A CASE-BASED DECISION SUPPORT SYSTEM FOR INDIVIDUAL STRESS DIAGNOSIS USING FUZZY SIMILARITY MATCHING

SHAHINA BEGUM, MOBYEN UDDIN AHMED, PETER FUNK, NING XIONG, BO VON SCHÉELE

School of innovation, design and engineering
Mälardalen University, SE-72123 Västerås, Sweden

Abstract. Stress diagnosis based on finger temperature signals is receiving increasing interest in the psycho-physiological domain. However, in practice, it is difficult and tedious for a clinician and particularly less experienced clinicians to understand, interpret and analyze complex, lengthy sequential measurements in order to make a diagnosis and treatment plan. The paper presents a case-based decision support system to assist clinicians in performing such tasks. Case-based reasoning is applied as the main methodology to facilitate experience reuse and decision explanation by retrieving previous similar temperature profiles. Further fuzzy techniques are also employed and incorporated into the case-based reasoning system to handle vagueness, uncertainty inherently existing in clinicians reasoning as well as imprecision of feature values. Thirty nine time series from 24 patients have been used to evaluate the approach (matching algorithms) and an expert has ranked and estimated similarity. On average goodness-of-fit for the fuzzy matching algorithm is 90% in ranking and 81% in similarity estimation which shows a level of performance close to an experienced expert. Therefore, we have suggested that a fuzzy matching algorithm in combination with case-based reasoning is a valuable approach in domains where the fuzzy matching model similarity and case preference is consistent with the views of domain expert. This combination is also valuable where domain experts are aware that the crisp values they use have a possibility distribution that can be estimated by the expert and is used when experienced experts reason about similarity. This is the case in the psycho-physiological domain and experienced experts can estimate this distribution of feature values and use them in their reasoning and explanation process.

Keywords: case-based reasoning, fuzzy logic, decision support system, classification, diagnosis.

1. INTRODUCTION

Stress is a common problem for many people in today's modern society. It is well known that increased stress level can lead to serious health problems. Stress has a side effect of reducing awareness of bodily symptoms and people often on a heightened level of stress may not be aware of it and first notice it weeks or months later when the stress is causing more serious stress related effects in the body and health (Von Schéele and Von Schéele 1999). Severe stress during long periods is highly risky or even life-endangering for patients with e.g. heart disease or high blood pressure. A computer-aided system that helps early detection of potential stress problems would bring vital benefits for treatment and recovery in both clinical and home environments.

Medical investigations have identified that finger temperature has a strong correlation with stress status for most people. Interpreting and analyzing finger temperature profiles for diagnosing severity of stress and other related dysfunctions is receiving increasing significance within the psycho-physiological domain. In doing this, clinicians are required to carefully inspect lengthy streams of measurements for capturing indicative characteristics and recognizing any possible disorders. It is a time-consuming and tedious task for humans to carry out. Further, understanding large variations of measurements from diverse patients requires knowledge and experience and without adequate support, errors of judgment could be made by a less experienced staff.

In this paper we present an approach to provide decision support for clinicians in analyzing and classifying finger temperature measurements. The aim of this research work is to help clinicians to diagnose individual stress level of a patient. The main approach is based on the use of case-based reasoning, a methodology receiving increased attention in the medical and psychological domain, e.g. as in (Bichindaritz 1996; Perner et al. 2003, Schmidt et al. 2006, and

Nilsson et al. 2006). The approach enables reuse of experience from previous cases with analyzed temperature and stress profiles. Three similarity matching functions have been established for this purpose to assess case similarity and relevance in this application scenario. The comparative performance of these three similarity functions have been evaluated empirically. 39 Measurements from 24 people have been collected and used in evaluations where the clinical expert has ranked and estimated their similarity. In order to verify the system, goodness-of-fit and absolute mean difference are calculated where the main goal of the evaluation is to see its performance in comparison to an expert's estimations. In this evaluation the suggested fuzzy similarity matching method yields the best performance concerning the rank of retrieved cases, i.e. producing a rank that is most consistent with domain expert opinions.

Fuzzy logic in combination with case-based reasoning shows some interesting features and may be of value in similar medical applications. Fuzzy theory has proved a powerful tool for representing and dealing with imprecision, vagueness and ambiguity arising from measurements, judgments and concepts. By using fuzzy set theory we can achieve more soft distinctions for making decisions that have a closer accordance with the experts' results. In (Burkhard and Richter 2000) it was identified that the central notion of similarity in CBR could be treated as fuzzy relation and a composite similarity measure could be constructed via fuzzy operations. Currently we have integrated fuzzy techniques into our system in such aspect where every crisp case index is fuzzified into a set of fuzzy sets for fuzzy matching, which makes similarity assessment more robust against known possibility distributions in values given by humans, noise and measurement errors.

The paper is organized as follows; section 2 gives an overview of our system being developed together with relevant background knowledge. Related work is outlined in section 3. Then in section 4 we explain the details of feature extraction from finger temperature signals. Section 5 presents three matching functions for retrieving and ranking similar cases, which is followed by reuse and retain schemes in section 6. The relative performance of three matching function is then evaluated and presented in section 7. Finally the paper is concluded by section 8 with summary and discussion.

2. METHOD AND SYSTEM OVERVIEW

Clinical studies show that finger temperature (FT) generally decreases with stress; however this effect of changes is very individual. The pattern of variation within a finger temperature signal could help to determine stress-related disorder. However, interpreting a particular curve and diagnosing stress level is difficult even for experts in the domain. In the proposed system, we use Case-based Reasoning (CBR) as it works well in such domain where the domain knowledge is not clear enough as in the psycho-physiological domain where even an experienced clinician might have difficulty expressing his knowledge explicitly. A fuzzy set theory is used to compose efficient matching method for finding most relevant cases by calculating similarities between cases. So, combinations of all such artificial intelligent techniques help us to build a computer-aided decision support system for diagnosing stress-related disorder and severity level of the disorder.

