From Tinnitus Data to Action Rules and Tinnitus Treatment

Xin Zhang[1], Zbigniew W. Ras[2,4], Pawel J. Jastreboff[3], Pamela L. Thompson[2,5]
[1] University of North Carolina, Department of Mathematics and Computer Science, Pembroke, NC 28372 USA,
[2] University of North Carolina, Department of Computer Science, Charlotte, NC 28223, USA,
[3] Emory University School of Medicine, Tinnitus & Hyperacusis Center, Atlanta, GA 30322, USA,
[4] Polish Academy of Sciences, Institute of Comp. Science, 01-237 Warsaw, Poland
[5] Catawba College, The Ralph W. Ketner School of Business, Salisbury, NC 28144, USA
xin.zhang@uncp.edu, ras@uncc.edu, pjastre@emory.edu, pthompso@catawba.edu

Abstract

Medical records contain feature values on specific dates, which brings a very unstable sampling rate. Therefore, popular temporal features, which have been successfully applied to sound data, failed to fully describe the vital behaviors of features and factors of patient and treatment. The authors developed a flexible temporal feature retrieval system based on grouping of similar visiting patterns with connection to an action-rules engine and observed interesting and useful results about how the changes of treatment factors affect the changes of recovery.

1. Introduction

Tinnitus, so-called “ringing in the ears”, which is not related to an external acoustic source or electrical stimulation [1], affects a significant portion of the general population around the world and causes significant suffering in around 10 million people of all ages in US. More so, medical researchers have found that most young, healthy people can percept tinnitus, given sufficient exposure of low level sounds [2]. However, most tinnitus patients are left with the well-known advice: “Learn to live with it.” The development of a neurophysiological model of tinnitus [3], and a corresponding new clinical approach have created a totally new treatment for tinnitus that results in significant improvement with high rate of success [4]. This method, Tinnitus Retraining Therapy (TRT), was developed by P.J. Jastreboff and it uses a combination of low level, broad-band noise and counseling to achieve the habituation of tinnitus, that is, the patients are no longer aware of their illness, except when they focus their attention on it, and even then tinnitus is not annoying or bothersome.

On one hand, TRT has provided relief for many patients and has generated a high volume of medical data. This data is typically in arrays or even nested-arrays format, which is not a suitable data representation for traditional data mining algorithms. On the other hand, due to the fact that objective methods are lacking to detect Tinnitus symptoms, it is also interesting for the clinical doctors to be able to learn the essences of Tinnitus from the audiological evaluation data. Thus, the authors initially started to explore the relationships among complex factors of the treatment and recovery patterns in different categories of patients for the purpose of optimizing the treatment process as well as learning the essence of the Tinnitus problems.

The rest of this section will focus on the basic domain knowledge necessary to understand Tinnitus Retraining Therapy data.

Tinnitus Retraining Therapy combines medical evaluation, counseling and sound therapy to successfully treat a majority of patients. Based on a questionnaire from the patient as well as an audiological test, a preliminary medical evaluation of patients is required before beginning TRT. Sensitive data from the medical evaluation is not contained in the tinnitus database O used in our research. Much of this data would contain information subject to privacy concerns, a consideration of all researchers engaged in medical database exploration. Some medical information, however, is included in the tinnitus database such as a list of medications the patient may take and other conditions that might be present, such as diabetes.

Patient categorization is performed after the completion of the medical evaluation, the structured interview guided by special forms (Jastreboff 1999) and audiological evaluation. This interview collects data on many aspects of the patient’s tinnitus, sound tolerance, and hearing loss. The interview also helps determine the relative contribution of hyperacusis and misophonia. A set of questions relate to activities prevented or affected (concentration, sleep, work, etc.) for tinnitus and sound tolerance, levels of severity, annoyance, effect on life, and many others. All responses are included in the database. As a part of audiological testing left and right ear pitch, loudness discomfort levels, and suppressibility is determined. Based on all gathered information a patient category is assigned. A patient's overall symptom degree is evaluated based on the summation of each individual symptom level, where a higher value means a worse situation.

