonomous and unobtrusive sensor nodes can form a body area network or body sensor network (BSN) [25]. Gathered physiological
parameters are routed through multiple wireless links in the BAN to a portable device (e.g., a smartphone) with constrained computational resources for real-time DAQ (Data Acquisition). Physiological signals in response to stimuli are collected and used as the EmI system input. Before these data streams are fed into the computational model, feature extraction techniques are exploited.

The first efforts toward affect recognition have focused on finding the link between users emotional state and its corresponding physiological state, translating low-level data captured by sensing devices to high-level abstractions expressed by a semantic annotation. The goal of the semantic annotator component is to build a semantic annotation combining physiological features and bidimensional emotional parameters: valence or evaluation and arousal or activation. Valence represents how positive or negative (i.e., pleasant or unpleasant) an emotion is, while arousal represents a passive/active scale, ranging from calm to excited. The key idea in the above model is to represent emotions with just a coordinate system conveying basic attributes. As a consequence, any emotion could be represented as a point in this space. Thus, each emotional state can be defined as a combination of these dimensions, e.g., anger can be characterized by high arousal and negative valence, happiness by low arousal and positive valence, and sadness by low arousal and negative valence. If the emotion is completely neutral it should be assigned to the center point of the space.

Exploiting the FuzzyDL-Learner concept emotion descriptions are automatically learned from biosignal features compiled into an OWL 2 ontology. Through non-standard inference services, the semantic annotation is compared with emotion descriptions contained in the ontology, created in a training step from a reference biosignal dataset. Non-standard inference services for semantic matchmaking, implemented in the Mini-ME reasoner, produce the most appropriate elicited user emotion(s). The system captures the best action to enhance user’s affective states, giving a user feedback.

A prototypical fuzzy ontology modeling the domain of interest has been defined, using fuzzy OWL. The logical foundation of fuzzy OWL are Fuzzy DLs, an extension to classical DLs with the aim of dealing with fuzzy/vague/imprecise information (for more details see [8][24]). Roughly, in Fuzzy DLs, there are fuzzy concepts (representing classes of objects), fuzzy roles (a.k.a. properties, joining pairs of objects), individuals (specific named objects) and fuzzy datatypes (or fuzzy concrete domains defined over an interval of the rational numbers). The important aspect to know is that, unlike usually, objects may be an instance of a fuzzy concept to some degree in [0, 1], while in the non-fuzzy case an object is either instance or not an instance of a concept. Axioms are statements which represent is-a relations between concepts. The logical statement has a degree of truth allowing to define new fuzzy concepts from other ones during the learning process. It is beyond the scope of this work to illustrate the details of Fuzzy DLs. We refer the reader to [8].

The FuzzyDL-Learner system is used to learn automatically to identify relationship between human affective states and bidimensional emotional characteristics. The main feature of the FuzzyDL-Learner system is that it allows to learn graded fuzzy OWL 2 descriptions of a selected target class in terms of specific inclusion axioms expressed in OWL EL [27], in which, fuzzy concepts may occur to improve both the accuracy of the description, as well as their readability. The learner uses the pFOIL-DL learning algorithm [15] to automatically induce fuzzy concepts descriptions. pFOIL-DL is inspired on FOIL [28], a popular Inductive Logic Programming algorithm for learning sets of rules. The three main differences from FOIL are: (i) pFOIL-DL uses a probabilistic measure to score concept expressions, (ii) it does not remove positive examples covered from the training set, but leaves it unchanged after each learned rule and (iii) it evaluates the goodness of an induced rule not independently of previously learned rules, but considering the whole set of learnt expressions. Additionally, FuzzyDL-Learner automatically fuzzifies the range of the real-valued bidimensional emotion parameters and finds relationship between emotions and the VA space. Furthermore, it may provide an automatic natural language translation of the learned classification emotion rules. The conjunction of the dimensional value intervals associated to each emotion as determined by the training set becomes the annotation for that emotion. In this way, each emotion can be described as the conjunction of qualitative features, valence and arousal. For instance, the following is a learned description for the emotion Fear:

\[
\exists \text{hasArousal}. \text{Arousal}_\text{low} \land \\
\exists \text{hasValence}. \text{Valence}_\text{high} \sqsubseteq \text{Fear}
\]

dictating that Fear is identified by low arousal values and high valence values, where low arousal (resp. high valence) are automatically determined as illustrated in Fig. 2. The output constitutes the annotated dataset and is the factual knowledge in the reference fuzzy Ontology.