2.1. Case-Based Reasoning

A case-based reasoning (CBR) (Aamodt and Plaza 1994; Watson 1997) method can work in a way close to human reasoning e.g. solves a new problem applying previous experiences. A clinician/doctor may start his/her practice with some initial experience (solved cases), then try to utilize this past experience to solve a new problem and simultaneously increases his/her case base. So, this method is getting increasing attention from the medical domain since it is a reasoning process that also is medically accepted. CBR has shown to be successful in a number of different medical applications (Nilsson and Sollenborn 2004). Aamodt and Plaza has introduced a life cycle of CBR (Aamodt and Plaza 1994) with four main steps as shown in Figure 1. Retrieve, Reuse, Revise and Retain present key tasks to implement such kind of cognitive model.
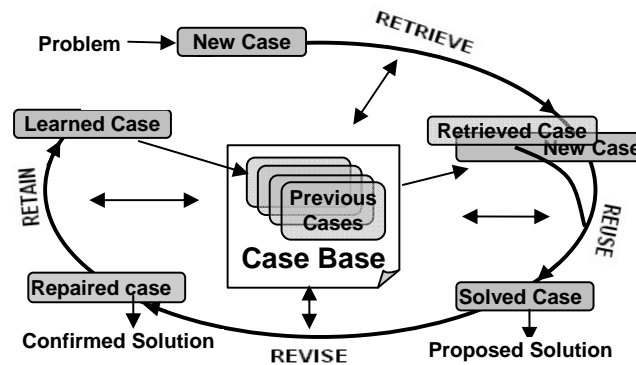
FIGURE 1. CBR cycle. The Figure is introduced by Aamodt and Plaza (Aamodt and Plaza 1994)

In the retrieval step, for any new problem the system tries to retrieve the most similar case(s) by matching previous cases from a case base. If it finds any suitable case that is close to a current problem then the solution is reused (after some adaptation and revision if necessary). A clinician may revise the selected case with solution and retain this solution along with the new problem into the case base. The CBR method in the proposed system is used to suggest recommendations for diagnosis of stress-related disorder for a new case by retrieving and matching previously solved similar problems from the case base.

2.2. Fuzzy Logic and Case-Based Reasoning

Fuzzy set theory has successfully been applied in handling uncertainties in various application domains (Jang, Sun, and Mizutani. 1997) including medical domain. Inexact medical entities can be defined using fuzzy sets. Fuzzy set theory was developed by Zadeh in 1965. It explains fuzziness existing in a human thinking process using fuzzy values instead of using a crisp or binary value. Use of fuzzy logic in medical informatics has begun in the early 1970s. In fuzzy CBR, fuzzy sets can be used in similarity measure (Bonissone and Cheetham 1998; Dvir, Langholz and Schneider 1999; Wang 1997). A discussion about the relationship between the similarity concept and several other uncertainty formalisms including fuzzy sets can be found in (Richter 2006). In the proposed application, fuzzy set theory is used for matching similarities between existing cases and a current case to model imprecise expert's knowledge in the psycho-physiological domain. It matches cases in terms of degrees of similarities [0-1] between attribute values of previous cases and a new case.

2.3. System Overview

A decision support system for diagnosing individual stress-condition based on finger temperature measurements works in several stages as illustrated in Figure 2. The first stage is the Calibration phase (Begum et al. 2006a) where the finger temperature measurement is taken using a temperature sensor to establish an individual stress profile. Feature extraction is the second stage described in section 4 where relevant features are extracted automatically from the outcome of the calibration phase. Finally, these extracted features thereafter help to formulate a new problem case and passed to the case-based reasoning cycle. The new case is then matched using different matching algorithms including modified distance function, similarity matrix and fuzzy similarity match, see details in section 5. The DSS can provide matching outcome in a sorted list of best matching cases according to their similarity values in three circumstances: when a new problem case is matched with all the solved cases in a case base (between subject and class), within a class where the class information is provided by the user and also within a subject, see more details in section 6.

Test step	Observation time	Con/Parameter	
1.	3 min	Base Line	
2.	2 min	Deep Breath	
3.	2+2 min	Verbal Stress	
4.	2 min	Relax	
5.	2 min	Math stress	
6.	2 min	Relax	

Measuring finger temperature of a stressed person in the **calibration phase**

Retrieving & parsing file

Finger temperature measurements along with other features, both from measurement and given by clinician

Data stored in a file

Confirmed solution

Feature extraction

Decision Support System for StressDiagnosis

Retain new case for further use

Case-base

Additional values from observation and final revision

Doctor/Clinician

Proposed solution

Retrieve solved case for matching

Similarity Matching Algorithms
* Fuzzy Similarity,
* Expert defined Matrix &
* Distance Function

