Thai Monosyllabic Words Recognition using Ant-Miner Algorithm

Saritchai Predawan[1], Chom Kimpan[1], and Chai Wutiwiwatchai[2]
[1]Faculty of Information Technology, Rangsit University, Thailand
[2]National Electronics and Computer Technology Center, Ministry of Science and Technology, Thailand

Abstract: *In this paper, Ant-Miner software is used to develop classification rules for Thai monosyllabic words. The hypothetical words used in this paper are composed of 65 command monosyllabic Thai words. The binary desired outputs were used during training 520 Thai words consist of 10 numerals and single-syllable, 65 words in each group were used for system evaluation. In order to improve recognition accuracy, initial consonants, vowels, final consonants and tonal level detected were conducted for speech preclassification. The parameters used in the metaheuritstic algorithms are optimized using pruning algorithm with the aim of improving the accuracy by generating minimum number of rule in order to cover more patterns. Thai monosyllabic words recognition using Ant-Miner yielded Thai monosysllabic words accuracy of recognition on test set of 88.65%, 87.69% and 91.54% for 50, 100 and 250 number of ants respectively.*

Keywords: *Thai monosyllabic words recognition, ant-miner algorithm, classification, Thai language.*

Received March 17, 2011; accepted May 24, 2011; published online August 5, 2012

1. Introduction

Thai language structure consists of 44 consonants, 24 vowels, and 5 tonal levels. These can be combined to a lot of words and also provide many ambiguous words. Thai language is a tonal language. There are five tones in Thai, low "ake", medium "saman", falling "toe", high "dtee" and rising "jattawa". The feature of speech that was used to classify the tone is the shape of fundamental frequency (F0) contour, which shown in Figure 1. There are several parameters that also have the effect on the shape of (F0) contour such as the gender and the age of speaker, the initial consonant, the final consonant and the duration of vowel (short or long). Intonations in Thai are used to distinguish the word's meanings [9]. Research in Thai tone extraction presented by Ramalingam [18] is compared with our work. Besides, several Thai speech recognition approaches that have been elaborated for years such as in [12, 14] are studied. The correctness values of these approaches are 74.9% [12] and 89.4% [14]. Many researchers have tried to investigate techniques of the global acoustic model, the acoustic model combination and context-independent model. All these techniques require large text and speech resources. The challenge is how to build the acoustic models from the limited data resources since some speech recognition applications do not required more than 90% of daily word coverage.

Rule discovery is an important data mining task since it generates a set of symbolic rules that describe each class or category in a natural way. The human mind is able to understand rules better than any other data mining model. However, these rules need to be simple and comprehensive; otherwise, a human won't be able to comprehend them. Evolutionary Algorithms have been widely used for rule discovery, a well-known approach being learning classifier systems.

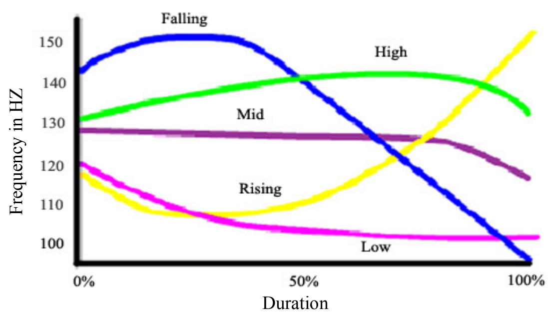

Figure 1. F0 normalize contour pattern of five Thai tones.

To our knowledge, Parpinelli *et al.* [13] were the first to propose Ant Colony Optimization (ACO) for discovering classification rules, with the system Ant-Miner. They argue that an ant-based search in more flexible and robust than traditional approaches. Their method uses a heuristic value based on entropy measure.

In [17], we presented Thai tone recognition using ant-miner algorithm, where the first to propose ACO for discovering classification rules, with the system Ant-Miner. We argue that an ant-based search is more flexible and robust than tradition approaches.

The remainder of the paper is organized as follow: In section 1, we present the over view of the Thai speech recognition system and the basic idea of the ant colony system. In section 2, acoustic modeling and

Ant-Miner algorithm are described. Section 3 describes speech database, experiments and results are illustrated in sections 4 and 5 notes conclusions.