The TRT emphasis is on working on the principle of differences of the stimuli from the background based on the fact that the perceived strength of a signal has no direct association with the physical strength of a stimulus, using a functional dependence of habituation effectiveness model. Therefore, once a partial reversal of hyperacusis is achieved, the sound level can be increased rapidly to address tinnitus directly.

Based on the above medical records, the authors developed a flexible temporal feature retrieving system with action rules [8, 9] generation device to reveal related changes among the symptoms as well as the TRT treatments.

2. System overview

The authors developed a flexible temporal feature retrieval system based on grouping the patients of similar visiting frequencies with connection to an action-rules engine, which consists of four modules: a data grouping device, a temporal feature extraction engine, a decision tree classification device, and an action rules generation device. See Figure 1. System overview.

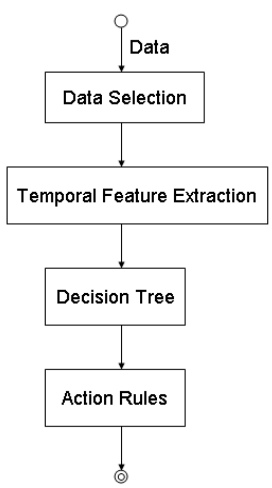

Figure 1. System overview

The data grouping device is to filter out less relevant records in terms of visiting duration patterns measured from the initial visits. The temporal feature extraction engine is to project temporal information into patient-based records for classic classifiers to learn effects of treatment as well as tinnitus upon patients. The decision tree classification device is to generate classification rules. The action rules generation device is to build action rules from certain pairs of classification rules [9].

3. Data selection and grouping

Records are grouped by similar visiting patterns, where the visiting history of each patient is discretized into durations, anchored from its initial visit date in terms of weeks, and serving as a seed for grouping. For example, a patient *p* who visited a doctor on July 8th, 2009, August 14th, 2009, and October 7st, 2009 is recorded as Table 1.

Table 1. An example of calculating visit duration

Visit ID	Duration (weeks)
1	6
2	14

The corresponding vector representation will have the form $v_p = [6, 14]$. This means that patient *p* visited the doctor five full weeks after his first visit and his last visit happened 13 weeks after his first visit (or 7 weeks after his second visit). In other words, patient p visited the doctor in the 6th week and 14th week counting in relation to his first visit. Now, assume that we have two patients denoted by *p*, *q*. Patient *p* visits are represented by a vector $v_p = [v_1, v_2, \ldots, v_n]$ whereas vector $v_q = [w_1, w_2, \ldots, w_m]$ represents visits of patient *q*. If $n \leq m$, then the distance $\rho(p,q)$ between *p*, *q* and the distance $\rho(q,p)$ between *q*, *p* is defined as

$$\rho(q,p) = \rho(p,q) = \frac{\sum_{i=1}^{n} |v_i - w_{J(i)}|}{n}, \qquad (1)$$

where $[w_{J(1)}, w_{J(2)}, \ldots, w_{J(n)}]$ is a subsequence of $[w_1, w_2, \ldots w_m]$ such that $\sum_{i=1}^{n} |v_i - w_{J(i)}|$ is minimal for all n-element subsequences of $[w_1, w_2, \ldots, w_m]$.
By $|v_i - w_{J(i)}|$ we mean absolute value of $[v_i - w_{J(i)}]$.

For example, if patient *p*, having four visits, is compared to patient *p'*, who had five visits; each of the three visits (second, third, fourth) of patient *p* shall be matched with a closest visit of *p'* and their difference shall be averaged.

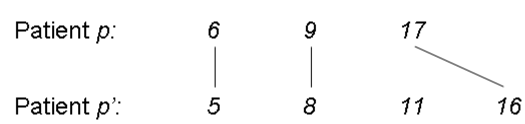

Figure 2. Matching for closest visit pattern

In the example shown in Figure 2, the distance $\rho(p,p') = 1$.