Old solved case

New problem case
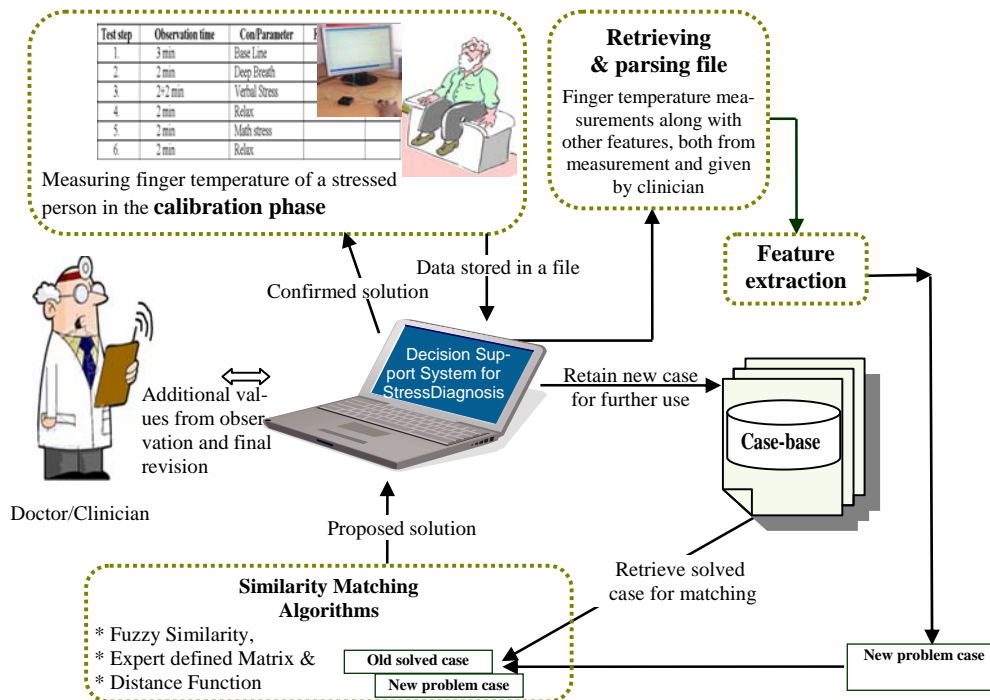
New problem case

FIGURE 2. System overview of a decision support system for stress diagnosis

A clinician thereafter revises the best matching cases and approves a case to solve a new problem case by using the solution of this old case; this confirmed solution is then prescribed to the patient. However, often an adjustment to the solution of the old case may require since a new problem case may not always be as same as an old retrieved case. However, there is no adaptation of the cases in the proposed system. This adaptation could be done by clinicians in the domain. In the medical system, there is not much adaptation, especially in a decision support system where the best cases are proposed to the clinician as suggestions of solutions and when the domain knowledge is not clear enough (Watson 1997). Finally, this new solved case is added to the case base functioning as a learning process in the CBR cycle and allows the user to solve a future problem by using this solved case, which is commonly termed as retain. Retaining of a new solved case could be done manually based on clinician or expert's decision. The decision support system is currently implemented as a prototype in Java so it is platform independent. An evaluation of the system performance compared to a domain expert/clinician is presented in section 7.

3. RELATED WORK

A procedure for diagnosing stress-related disorders has been put forward by Nilsson et al. (2006) according to which stress-related disorders are diagnosed by classifying the heart rate patterns analyzing both cardio and pulmonary signals, i.e., physiological time series and used as a research tool in psycho-physiological medicine. This was an initial attempt to use a DSS in a previously unexplored domain e.g. psycho-physiological medicine. This tool is more suitable to use in clinical environment whereas the DSS, diagnosing stress-related disorder analyzing the finger temperature signal, proposed in this paper can be developed as a tool to be used by people who need to monitor their stress level during everyday situations e.g. in home and in work environment for health reasons.

In our previous work (Begum et al. 2006a), a stress diagnosing system using CBR has been designed based only on the variation of the finger temperature measurements, but this previous research does not addressed whether any other factors that could also be used in diagnosing individual stress level. In the earlier research (Begum et al. 2007) we have further demonstrated a system for classifying and diagnosing stress level, exploiting finger temperature graphs and other features. This system relies on CBR as well as on fuzzy sets theory. In extracting features from FT signal we have considered step 3, 4 and 5 (calibration phase, see Begum et al. 2006a) and investigated the tempera-

ture variation of these steps. The current paper presents a result of the evaluation of a computer-aided stress diagnosis system in comparison to a domain expert/clinician. In this system fuzzy similarity matching is applied in CBR-retrieval. In addition, in extracting features from signal data we have considered step 2 to step 6 of the calibration phase.

Apart from the psycho-physiological domain, CBR has been applied in several others diagnosis/classification tasks in the medical domain. Montani et al. (2001) has combined case-based reasoning, rule-based reasoning (RBR), and model-based reasoning to support therapy for diabetic patients. Auguste (Marling and Whitehouse 2001) project has been developed for diagnosis and treatment planning in Alzheimer's disease. This is a hybrid system that combines CBR and RBR. MNAOMIA (Bichindaritz 1996) has been developed for the domain of psychiatry. CARE-PARTNER (Bichindaritz, Kansu and Sullivan 1998) is a decision support system developed in stem cell transplantation. The system uses a multi modal reasoning framework combining CBR and RBR. BOLERO (Lopez and Plaza 1993) is a successfully applied medical CBR diagnosis system in diagnosing pneumonias applies fuzzy set theory for representing uncertain and imprecise values. A CBR technique with fuzzy theory has been used for the assessment of coronary heart disease risk in (Schuster 1997). A CBR approach to dose planning in Radiotheraphy has been proposed by Song et al. in (2007) where fuzzy set theory is applied for measuring the similarity. A CBR system for cancer diagnosis has been proposed by (Diaz, Florentino, and Corchado 2006) which combine fuzzy case representation, a neural network to cluster the cases and a set of rules for the classification. All these projects and others (Gierl 1993, Schmidt and Gierl 2002, and Perner et al.2003) show significant evidence of successful implementations of the CBR techniques in the medical domain. Nevertheless, the application of CBR in the psycho-physiological domain has been limited so far. Therefore, to our knowledge, research work addressed in this paper for providing decision-support to clinicians in the psycho-physiological medicine is of great significance in applying CBR and other artificial intelligence techniques in medical domain.

4. FEATURES EXTRACTION AND CASE FORMULATION

Extracting appropriate features is of great importance in performing accurate classification in a computer-aided system whereas in manual process an experienced clinician often classify FT signal without being pointed out intentionally all the features he/she uses in the classification. A standard procedure followed by clinicians to establish a person's stress profile has already been discussed concerning the calibration phase (Begum et al. 2006a) whereby an experienced clinician manually evaluates the FT measurements during different stress conditions as well as in non-stressed (relaxed) conditions to make an initial diagnosis. In this phase, the finger temperature is measured using a temperature sensor connected to a computer and the temperature is observed in 6 steps (1. Baseline, 2. Deep-breath, 3. Verbal-stress, 4. Relax with positive thinking, 5. Math-stress and 6. Relax). After the test, a person is requested to answer some questions for instance, when he/she had his/her meal, food habit, food allergy and so on. The output from the calibration phase is then used in extracting significant features and afterwards a new case is formulated employing these extracted features.

The FT sensor measurements are recorded using software which provides filtered data to the system. This signal data and answer to the questions from the calibration phase are then stored in a file in the local device and exported to the DSS. From the exported file, system retrieves 15 minutes finger temperature measurements (time, temperature) in 1800 samples, together with other numeric (age, room-temperature, hours since meal, etc) and symbolic (gender, food and drink taken, sleep at night, etc) features. In fact, dealing with sensor signal is more complex than human designed features such as age, gender, room temperature etc.
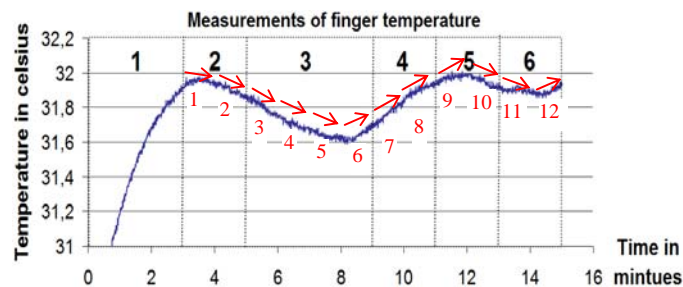
FIGURE 3. Changes in FT data against time during different stress and non-stress condition

Figure 3 displays skin temperature of the finger during both the stress and non-stress conditions. As can be seen, after analyzing a number of finger temperature signals, the temperature is rising and falling against time and after an initial increase, finger temperature decreases in stress condition (step 3) and increases in relax condition (step 4). According to closer discussion with clinicians on the interpretation of such graph, it is concluded that in general, the finger temperature could decrease with stress and increase in relax state and the changes between the steps are also of importance for the clinicians. A standardization of the slope that is using negative and positive angles makes it more visualise and gives a terminology to a clinician for reasoning about stress.