2. Acoustic Model

The first and foremost principle of the speech recognition that makes the system useful and powerful is the acoustic models. Acoustic modeling is a very important process because it directly effects the search speed, and accuracy. The design factors of acoustic modeling include the number of models which are suitable for the language coverage and the size of speech training database. The size of training database directly impacts the system performance. The number of acoustic models corresponds to linguistic knowledge of target language [2]. Hence many acoustic modeling techniques are applied by following the linguistic knowledge. In this section, the basic of Thai phones and acoustic modeling techniques are described.

The single and double Thai consonants are shown in Tables 1 and 2 respectively, and Table 1 indicates differences between the initial consonant and final consonant whereas Table 3 is Thai vowel symbols. These symbols of Thai phonemes are a little bit different from those of another Thai linguist such as Luksaneeyanawin [11]. But in the case of Thai tone, the digits 0 to 4 are used to represent the five tones, which are middle, low, falling, high and rising, respectively.

Table 1. Initial and final consonant symbol (26 phonemes and 12 phonemes).

Consonant	Phoneme		Consonant	Phoneme	
	Initial (C_i)	Final (C_f)		Initial (C_i)	Final (C_f)
ก	k	k^	บ	b	p^
ข,ค,ฆ	kh	k^	ป	p	p^
ง	ng	ng^	ผ,พ,ภ	ph	p^
จ	c	t^	ฝ,ฟ	f	p^
ฉ,ฌ,ช	ch	t^	ม	m	m^
ซ,ศ,ษ,ส	s	t^	ร	r	n^
ญ,ย	j	j^	ล,ฬ	l	n^
ฎ,ด	d	t^	ว	w	w^
ฏ,ต	t	t^	ห,ฮ	h	-
ฐ,ฑ,ฒ,ถ,ท,ธ	th	t^	อ	z	-
ณ,น	n	n^	Foreign language	br,bl,fr,fl,dr	f^,s^,ch^,l^

Table 2. Double Thai consonant (12 phonemes).

Double Consonant	Phoneme Symbol	Double Consonant	Phoneme Symbol	Double Consonant
ปร	pr	กร	kr	ปร
ปล	pl	กล	kl	ปล
พร	phr	กว	kw	พร
พล	phl	คว	khr	พล
ตร	tr	คล	khl	ตร
ทร	thr	ขว	khw	ทร

Table 3. Thai vowel symbol (24 phonemes).

Tongue Height	Tongue Advancement		
	Front (Short/Long)	Central (Short/Long)	Back (Short/Long)
Close	i, ii (อิ, อี)	v, vv (อึ, อือ)	u, uu (อุ, อู)
Mid	e, ee (เอะ, เอ)	q, qq (เออะ, เออ)	o, oo (โอะ, โอ)
Open	x ,xx (แอะ, แอ)	a, aa (อะ, อา)	@, @@ (เอาะ, ออ)
Diphthongs	ia, iia (เอียะ, เอีย)	va, vva (เอือะ, เอือ)	ua, uua (อัวะ, อัว)

2.1. Thai Phone Model

The general forms of Thai syllables are C_iV and C_iVC_f and the tone is marked onto each syllable. Five different tones in Thai are divided into two groups:

1. The static group-high, middle, and low tone.
2. The dynamic group-rising and falling tone.

Thai phonetic system has 21 single consonants, 12 double consonants, 24 vowels, and more than 5 double consonants that use for pronouncing the foreign word. A number of Thai phones are 74 (38+24+12) as shown in Table 4.

Table 4. Thai phone symbols.

Type of Phone	Phone Symbols	
Initial Consonants	Single	k, kh, ng, c, ch, s, j, d, t, th, n, b, p, ph, f, m, r, l, w, h, z
	Double	pr, pl, tr, kr, kl, kw, phr, phl, thr, khr, khl, khw, br, bl, fr, fl, dr
Vowels	Short	a, i, v, u, e, x, o, @, q
	Long	aa, ii, vv, uu, ee, xx, oo, @@, qq
	Diphthong	ia, iia, va, vva, ua, uua
Final Consonants	k^, ng^, j^, t^, n^, p^, m^, w^, ch^, f^, l^, s^	

That also include 5 initial consonants *(/br/, /bl/, /fr/, /fl/, /dr/)* and 4 final consonants *(/f^/, /s^/, /ch^/, /l^/)* for foreign words. The character "^" is used to denote the difference between the initial consonants and the final consonants. More details are given in [8, 11]. All of Thai phone symbols used in this paper are compatible with those used by other Thai researchers or linguists [11]. The Thai syllable structure is composed of three different sound systems as follows [10]:

1. The system of consonants consists of 33 consonantal units, 21 consonants and 12 consonant clusters.
2. The system of Thai vowels consists of 18 monophthongs and 6 diphthongs. The monophthongs are qualitatively 9 different vowels, each of which has two members, short and long. Each of three different diphthongs also has 2 quantitatively different members.
3. The system of tones consists of 5 tones. There are 3 kinetic or relatively leveled tones, the high (H), the

mid (M), and the low (L), and 2 dynamic or contour tones, the falling (F) and the rising (R).