The subsequent classification task exploits a semantic-based matchmaking process computing non-standard inferences, implemented in the Mini-ME embedded reasoning engine [10]. In a generic setting, given a request R and a set of resources S expressed w.r.t. a reference ontology, semantic matchmaking allows to find and rank the best matching resources through non-standard inference procedures called Concept Contraction and Concept Abduction [16]. If R and S have conflicting characteristics, Concept Contraction determines new concepts G (Give up) and K (Keep); G is the explanation about what in R is incompatible with S, while K represents the compatible part. In addition, a penalty value is computed, which is the semantic distance of the description w.r.t. the request. Otherwise, if there is compatibility between R and S but R does not match S fully, Concept Abduction extracts the concept expression H (Hypothesis), expressing what should be hypothesized in S in order to completely satisfy R. A related penalty value of a service A w.r.t. a request B is computed as:

\[
d(A, B) = 100(1 - \frac{\text{penalty}_C + \text{penalty}_A}{\max \text{penalty}_C})
\]

where penalty_C and penalty_A are the penalty induced by Concept Contraction and Concept Abduction between
each service/resource annotation and the request. Penalty is normalized w.r.t. the maximum possible semantic distance from the request A, i.e., the one of the most generic DL concept (denoted ⊓); this distance depends only on the reference ontology. Finally, semantic affinity is expressed by a percentage of scores and the service with highest rank is selected by the requester.

We have adapted the semantic matchmaking problem to the discovery of user emotions in the following way. The request is defined by a semantic description expressed as the logical conjunction of information about emotional bidimensional features extracted from unlabeled user input. Resources are the semantic descriptions populating the previously annotated dataset. The annotated dataset and fuzzy concepts descriptions are fed to the matchmaking reasoner, and ranked penalty values associated with a logic-based explanation are obtained from the semantic matchmaking process. The emotion with the lowest penalty is identified as the best matching emotion from the current user’s biosignals.

IV. CASE STUDY: EMOTIONS IN VIEWING SAFETY VIDEO INSTRUCTIONS

The proposal explores how emotions affect consumer decision making in Emotional or Experiential Marketing [29]. In contrast to traditional marketing, emotional marketers focus to understand what influences consumer decision-making and stimulates their sense and minds. Consumer experience is not often caused by rational choice: typically, emotional responses drive human opinion and experience. The selected reference scenario and case study aims to capture passengers cruise emotions when viewing safety video instructions immediately after sailing. The video provides clear instructions and explains in detail the actions each person on board should follow in the event of an emergency. Passengers convey the emotion elicited, after viewing, in terms of valence and arousal. The subsequent FuzzyDL-Learner in conjunction with semantic-based matchmaking makes a detailed analysis of emotions and behaviors fully transparent to the user. The feedback is to suggest what parts of the safety video should be changed and in what way in order to favor cruiser engagement and serenity, so that he/she feel safe to travel with a cruise company and will enjoy their stay.

The freely accessible DECAF [17] database for affect recognition and tagging was used to assess the feasibility of our approach. It is a multimodal dataset for decoding user physiological responses to multimedia content: it consists of a collection of peripheral physiological signals and multi-modal recordings, taken from 30 healthy subjects. The records incorporate magnetoencephalogram, horizontal electrooculogram, electrocardiogram, trapezius electromyogram, and near infrared facial video signals. The participants watched 36 emotional videos and gave feedback in terms of valence and arousal. Arousal expresses the intensity of the emotional feeling built up when a subject watched a safety video, ranging on a discrete scale of 0 (very calm) to 4 (very aroused), valence refers to how was the feeling after watching a clip on a scale of -2 (unpleasant) to 2 (very pleasant). The chosen video clips are also associated with emotional tags.

Our prototypical system exploits the Fuzzy-DL Learner and the Mini-ME matcher for emotion detection. The workflow starts with valence, arousal participants’ self-assessment ratings information. To make sense of the data, z-score normalization rescaling is required. For each video normalized arousal and valence scores are calculated by taking the mean and standard deviation of arousal and valence ratings listed in [17], considered as ground truth. In order to enable a fully automated emotion annotation and matchmaking process,
the above meaningful emotional features are translated to an OWL ontology. A two-step modeling was devised to tie VA parameters to emotions. The former exploits bidimensional features as input to the FuzzyDL-Learner in order to fuzzify valence and arousal and to create a fuzzy OWL ontology by identifying axioms that express each emotional label. The latter maps the DECAF dataset to an annotated dataset to transform raw data into higher level knowledge according to previous fuzzified arousal/valence space.