It can be easily checked that $\rho(q,p)$ is a tolerance relation.

Threshold is applied to filter out patient records with large distance values to form a tolerance class, where all group members have similar visiting patterns; therefore visit-related temporal features can be computed for all group members.

For instance, let us assume that we have 8 patients p_1, p_2, p_3,…, p_8 with doctor's visits assigned to them which are represented by vectors:
$v_{p1} = [3, 8, 12, 20]$, $v_{p2} = [4, 7]$, $v_{p3} = [5, 12, 21, 30]$, $v_{p4} = [7, 21, 29]$, $v_{p5} = [12, 22]$, $v_{p6} = [13, 19, 29]$, $v_{p7} = [2, 13, 19, 31, 38]$, $v_{p8} = [7, 12, 20]$.

The threshold value $\rho=1$ is set up as a minimal distance between vectors representing patients.

The following tolerance classes containing more than one element will be constructed:

$TC_{\rho=1}(v_{p2}) = \{v_{p1}, v_{p2}\}$, $TC_{\rho=1}(v_{p4}) = \{v_{p4}, v_{p3}\}$, $TC_{\rho=1}(v_{p5}) = \{v_{p5}, v_{p1}, v_{p3}, v_{p8}\}$, $TC_{\rho=1}(v_{p6}) = \{v_{p6}, v_{p7}\}$, $TC_{\rho=1}(v_{p6}) = \{v_{p6}, v_{p1}\}$.

We say that $TC_{\rho=1}(v_{p2})$ is generated by p_2 *and similarly* $TC_{\rho=1}(v_{p4})$ is generated by p_4.

The ultimate goal of constructing tolerance classes is to identify right groups of patients for which useful temporal features can be built. By increasing the threshold value, we get larger classes for the process of knowledge extraction but the information included in temporal features will be less accurate. On the other hand, if the threshold value is too small, the size of tolerance classes might be also too small in order to get any useful information through the knowledge extraction process.

4. Temporal feature extraction

The dataset associated with a tolerance class which is generated by patient *p* contains records describing patients who visited their doctor at during similar weeks as the patient *p*. Data referring only to these visits are stored in tuples representing all patients in this tolerance class. In other words, if patient *p* generates a tolerance class $TC_{\rho=1}(v_{p2})$ where $v_{p2} = [4, 7]$ and another patient *p1* has a vector representation $v_{p1} = [3, 8, 12, 20]$ of his doctor's visits, then *p1* has a vector representation *[3, 8]* relative to $TC_{\rho=1}(v_{p2})$. This way all patients associated with the same tolerance class have the same number of doctor's visits and all these visits occurred approximately at the same time.

This paper involves the construction of a collection of databases D_p, where *p* is a patient and D_p corresponds to $TC_\rho(v_p)$, for the purpose of classifiers construction based on the database *O* that was mentioned in the previous section. This paper uses the term "attribute" to refer to a column in the table from the database *O* and the term "feature" to refer to a column in the database D_p. Also, due to the intuition of the process in each visit, recovery of any patient with only one visit cannot be evaluated. Therefore, such records have been removed during the experiments. During a period of medical treatment, a doctor may change the treatment from one category to another based on the specifics of recovery of the patients and the symptoms of a patient may vary as a result of the treatment. Additionally, the category of patient may change over time (e.g., hyperacusis can be totally eliminated and consequently the patient may move from treatment category 3 to 1). Other typical categorical features which may change over time in the database *O* include sound-instrument types as well as visiting frequencies. Statistical and econometric approaches to describe categorical data have been well discussed in [6].

In terms of continuity, there are two types of data: one is numerical, such as scores for emotions, functions, and catastrophes related to the tinnitus problems; the other is categorical, such as instruments used in the therapy and patient categories. In terms of stability, there are other two types of data: one is stable; the other is flexible [8]. In this research, stable is defined relative to others: an attribute should have the same value along time throughout the most of the records.