The smallest construction of sounds or syllables in Thai is composed of one vowel unit or one diphthong, one two or three consonants, and a tone. The construction can be represented with the structure as illustrated in equitation 1:

$$S=C_i(C_i)V^t(V)(C_f) \tag{1}$$

where: C_i is initial consonant, C_f is final consonant, V is vowel, t is tone.

In our study we divide syllable into phonemes based on the time that each phoneme occurs and their characteristics. These phonemes are: 1). The initial consonant, 2). The vowel, 3). The secondary vowel, 4). The syllable ending, and 5). The tonal. Considering the following examples in Table 4.

In forms of either $/C_i\_V\_T/$ or $/C_i\_V\_C_f\_T/$, where C_i denotes the initial consonant (including single and double consonant), V denotes the vowel (both short and long vowel), C_f denotes the final consonant (some single consonants), and T denotes the tone Thai has approximately 43,255 vocabularies [7] and five tones. The characteristics of five tones are separated by the shape of fundamental patterns (F_0) [19]. In each tone, it has individual different meaning. Thus, tone is one of the main problems in speech recognition accuracy. For example, "khaa" have meanings according to each tone as show in Table 5 [17].

Table 5. Example of "khaa" in five meanings.

No.	Thai Words	Phonemic Transcription	English Translation	Tone
1	คา	khaa	a kind of grass	0
2	ข่า	kh`aa	galingale	1
3	ฆ่า	khˆaa	to kill	2
4	ค้า	kh'aa	to trade	3
5	ขา	khˇaa	a leg	4

2.2. Tone in Thai Language

Thai is a tonal language which makes it very different to Western languages. The phonetic structure of Thai is based primarily upon the monosyllable. Thai syllable structure consists of $/C(C)V(:)(C)^T/$ where C, V, and T represent a consonant, vowel, vowel length and lexical tone, respectively [17]. Each syllable has a choice between five distinct tones: low, mid, rising, high and falling with 4 mark symbols respectively. Tone is a supra-segmental that lies on a group of voiced segment and usually associated with vowels.

The level of tone depends on frequency of vocal cords vibration. If high frequency of vibration the level of tone will be high, on the other hand the tone level is low if the frequency of vibration is low. When a tone of vowel is occurred in the syllable, the other tone in that syllable will be changed to the same as vowel tone. Normally, Tone is occurred together of vowel. There are 2 groups of standard Thai language tone: level tone and contour tone [10]. Level tone means the level of tone is stable from beginning to the end. There are three phones of this tone:

1. *Low Tone:* Is the phone tone "Ek" which started at level of 118Hz and then decrease a bit until close to the end syllable around 110Hz. This is a level tone with no inflection but lower in pitch than common tone. The symbol for this phone tone is /1/ for example: ป่า /Pa1/.
2. *Mid Tone:* Is the phone tone "Saman" which started at level of 120Hz and then decrease a bit until close to the end syllable around 112Hz. This is spoken in the speaker's ordinary tone of voice without any inflection. It is the tone used in English for ordinary conversation. The symbol for this phone tone is /0/ for example: ปา /Pa0/.
3. *High Tone:* Is the phone tone "Tri" which started at level of 125Hz and then increasing around 134-140Hz until close to the end syllable and decrease a bit when close to the end syllable. This is a uniform tone pitched well above the level of the speaker's normal voice. The symbol for this phone tone is /3/ for example: ป๊า /Pa3/.

Contour tone means the level of tone is change from the beginning. It can be from high to low and vice versa between the periods of utterance for each syllable and high frequency changing between the initial and final syllable. There are two phones of this tone:

1. *Rising Tone:* Is the phone tone "Chattawa". It starts from low level around 110Hz, decreasing a bit and then change to high rapidly and at the end of syllable the frequency is around 140Hz. This as the name implies has a rising inflection. The symbol for this phone tone is /4/ for example: ป๋า /Pa4/.
2. *Falling Tone:* Is the phone tone "Tho". It starts from high level around 140Hz, increasing a bit and then change to low rapidly till lowest at the end of syllable around 100Hz. This is an emphatic and heavily accented tone with a falling inflection. The symbol for this phone tone is /2/ for example: ป้า /Pa2/.