Therefore, we calculate the derivative of each step to introduce "degree of changes" as a measurement of the finger temperature changes. A low angle value, e.g. zero or close to zero indicates no change or stable in finger temperature. A high positive angle value indicates rising finger temperature, while a negative angle, e.g. -20° indicates falling finger temperature. Step1 (baseline) is used normally to stabilize the finger temperature before starting the test hence this step has not been considered and the clinician also agreed on this point. Each step is divided by one minute time interval (4 minutes step3 is extracted as 4 features) and each feature contains 120 sample data (time, temperature). Thus 12 features are extracted from the 5 steps (step 2 to 6) and named as *Step2_Part1*, *Step2_Part2*, *Step3_Part1*, ………, *Step6_Part1*, *Step6_Part2* as shown in Figure 3. First, a slope of the linear regression line has been calculated through the data points, as y is temperature (in Celsius) and x is time (in minute) by equation 1 for each extracted feature from the measurement.

$$slope_f = \frac{\sum_{i=0}^{n}(x - \bar{x})(y - \bar{y})}{\sum_{i=0}^{n}(x - \bar{x})^2}$$

(1)

Where *f* denotes the number of features (1 to 12 see Figure 4), *i* is the number of samples (1 to 120) and $\bar{x}, \bar{y}$ is average of the samples. Then this slope value is converted to arctangent as a value of angle in radians (*-pi/2* to *+pi/2*) and finally expressed arctangent value in degrees by multiplying *180/PI*. The converting function from radians to degree is described in equation 2 where *PI* is 3.14 as a standard value. So these 12 features contain degree values comprising 120 sample data (time, temperature). Instead of keeping the sample data these degree values are used or represented as features.

$$degree_f = \left[\tan^{-1}(slope_f)\right] * \frac{180}{PI}$$

(2)

Five other features which have also been extracted from the sensor signal are *start temperature* and *end temperature* from step2 to step6, *minimum temperature* of step3 and step5, *maximum temperature* of step4 and step6, and *difference between ceiling and floor*. Finally, 17 (12+5) features are extracted automatically from the fifteen minutes (1800 samples) FT measurements signal data. Then a new case is formulated with 19 features as a total keeping in a vector above 12 features and adding *hours since last meal and gender*.

The DSS thereafter formulates a new problem case combining this generated extracted features and human defined features. This new formulated case is then applied in diagnosing stress and making treatment plan by using the CBR cycle.

5. CASE RETRIEVAL AND MATCHING

Case retrieval is the major phase in CBR cycle where matching between two cases plays vital role because near-most or most relevant solved cases could be retrieved if a superior matching algorithm exists. To be more cautious, the proposed DSS used three different matching algorithms and in three different matching prospects. The retrieval step is essential especially in medical applications since missing similar cases may lead to less informed decision. The reliability and accuracy of the diagnosis systems depend on the storage of cases/experiences and on the retrieval of all relevant cases and their ranking. To solve and store a new case the DSS used 19 features in total. Of which 12 features are *Step2_Part1*, *Step2_Part2*, *Step3_Part1*, ………, *Step6_Part1*, *Step6_Part2* and 7 other features are *start temperature* and *end temperature* from step2 to step6, *minimum temperature* of step3 and step5, *maximum temperature* of step4 and step5, *difference between ceiling and floor, Hours since last meal and gender*.

In the DSS three implemented matching algorithms are 1) modified distance function for calculating similarity where distance between two cases are used as similarity value 2) similarity matrices defined by the expert where distance between two cases are converted into similarity values using matrices and 3) fuzzy set theory to calculate similarity between two cases. Similarity measurement is taken to assess the degrees of matching and create the ranked list containing the most similar cases retrieved by equation 3

$$Similarity \quad (C,S) = \sum_{f=1}^{n} w_f * sim(C_f, S_f)$$
(3)

Where C is a current/target case, S is a stored case in the case base, w is the normalized weight defined by equation 4, n is the number of the attributes/features in each case, f is the index for an individual attribute/feature and sim (C_f, S_f) is the local similarity function (see sections 6.1, 6.2 and 6.3) for attribute f in cases C and S.

$$w_f = \frac{lw_f}{\sum_{f=1}^{n} lw_f}$$
(4)

Here, the *Local weight (lw)* defined by experts, assumed to be a quantity reflecting importance of the corresponding feature, *Normalized weight (w)* is calculated by equation 4. Generally there are two ways to specify the values of weights for individual features. One way is to define weights by experts in terms of domain knowledge, while the other is to learn or optimize weights using the case library as information source. In this project, both of these approaches have been implemented to create suitable weight values. The performance of both expert weights and automatic weights (learnt from the case base) in similarity evaluation will be evaluated in section 7.

5.1. Modified Distance Function

Distance between the features of two cases (C, S) can be calculated by one dimensional Euclidean distance function. Hence all the symbolic features are converted into numeric values before calculating the distance for example, for a feature 'gender' male is converted to one (1) and female is two (2). However, we normalized the distance values from 0 to 1 using equation 5, where 1 indicates no distance and 0 far away (largest distance).

$$sim(C_f, S_f) = 1 - \frac{abs(C_f - S_f)}{Max(C_f, S_f) - Min(C_f, S_f)}$$
(5)

Function sim (C_f, S_f) in equation 5 represents local similarity and function *abs* is used to get an absolute value of (C_f - S_f). *Max* retrieves the maximum feature value for a feature f between the whole case base and a query case C and *Min* retrieves the minimum feature value for a feature f between the whole case base and a query case C.

5.2. Similarity Matrix

For the numeric features, distance between two features is calculated through the one dimensional Euclidean distance function. After calculating the distance, this value is converted using the local similarity values as depicted in Table 1 where the similarity values for different features are defined by a domain expert. But the similarity between two symbolic features is calculated directly using matrix without calculating the distance. For example, similarity between same genders is defined as 1 otherwise 0.5, as can be seen from Table 1.