Under all the above assumptions, the transform from visit-based format *O* to patient-based format D_p for each tolerance class $TC_\rho(v_p)$ is quite straightforward.

Let us assume that $TC(v_p)$ is a tolerance class generated by patient *p* where $v_p = [v_1, v_2, \ldots, v_n]$ and $[w_{J(1)}, w_{J(2)}, \ldots, w_{J(n)}]$ is a vector representation of patient *q* relative to $TC(v_p)$. We also assume that $J(0)=1$.

Now, assume that *A* is a numerical attribute which time-dependent values for patient *q* are given as a vector $[a_{J(1)}, a_{J(2)}, \ldots, a_{J(n)}]$.

For each patient $q \in TC_\rho(v_p)$, if *n* is an even number, we compute the temporal feature value $A_1(q)$ to describe the derivative of an attribute *A* against a number of rounded weeks between his doctor's visit $J(0)$ and $J(n)/2$. We also compute $A_2(q)$ to describe the derivative of an attribute *A* against a number of rounded weeks between his doctor's visit $J(n)/2$ and $J(n)$. Finally, we compute $A_3(q)$ to describe the derivative of an attribute *A* against a number of rounded weeks between his doctor's visit *J(0)* and *J(n)*.

$$A_1(q) = \frac{[a_{J(n)/2} - a_{J(0)}]}{\left|w_{J(n)/2} - w_{J(0)}\right|} \qquad (\mathbf{2})$$

$$T_1(q) = /a_{J(n)/2} - a_{J(0)}/ \qquad (\mathbf{3})$$

$$A_2(q) = \frac{[a_{J(n)} - a_{J(n)/2}]}{\left|w_{J(n)} - w_{J(n)/2}\right|} \qquad (1)$$

$$T_2(q) = /a_{J(n)} - a_{J(n)/2}/ \qquad (2)$$

$$A_3(q) = \frac{[a_{J(n)} - a_{J(0)}]}{\left|w_{J(n)} - w_{J(0)}\right|} \qquad (3)$$

$$T_2(q) = /a_{J(n)} - a_{J(0)}/ \qquad (4)$$

When *n* is an odd number, we developed the temporal feature $A_4(q)$ to describe the derivative of an attribute against time of rounded week of its first visiting duration. Temporal feature $A_5(q)$ is similar to $A_3(q)$.

$$A_4(q) = \frac{[a_{J(1)} - a_{J(0)}]}{\left|w_{J(1)} - w_{J(0)}\right|} \qquad (\mathbf{5})$$

$$T_4(q) = /a_{J(1)} - a_{J(0)}/ \qquad (6)$$

$$A_5(q) = \frac{[a_{J(n)} - a_{J(1)}]}{\left|w_{J(n)} - w_{J(1)}\right|} \qquad (7)$$

$$T_5(q) = /a_{J(n)} - a_{J(0)}/ \qquad (8)$$

New features, $A_6(q)$ and $T_6(q)$ are defined similarly to $A_1(q)$ and $T_1(q)$. Finally, $A_7(q)$ and $T_7(q)$ are defined similarly to $A_2(q)$ and $T_2(q)$.

5. Decision tree and action rules extraction

Action rules are rules extracted from an information system that describe a possible transition of objects from one state to another with respect to a distinguished feature called a decision feature. Features used to describe objects are partitioned into stable and flexible. Values of flexible features can be changed. This change can be influenced and controlled by users. Action rules mining initially was based on comparing profiles of two groups of targeted objects - desirable and undesirable [8].

The concept of an action rule was introduced by Ras and Wieczorkowska and defined as a term $\omega\wedge[\alpha\rightarrow\beta]\Rightarrow[\phi\rightarrow\psi]$, where ω is a conjunction of fixed condition features shared by both groups, $[\alpha\rightarrow\beta]$ represents proposed changes in values of flexible features, and $[\phi\rightarrow\psi]$ is a desired effect of the action [8]. Symbol $\wedge$ is interpreted as logical *and*.