Seven emotional classes were considered to induce concept descriptions: Amusement, Anger, Disgust, Excitement, Fear, Fun and Shock. The discretization method adopted by pFOIL-DL partitioned valence and arousal numeric datatypes into 5 fuzzy sets (veryLow, low, medium, high, veryHigh) with associated membership functions, as depicted in Figure 2. The learned expressions are reported in Table I. These axioms form our fuzzy OWL ontology. As a matter of example, e.g., anger has been characterized by high arousal and negative valence while amusement by low arousal and positive valence.

The goal of the second step is to build the annotated dataset, connoting each valence, arousal according to the fuzzy interval obtained in the previous phase. For example, assume subject with ID 9, after viewing a video, may reports V=0 and A=2. Then, self-assessment valence/arousal ratings provided by participants are processed as follows:

1) Valence/arousal ratings are z-score normalized considering ground truth mean and standard deviation in the training set. The chosen video clip has $\mu_A=1.20$, $\sigma_A=0.96$, $\mu_V=1.56$ and $\sigma_V=0.50$. The normalized ratings are, thus, $A=0.83$ and $V=-3.12$.

2) The semantic description of the subject, according to the reference ontology, is composed. According to fuzzy concepts obtained previously, normalized ratings are both in the low range, so a semantic description is expressed as:

   SubjectId: 9  \( \bigwedge \) hasArousal.low \( \exists \) hasArousal \( \exists \) hasValence.low \( \exists \) hasValence

3) Annotated dataset and concept descriptions learned by FuzzyDL-Learner are then fed to the matchmaker in order to detect the subject’s emotion(s). In the case under examination, ranked penalty obtained from the semantic matchmaking process are: Amusement:16.18, Fun:27.27, Excitement:36.36, Disgust:42.86, Fear:45.45, Shock:57.18, Anger:60.53. Amusement has the lowest semantic distance and therefore the best matching emotion.

4) Finally, based on the elicited emotion, the most suitable feedback could be applied in order to improve the emotional condition or prevent harmful health states of the passengers.

V. EXPERIMENTS

The experimental setup used to test the accuracy of our implementation consisted of a smaller sample number than DECAF [17]. 30 subjects were involved in the experiment watching 20 emotional videos, making a total of 600 individual records. The chosen video clips were shown in random order eliciting 7 emotions, namely amusement, anger, disgust, excitement, fear, fun ad shock.

Table II shows the confusion matrix of emotions classification. On a total of 600 instances, 373 were correctly classified with an accuracy of 62.17%. A graphical representation of the results is shown in Fig. 3. The overall weighted classification precision, recall and F-Measure are 0.735, 0.660 and 0.695 respectively.

A relevant issue is the user subjectivity associated with emotional perception: values assigned to a given impression may be subjective. For instance, Fun emotion has been mistakenly classified as Amusement because both are pleasant and not aroused emotions. For this reason, tolerance is a crucial factor to be considered. In summary, our preliminary results reveal that the semantic-based approach in conjunction with fuzzy ontology-based approach seems to be a promising route towards improving standard machine learning based emotion classification techniques.

VI. CONCLUSION AND FUTURE WORK

The paper presented early work on a novel framework for emotion recognition from biosignals via semantic annotation and matchmaking in conjunction with a fuzzy ontology-based approach. Raw sensor data, without any descriptive metadata, have limited use as they are hard to discover, integrate or interpret. One challenge is to develop and test a framework for expressing and classifying complex patterns from biosignals, allowing emotion recognition through emotional model. To this end, FuzzyDL-Learner extracts fuzzy emotional concepts creating a fuzzy OWL ontology. Then, by exploiting a matchmaker, the semantic descriptions of the test sample are compared with annotations contained in the fuzzy ontology. The matchmaker returns then the most similar emotion as output.

<table>
<thead>
<tr>
<th>Real/Predict</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>102</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>b</td>
<td>5</td>
<td>111</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c</td>
<td>19</td>
<td>7</td>
<td>21</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>d</td>
<td>4</td>
<td>15</td>
<td>3</td>
<td>36</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>e</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>f</td>
<td>41</td>
<td>21</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>46</td>
<td>2</td>
</tr>
<tr>
<td>g</td>
<td>11</td>
<td>12</td>
<td>2</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>17</td>
</tr>
</tbody>
</table>

Figure 3. Classification result.
The proposed approach is currently under prototypical implementation. Experimental evaluation on proper testbeds is ongoing and will allow to assess effectiveness w.r.t. the state of the art in AC. A further endeavor is validating the reference dataset quality and improving the accuracy of the proposed framework.

REFERENCES