TABLE 1. Example of the expert defined matrices used to calculate similarity

Similarity for step		Similarity for ceiling/floor		Hours since last meal						Similarity for gender	
Distance	sim		sim	T/ S	0	1	2	3	>4	m	f
0-2 degree	1	<0.3	1	0	1	0.8	0.6	0.4	0	m 1	0.5
>2 and <4	0.8	0.3-0.5	0.8	1	0.8	1	0.8	0.6	0.4	f 0.5	1
>4 and <6	0.6	0.5-0.7	0.6	2	0.6	0.8	1	0.8	0.6		
>6 and <8	0.4	0.7-0.9	0.4	3	0.4	0.6	0.8	1	0.8		
>8 and <10	0.2	0.9-1.1	0.2	>4	0	0.4	0.6	0.8	1		
>10	0	> 1.1	0								

5.3. Fuzzy Similarity

Many crisp values both from measurements and given by a clinician are known to have a possibility distribution often known by experts and used in their reasoning. We propose that this dimension and domain knowledge is represented by fuzzy similarity, a concept well received by clinical experts. Representation of a similarity value using a matrix as shown in Table 1 often shows a sharp distinction which may provide an unreliable solution in domains where it is known that these values are less exact. Fuzzy similarity matching reduces this sharp distinction. After discussions with clinical experts a triangular membership function (*mf*) replaces the crisp input feature with a membership grade of 1. For instance, as shown in Figure 4 a current case has the lower and upper bounds 2.5 and 7.5 represented by an *mf* of grade 0 and an input value 5 is represented by an *mf* grade of 1 (fuzzy set *m1*). Again an old case has the lower and upper bounds -1.5 and 4.5 represented by an *mf* grade of 0 and an input value 3 is represented by an *mf* grade of 1 (fuzzy set *m2*). In both cases, the width of the *mf* is fuzzified by 50% in each side. Then by applying fuzzy intersection between the two fuzzy sets *m1* and *m2* we get a new fuzzy set *om* which represents the overlapping area between *m1* and *m2*.
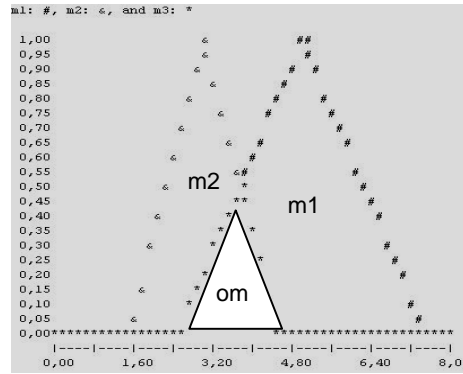

FIGURE 4. Fuzzy similarity using triangular membership functions

Similarity between the old case and the new case is now calculated using equation 6 where area of each fuzzy set (*m1*, *m2* and *om*) is calculated. The similarity equation according to (Dvir et al. 1999) is defined as-

$$sim(C_f, S_f) = s_f(m1, m2) = \max(om/m1, om/m2) \tag{6}$$

Where, $s_f(m1, m2)$ calculates similarity between two features *f* of new and old cases. When the overlapping area (*om*) is bigger then the similarity between two features will also increase and for two identical fuzzy sets the similarity will reach unity.

6. REUSE AND RETAIN

The objective of this implemented system is the diagnosis of an individual's stress condition where the main functionality lies in solving a new problem case by using solution of past solved cases. However, solution of a past case often requires adaptation to find a suitable solution for the new case. This adaptation might often be a combination of two or more solutions of cases from the retrieved cases. Specially, in medical domains the domain knowledge is often not well understood as in circumstances of diagnosing stress related to psycho-physiological issues. Therefore, retrieving a single matching case as a proposed solution may not be sufficient for the DSS in this domain. So, the proposed system retrieved a list of ranked cases in three matching circumstances shown by the indicators 1, 2 and 3 in Figure 5. The three yielded matching circumstances are: 1) a ranked list by the system for a current/new case matching with all the other cases in a case base 2) a sorted list of matched cases that matches a current/new case with the same subjects'/patients' cases and 3) presented best matched cases when a new problem case is matched with the solved cases in the same class where case-class is given by the user. In all the circumstances ranked list of cases are presented on the basis of their similarity value and the identified class. The solution for a retrieved old case, that is diagnosis and treatment suggestions, are also shown using indicator 5 in Figure 5. Indicator 6 shows comparison of FT measurement between a new case and old case where FT values are plotted through line chart and user can use different matching algorithms by selecting specific method shown by indicator 7. It can be seen using indicator 4 in Figure 5, details of the matching information for a new case with an old case is displayed thereby clinicians/users get an opportunity to see more details of the matching cases which may help to determine if the solution is reusable or require an adaptation for a new problem.
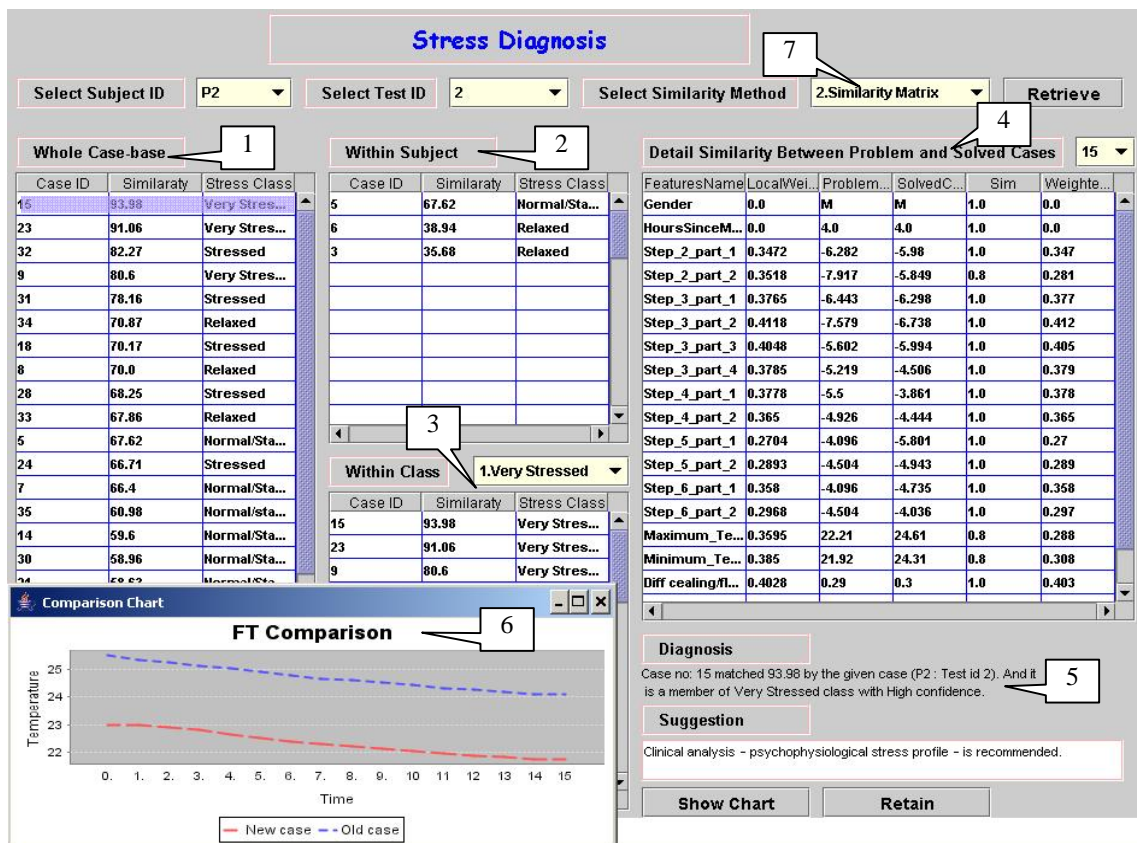