To a decision system corresponding to $TC_\rho(v_p)$, authors applied tree classifiers to get classification rules and the action rules construction method proposed in [9]. To reduce the number of values for numerical attributes in the database we use a classical method based either on entropy or Gini index resulting in a hierarchical discretization. Before using any flexible attribute in the process of a decision tree construction, all stable attributes have been used first. This way the decision table is split into a number of decision sub-tables leading to them from the root of the tree by uniquely defined paths built from stable attributes [5]. Each path defines a header in any action rule extracted from the corresponding sub-table.

6. Experiments and results

Tools used in this research include SAS with connection to Microsoft Access databases as well as WEKA. The authors explored a database of 215 patients who have at least four visits during the process of the TRT, with 32 features. A threshold of 2.5 was applied to distance during the data selection, where sub-datasets by 14 seeds out of 215 were formed, which became a new database with 747 records.

Table 2. Seeds generated with distance ≤ 3 weeks

Seed Patient ID	Size of the Subset
03093	62
02038	61
01067	58
04098	57
05024	53
03075	53
05013	52
05011	52
00067	52
04062	51
04008	51
00026	49
01052	48
01038	48
total	747

Seed patient records appearing above all have three visits in total. This is due to the nature of the data selection algorithm: we may select close visits from the paired patient only when the paired patient has more visits than the seed patient; therefore, records with small visit numbers tend to collect more similar patterns around them. The features applied in this research include: A_1, A_2, A_3, T_1, T_2, T_3 for the total score of negative emotions, A_1, A_2, A_3, T_1, T_2, T_3 for the total score of functional problems, A_1, A_2, A_3, T_1, T_2, T_3 for the total score of catastrophe, the most important problem in the first visit, the most important problem in the second visit, the most important problem in the third visit, the sound instrument used in the first visit, the sound instrument used in the second visit, the sound instrument used in the third visit, the follow-up method after the first visit, the follow-up method after the second visit, the follow-up method after the third visit, the dependency on the presence of hyperacusis in the first visit, the dependency on the presence of hyperacusis in the second visit, the dependency on the presence of hyperacusis in the third visit, the real ear measurements in the first visit, the real ear measurements in the second visit, the real ear measurements in the third visit, patient category, the treatment category in the first visit, the treatment category in the second visit, the treatment category in the third visit. The decision attribute relates to whether the symptom of either the negative emotions, or functional problems, or catastrophe, or the overall is relieved or not.

Table 3. Seed generated with distance ≤ 4 weeks

Seed Patient ID	Size of the Subset
03071	47
total	47

The above seed patient record has four visits in total.

The features applied in this research include: A_3, A_4, A_5, A_6,, A_7, T_3, T_4, T_5, T_6, T_7 for the total score of negative emotions, A_3, A_4, A_5, A_6, A_7, T_3, T_4, T_5, T_6, T_7 for the total score of functional problems, A_3, A_4, A_5, A_6, A_7, T_3, T_4, T_5, T_6, T_7 for the total score of catastrophe, the most important problem in the first visit, the most important problem in the second visit, the most important problem in the third visit, the most important problem in the fourth visit, the sound instrument used in the first visit, the sound instrument used in the second visit, the sound instrument used in the third visit, the sound instrument used in the fourth visit, the follow-up method after the first visit, the follow-up method after the second visit, the follow-up method after the third visit, the follow-up method after the fourth visit, the dependency on the presence of hyperacusis in the first visit, the dependency on the presence of hyperacusis in the second visit, the dependency on the presence of hyperacusis in the third visit, the dependency on the presence of hyperacusis in the fourth visit, the real ear measurements in the first visit, the real ear measurements in the second visit, the real ear measurements in the third visit, the real ear measurements in the fourth visit, patient category, the treatment category in the first visit, the treatment category in the second visit, the treatment category in the third visit, the treatment category in the fourth visit. The decision attribute is about if the symptom of either the negative emotions, or functional problems, or catastrophe, or the overall is relieved or not.

Decision tree study was performed using C4.5 [7], a system that incorporates improvements to the ID3 algorithm for decision tree induction. C4.5 includes improved methods for handling numeric attributes and missing values, and generates classification rules from the trees [10]. In this study C4.5 was used in order to evaluate positive recovery from functional problems, negative emotions, and catastrophe. All classifiers were 10-fold cross validation with a split of 90% training and 10% testing.

The authors discovered the following action rules.

Rule 1. generated from seed 02038: [(*patient category*=2) $\wedge$ (A_1 *of total score* $\leq$ 3.7) $\wedge$ (*initial real ear measurements*, y→n)] $\Rightarrow$ (*positive recovery of catastrophe*, n→y)
Supports: 4, confidence: 75.3%

Meaning of the above rule: if patients have hearing loss as a significant subjective problem and tinnitus as a significant problem, have *A1* of the total score less than 3.7, having real ear measurements in the first visit or not decides if they will have improvements in terms of catastrophe scores after TRT.

Rule 2. generated from seed 02038: [(*patient category*=3) $\wedge$ (A_1 *of total score* $\leq$ 3.7) $\wedge$ (*follow up method*, "counseling" →"telephone") $\Rightarrow$ (*positive recovery of catastrophe*, n→y)]
Support: 2, confidence: 92.3%

Meaning of the above rule: if patients have the primary problem of hyperacusis and are treated for this condition with a specific TRT protocol that involves use of wearable sound generators or combination instruments and have *A1* of the total score less than 3.7, the change of follow up method from counseling to telephone indicates improvements in terms of catastrophe scores after TRT. More so, the rule strongly suggests that telephone method in the mentioned condition means improvements on catastrophe, where the side of the action rule has a support of 13.

Rule 3. generated from seed 04062: [(*patient category*=3) $\wedge$ (A_1 *of total score* $\leq$ 1.1) $\wedge$ (*initial dp*=n) $\Rightarrow$ (*positive recovery of negative emotion*=y)]
Support: 9, confidence: 69.2%

Meaning of the above rule: if a patient has the primary problem of hyperacusis and is treated for this condition with a specific TRT protocol that involves use of wearable sound generators or combination instruments and has *A1* of the total score less than 1.1, and there is no dependency on the presence of hyperacusis, he or she may have improvements in terms of negative emotions score after TRT.

Rule 4. generated from seed 04062: [(*patient category*=4) $\wedge$ (A_1 *of total score* $\leq$1.1) $\wedge$(T_1 *of total score*, $\leq$ 2→>2) $\Rightarrow$ (*positive recovery of negative emotion*, n→y)]
Support: 4, confidence: 66.7%

Meaning of the above rule: if patients are relatively uncommon and suffer from a condition in which their tinnitus or their hyperacusis is significantly worsened because of exposure to certain types of sounds, and if their A_1 of the total score is not greater than 1.1 and their T_1 of the score of catastrophe changes from less than or equal to -2 to greater than -2, then they may begin to have improvements on negative emotions. More so, when the T_1 of the score of catastrophe is greater than -2, the mentioned typed of patients will always have improvement on negative emotions, as this side of the mentioned action rule has support of 7 and confidence of 100%.

Rule 5. generated from seed 02038: [(*patient category*=3) $\wedge$ (A_1 *of total score* $\leq$1.1) $\wedge$(T_2 *of total score* $\leq$ 12) $\wedge$(*most important problem*, H→(T | L))]$\Rightarrow$(*positive recovery of negative emotion*, n→y)
Support: 7, confidence: 63.2%

Meaning of the above rule: if a patient has the primary problem of hyperacusis and is treated for this condition with a specific TRT protocol that involves use of wearable sound generators or combination instruments, and has A_1 of the total score less than 1.1 and T_2 of the total score less than or equal to 12, the change of the most importance problem from hyperacusis to tinnitus or hearing loss indicates improvement on negative emotions.