FIGURE 5. A screenshot of the stress diagnosis system.

Users can adapt solutions i.e. it could be a combination of two solutions from the list of retrieved and ranked cases in order to develop a solution to the problem in the new case. Then clinician/expert determines if it is plausible solution to the problem and he/she could modify the solution before approved. Then the case is sent to the revision

step where the solution is verified manually for the correctness and presented as a confirmed solution to the new problem case. In the retention step, this new case with its verified solution can be added to the case base as a new knowledge.

7. EVALUATION

After implementing the proposed DSS, performance of the system has been evaluated where the evaluation is conducted on the similarity matching. System performance in terms of accuracy has been compared with experts in the domain where the main goal is to see how close the system could work compared to an expert. The accuracy of the system as compared to the expert is calculated using a statistics *square of the correlation coefficient* or *Goodness-of-fit* (R^2) (Carol 2002). *Absolute mean difference* is also calculated to determine the deviation between expert and the system. The case base is initialized with 39 reference cases classified by the domain expert and the classification of sensitivity to stress has been denotes as Very Relaxed, Relaxed, Normal/Stable, Stressed and Very Stressed.

7.1. Similarity Matching

We have discussed in the earlier section (section 5) about the three matching algorithms implemented in this system and now the performance of these algorithms is evaluated in this section. For the evaluation we have chosen randomly three subsets of cases and three query cases, the subsets are as follows: 1) Set A: {7 cases} with query case id 4, 2) Set B: {11 cases} with query case id 16 and 3) Set C: {10 cases} with query case id 28. All the three sets have been sorted according to the similarity with the query case decided by a domain expert (human reasoning). The sorted cases are then converted to the rank numbers, i.e., the position of a case in the ranking. Likewise, the evaluation process is designed for the three algorithms including distance, matrix, and fuzzy matching (see details in section 6), used in the system. Top six cases from each set according to the expert's ranking are used as standard for the evaluation process where both the similarity values and the ranking numbers are considered; one example evaluation result for case Set A is shown in Table 2.

TABLE 2. Similarity matching for Set A with case id 4 in comparison with a clinical expert

Matching (Query, Set A)	Expert		Modified Distance		Similarity Matrix		Fuzzy Similarity	
	Ranking	Similarity (%)	Ranking	Similarity (%)	Ranking	Similarity (%)	Ranking	Similarity (%)
4, 15	1	96	1	95	1	93	1	94
4, 23	2	95	2	94	2	90	2	89
4, 9	3	94	4	87	4	79	3	84
4, 24	4	75	5	84	6	65	4	74
4, 31	5	70	3	92	3	80	5	70
4, 7	6	65	6	80	5	67	6	65
Goodness-of-fit			0.69	0.43	0.51	0.60	1.00	0.94
Absolute Mean Difference			0.67	9	1.00	8	0	3

In the table above, the 1st column describes identification of two matching cases (Query and Set A). For instance query case id 4 is matched with case id 15 in Set A. Gray coloured columns represent the position of each case ranked by the expert and other three algorithms. The rest of the columns display the similarity value of each case both by the expert and the system using the three algorithms. Last two rows show the value of goodness-of-fit (R^2) and absolute (ABS) mean difference, calculated on the basis of the ranking and similarity values identified by the expert and the system. According to the R^2 value and absolute mean difference both in ranking and in similarity, fuzzy similarity matching algorithm shows better performance than the other algorithms on the example Set A compared with the expert's opinion.

TABLE 3. Average Goodness-of-fit and absolute mean difference for three matching algorithms

Similarity Algorithms	Goodness-of-fit	Absolute Mean Difference

in Average (Set A, B and C)	Ranking	Similarity (%)	Ranking	Similarity (%)
Modified Distance	0.52	0.38	1.00	9.33
Similarity Matrix	0.43	0.33	1.00	8.00
Fuzzy Similarity	0.90	0.81	0.33	11.67

Table 3 shows the average outcome across the three subsets: Set A, Set B and Set C in terms of the goodness-of-fit (R^2) and absolute mean difference for evaluating three algorithms (Using distance, matrix and fuzzy-logic), comparing expert's ranked cases with the cases ranked by the system and with their similarity value as well. As can be seen from Table 3, similarity matching algorithm using fuzzy logic seems to be reliable both in similarity and ranking value and it outperforms the other two matching algorithms. According to the calculated average R^2 between the yielded (using fuzzy-logic) and the desired values (using domain expert) for the ranking and similarity assessments across three case subsets, it is showed that the fuzzy matching can yield results coinciding with expert suggestions with 90% in ranking and 81% in similarity evaluations. Regarding the average absolute mean difference on the three sets (A, B and C) fuzzy-logic has lower error (i.e. 0.33) in ranking compared to others although the error in similarity (i.e. 11.67) is higher than the others. The error may be due to the fact that the fuzzy similarity matching algorithm depends on the width of the fuzzy set membership functions (*mf*) and in our system the width of the *mf* is fuzzified by 50% in each side. Overall, using fuzzy logic the proposed system can retrieve more relevant cases which have been chosen by an expert/clinician comparing the other two algorithms.