Rule 6. generated from seed 00026 : [(T_1 *of total score* $\leq$ 2) $\wedge$ (A_2 *of total score* $\leq$ 12) $\wedge$ (*the most important problem in the last visit* is "Tinnitus") $\wedge$ (*rem of the second visit*, y→n)]$\Rightarrow$ (*positive recovery of negative emotion*, n→y)
Support: 8, confidence: 66.7%

Meaning of the above rule: if a patient has T_1 of the total score not great than 2 and T_2 of the total score not greater than 12, and if tinnitus is the most important problem in the last visit, stopping the real ear measurements in the second visit means improvement of the negative emotions.

Rule 7. generated from seed 00026: [(T_1 *of total score* > -4) $\wedge$ (T_2 *of total score* <= -2) $\wedge$(T_2 *of* catastrophe > -4) $\wedge$ ((p2 = T) |(p2 = L)] $\Rightarrow$ (ta_sc_f13$\leq$0)
Support: 16, confidence: 76.2%

Meaning of the above rule: if a patient has T_1 of the total score greater than -4 and T_2 of the total score not greater than -2 and T_2 of the catastrophe greater than -4 and the most important problem of the second visit is either tinnitus or sound loss, then he or she will have improvement in terms of functional problems.

7. Conclusion

TRT is a complex treatment process, which generates a high volume of matrix data over time: some attributes have relatively stable values while others may be subject to change as the doctors are tuning the treatment parameters while symptoms of patients are altering. Understanding the relationships between and patterns among treatment factors helps to optimize the treatment process. The authors investigated new flexible temporal features to describe the sparse visiting records of patients for the purpose of classification tree construction. Interesting action rules about the relationship among treatment factors and symptoms were revealed during the experiments. As TRT itself has been successful, the record of treatment reflected in the database has unbalanced results as there are much more positive recovery examples than negative ones. Therefore, as the database grows and more negative examples are collected, the confidence of the action rules will increase.

8. Acknowledgment

This material is based in part upon work supported by the Ministry of Science and Higher Education in Poland under Grant *N N519 404734* and by AMVIS Program CONTACT in Czech Republic under Grant *ME 913*.

9. References

[1] McFadden D. Tinnitus: facts, theories, and treatments. Washington (DC): National Academy Press; 1982.

[2] Heller MF, Bergman M. Tinnitus in normally hearing persons. Ann Otol Rhinol Laryngol 1953; 62:73–93.

[3] Jastreboff, P.J. Phantom auditory perception (tinnitus): mechanisms of generation and perception. Neurosci.Res.:221-254, 1990.

[4] Jastreboff, P.J., Gray, W.C., Gold, S.L. Neurophysiological approach to tinnitus patients. Am.J.Otology, 17:236-240, 1996.

[5] L.-S. Tsay and Z.W. Ras (2006), Action rules discovery system DEAR3, in Foundations of Intelligent Systems, Proceedings of ISMIS 2006, Bari, Italy, LNAI, No. 4203, Springer, 483-492

[6] D. Powers and X. Yu (1999) Statistical Methods for Categorical Data Analysis, Academic Press.

[7] J.R. Quinlan (1993) C4.5: Programs for Machine Learning San Mateo, CA. Morgan Kaufmann.

[8] Z.W. Ras, A. Wieczorkowska (2000), Action-Rules: How to increase profit of a company, in Principles of Data Mining and Knowledge Discovery, Proceedings of PKDD 2000, Lyon, France, LNAI, No. 1910, Springer, 587-592

[9] Z.W. Ras and A. Dardzinska (2009), Action Rules Discovery Based on Tree Classifiers and Meta-actions, in Proceedings of the 18th International Symposium on Foundations of Intelligent Systems, Prague, Czech Republic, 66-75.

[10] I.H. Witten and F. Eibe (2005) Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers.

[11] Zhang, X., Ras, Z.W.: Analysis of Sound Features for Music Timbre Recognition. Invited paper, in proceedings of the IEEE CS International Conference on Multimedia and Ubiquitous Engineering, April 26-28, in Seoul, Korea, 3-8. (2007).