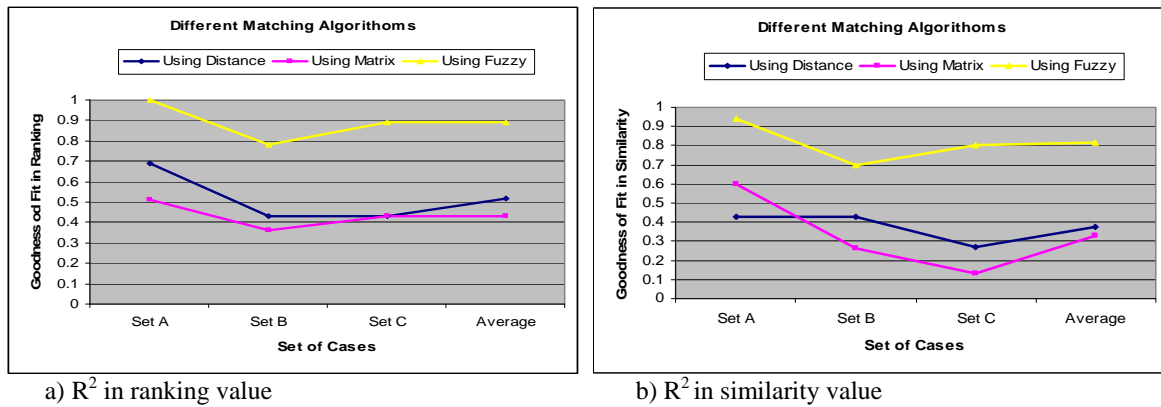a) R^2 in ranking value

b) R^2 in similarity value

FIGURE 6. Comparison among three different matching algorithms.

Comparison charts of the three matching algorithms using the three sets according to their goodness-of-fit (R^2) is presented in Figure 6, where a) shows calculated R^2 for ranking values and b) shows R^2 for similarity values.

Until now all the results we have discussed are based on weights defined by domain expert. As another alternative we also attempted to discover proper weight values by learning from the case base. The basic idea followed is to distinguish individual features in terms of discriminating powers (Funk and xiong 2007) on the discritized universes of features. The weight of an individual feature is simply defined to equal the metric of discriminating power, which further can be estimated using samples in the case base. The performance of such automatic learnt weights is demonstrated in Table 4 as compared to expert results in similarity evaluation and case ranking using fuzzy similarity matching algorithm.

TABLE 4. Performance of automatic weighting.

Similarity (Query, Case Sets)	Goodness-of-fit in Ranking	Goodness of Fit in Similarity	Absolute Mean Difference in Ranking	Absolute Mean Difference in Similarity (%)
4, Set A	0.89	0.92	0.33	4
16, Set B	0.60	0.55	1.00	21
28, Set C	0.79	0.58	0.67	18

Average	0.84	0.68	0.67	14.67

We see from table above that, with 14.67% as the mean difference in similarity and 0.67 as the mean position error in ranking, the automatic weights are satisfying by producing good results close to expert evaluations. This also indicates that weight learning from the case base is a feasible solution that would help when domain knowledge is not available.

8. SUMMARY AND DISCUSSION

This paper presents a computer-aided decision support system for analyzing and diagnosing stress-related disorders based upon finger temperature signals. Our work to date features three main points, namely feature extraction from time series data, case-based reasoning, and fuzzy information processing. Feature extraction is tasked to "dig out" key characteristics from original signals to reach a concise yet sufficient description of problems. The success for this heavily relies on domain knowledge and 19 time-based features have been identified and confirmed through cooperation with domain experts. The method of case-based reasoning is employed to make recommendations for stress diagnosis by retrieving and comparing with previous similar cases in terms of features extracted. Moreover, fuzzy techniques are incorporated into our CBR system to better accommodate uncertainty in clinicians reasoning as well as imprecision in case indexes. All such ideas have been implemented and validated in a prototypical system.

Feature weighting is another important issue under investigation in our project. With available test data we have recognized that the extracted features have different importance and proper weightings for them plays a crucial role for system performance.

So far we have two sets of weight values, both of which offered acceptable system performance in evaluation. One of such weight sets was exclusively defined by an experienced domain expert, and the other set was learnt from the case base by applying the so called discriminating power (Funk and Xiong 2007) on discretized universes of individual features. The automatic learnt weights have shown to perform sufficiently close to an expert in identifying similar cases, sufficiently good bearing in mind that different expert have different opinions and that there is no exact answer. We conjecture there would be two reasons for this inferiority. The first lies in the fact that there are merely 39 cases in the current case library and this low number of samples may degrade the reliability of weights achieved. The second and possibly more important is the lack of expert preference information in the case base. One of our future research directions will be optimization of feature weights by directly utilizing case preferences of expert as learning signals.

References

AAMODT, A., and E. PLAZA. 1994. Case-based reasoning: Foundational issues, methodological variations and system approaches. Artificial Intelligence Communications, 7:39– 59.

BEGUM, S., M. U. AHMED, P. FUNK, N. XIONG, and B. VON SCHÉELE. 2007. Classify and Diagnose Individual Stress Using Calibration and Fuzzy Case-Based Reasoning. In proceedings of 7th International Conference on Case-Based Reasoning, Edited by Weber and Richter, Springer, Belfast, Northern Ireland, pp. 478-491

BEGUM, S., M. U. AHMED, P. FUNK, N. XIONG, and B. VON SCHÉELE. 2006a. Using Calibration and Fuzzification of Cases for Improved Diagnosis and Treatment of Stress. In Proceedings of the 8th European Workshop on Case-based Reasoning in the Health Sciences, pp. 113-122.

BEGUM, S., J. WESTIN, P. FUNK, and M. DAUGHERTY. 2006b. Induction of adaptive neuro-fuzzy inference systems for investigating fluctuations in Parkinson's disease. In Proceedings of 23rd Annual Workshop of the Swedish Artificial Intelligence Society. Edited by P. Eklund, M. Minock, H. Lindgren. Pp. 67-71

BICHINDARITZ, I., E. KANSU, and K.M. SULLIVAN. 1998. Case-based reasoning in care-partner: Gathering evidence for evidence-based medical practice. In Advances in CBR: The Proceedings of the 4th European Workshop on Case Based Reasoning, pp. 334–345.

BICHINDARITZ, I., 1996. MNAOMIA: Improving case-based reasoning for an application in psychiatry. Artificial Intelligence in Medicine: Applications of Current Technologies, Standford, CA, pp. 14–20.

BONISSONE, P., and W. CHEETHAM, 1998. Fuzzy Case-Based Reasoning for Residential Property Valuation, Handbook on Fuzzy Computing (G 15.1), Oxford University Press.

BURKHARD, H.-D., and M. M. RICHTER, 2000. On the notion of similarity in case based reasoning and fuzzy theory. In Soft Computing in Case-Based Reasoning, Edited by Sankar, Tharam and Daniel, Springer-Verlag, London, UK , pp. 29 - 45

CAROL, C.H. 2002. Goodness-Of-Fit Tests and Model Validity. Birkhäuser, ISBN 0817642099.

DIAZ, F., F. FDEZ-RIVEROLA, and J.M. CORCHADO 2006. Gene-CBR: A Case-Based Reasoning Tool for Cancer Diagnosis Using Microarray Data Sets. Computational Intelligence, Vol. 22, pp. 254-268.

DVIR, G., G. LANGHOLZ, M. SCHNEIDER, 1999. Matching attributes in a fuzzy case based reasoning. Fuzzy Information Processing Society, pp. 33–36.

FUNK, P., and N. XIONG. 2007. Extracting knowledge from sensor signals for case-based reasoning with longitudinal time series data. Case-Based Reasoning on Images and Signals. *Edited by* Petra Perner, Springer Verlag, pp. 247-284.

GIERL, L., 1993. ICONS: Cognitive basic functions in a case-based consultation system for intensive care. *In* Proceedings of Artificial Intelligence in Medicine. Andreassen S et al., eds., pp. 230-236.

JANG J.S.R., C.T. SUN, and E. MIZUTANI. 1997. Neuro-fuzzy and Soft Computing. A computional approach to learning and machine intelligence. Prentice Hall, NJ. ISBN 0-13261066-3.

LOPEZ, B., and E. PLAZA. 1993. Case-based learning of strategic knowledge Machine Learning EWSL-91, Lecture Notes in Artificial Intelligence, *edited by* Kodratoff, Springer-Verlag, pp.398-411.

MARLING, C., and P. WHITEHOUSE. 2001. Case-based reasoning in the care of Alzheimer's disease patients. *In* Case-Based Research and Development, pp.702–715.

MONTANI, S., P. MAGNI, A.V. ROUDSARI, E. R. CARSON, and R. BELLAZZI. 2001. Integrating Different Methodologies for Insulin Therapy Support in Type 1 Diabetic Patients. *In* proceedings of the 8th Conference on Artificial Intelligence in Medicine. Springer. pp.121-130.

NILSSON, M., P. FUNK, E. OLSSON, B. H. C. VON SCHÉELE, and N. XIONG. 2006. Clinical decision-support for diagnosing stress-related disorders by applying psychophysiological medical knowledge to an instance-based learning system. Artificial Intelligence in Medicine, pp. 159-176.

NILSSON, M and M. SOLLENBORN. 2004. Advancements and trends in medical case-based reasoning: An overview of systems and system development. *In* proceedings of the 17th International FLAIRS Conference, Miami Beach, Fl, pp. 178-183.

PERNER, P. 2007. Introduction to Case-Based Reasoning for Signals and Images. Case-Based Reasoning on Signals and Images. *Edited by* Petra Perner, Springer Verlag, pp. 1-24.

PERNER, P., T. GUNTHER, H. PERNER, G. FISS, and R. ERNST. 2003. Health Monitoring by an Image Interpretation System- A System for Airborne Fungi Identification. *In* Proceedings of the 4th International Symposium on Medical Data Analysis, SMDA '03, Springer, pp. 64-77.

PLAZA, E. and J-L. ARCOS. 1993. A reactive architecture for integrated memory-based learning and reasoning. *In* Proceedings of the First European Workshop on Case-Based Reasoning, ed Richter, Wess, Altho and Maurer, Vol.2, pp.329-334.

PLAZA, E. and R. L. MANTARAS. 1990. A case-based apprentice that learns from fuzzy examples, Methodologies for intelligent Systems 5, Elsevier, pp. 420-427.

RICHTER, M. M. 2006. Modeling Uncertainty and Similarity-Based Reasoning - Challenges, *In* Proceedings of the 8th European Workshop on Uncertainty and Fuzziness in CBR, pp. 191-199.

RISSLAND, E. and D. SKALAK. 1989. Combining case-based and rule-based reasoning: A heuristic approach. *In* Proceedings IJCAI-89, Detroit, MI, pp. 524-530.

SCHMIDT, R., W. TINA, and G. LOTHAR. 2006. Predicting Influenza Waves with Health Insurance Data. Computational Intelligence, Vol. 22, pp. 224-237.

SCHMIDT, R., and L. GIERL. 2002. Prognostic Model for Early Warning of Threatening Influenza Waves. *In* Proceedings of the 1st German Workshop on Experience Management, pp. 39-46.

SCHUSTER, A. 1997. Aggregating Features and matching Cases on Vague Linguistic Expressions, *In* Proceedings of International Joint Conferences on Artificial Intelligence (1997) 252-257.

SONG, X., P. SANJA, and S. SANTHANAM. 2007. A Case-Based Reasoning Approach to Dose Planning in Radiotherapy. *In* Proceedings of the 8th European Workshop on Case-based Reasoning in the Health Sciences, pp. 348-357

VON SCHÉELE, B. H. C., and I. A. M. VON SCHÉELE. 1999. The Measurement of Respiratory and Metabolic Parameters of Patients and Controls Before and After Incremental Exercise on Bicycle: Supporting the Effort Syndrome Hypothesis. Applied Psychophysiology and Biofeedback, Vol. 24, pp.167-177.

WANG, W. J. 1997. New similarity measures on fuzzy sets and on elements. Fuzzy Sets and Systems, pp. 305–309.

WATSON, I. 1997. Applying Case-Based Reasoning: Techniques for Enterprise Systems. Morgan Kaufmann Publishers Inc, 340 Pine St, 6th floor, San Fransisco, CA 94104, USA.