In the case of indistinct tone classification words such as "khaa", "maa" and "paa". They are trained into each five tones. The speech recognition results are difficult to recognize because they have similar likelihood scores. To improve the recognition performance, researchers and developers have to use many recognition patterns to compute the recognition results [6]. Thai has five distinctive tones and each tone is well represented by its fundamental frequency (F_0) pattern [11]. F_0 is considered to use tone classification as an input. Several interacting factors affect F_0 realization of tones. Several factors, including tonal coarticulation, stress, and intonation, may affect the tone patterns of Thai speech. The F_0 pattern of a syllable affected by the F_0 patterns of neighboring syllables is referred to as tonal coarticulation [5]. The F_0 contour patterns of the unstressed syllables are generally different from the stressed ones [15]. The intonation effect makes the F_0 contour of the utterances decline gradually [16]. The statistical data analysis of the acoustical features, including duration, energy, and fundamental frequency (F_0) of stressed and unstressed syllables are computed from the training sets uttered by all speakers. It was found that duration and normalized energy are the effective features for distinguishing between the stressed and unstressed syllables, while the mean normalized F_0 did not signal the stress function of the syllables. Intonation affects the tone patterns by making them decline gradually. Within each tone, the mean F_0 of the preceding syllable is higher than the succeeding syllable, and the mean F_0 is lowest at the ending syllable of the sentence. This datum suggests that the mean F_0 can he used to deal with the intonation effect. The tone pattern of a syllable is also affected by tone patterns of the neighboring syllables due to the anticipatory and carryover coarticulation. It is evident that F_0 contours of both stressed and unstressed syllables are subject to modification by the preceding and succeeding syllables.

There are two different types of Thai syllable: stressed and unstressed syllable. Stressed syllable is the syllable that can be stressed individually in normal speaking. Its consist of at least three parts: initial consonant vowel and tone as (C_i-V-T) or maximum five parts of two initial consonants (cluster consonant) one vowel one final consonant and tone (C_iC_i-V-C_f-T). Unstressed syllable is the syllable that appears in the unstressed position in normal speaking. The Thai phoneme set consists of 21 consonantal phonemes, 17 consonantal cluster phonemes, and 24 vowels as shown in Table 4. To determine the tone of any syllable the following three factors have to be considered.

1. *Class of the Initial Consonant:* The Thai 44 consonants are divided up into three groups known respectively as High, Middle and Low class consonants and the first thing to look at in determining the tone of a word or syllable is the class of the initial consonant. There are many cases where the letter "ห" as an initial consonant is silent and there are a few cases where the letter "อ" as an initial consonant is also silent, but this makes no difference to the rule, the tone is still governed by the class of the initial consonant even though it can be a silent consonant.

   The High class consonants are:

   ข-kh ฉ-ch ฐ-ถ-th ผ-ph ฝ-f ศ-ษ-ส-s ห-h

   The Middle class consonants are:

   ก-k จ-c ฎ-ด-d ฏ-ต-t บ-b ป-p อ-z

   All the remainders are Low class consonants:

   พ-ภ-ph ฟ-f ฑ-ฒ-ท-ธ- th ค-ฅ-ฆ-kh ซ-s ฮ-h ช-ฌ-ch ง-ng ญ-ย-j น-ณ-n ร-r ว-w ม-m ฬ-ล-l

2. *The Final Sounded Consonant:* All words which do not end in a vowel sound must have either m, n, ng, k, p, or t as the final sound. Although this is strictly true, but in some conversation the final consonant is often slurred and particularly after a long vowel, the final "p" may sound more like a "b" and the final "t" more like a "d". Where there is no tone mark, the tone of the syllable or word will depend on both the class of the initial consonant and on whether it ends with the m, n, ng sounds or the k, p, t sounds. It should be noted that a final consonant with the sign -์ over it is not sounded and hence can have no effect on the tone.

3. *The Type of Final Vowel:* If the word has no tone mark and ends in a final vowel, the tone is dependent on whether this final vowel is a long or short one. The short vowels for tonal purposes are -ะ, -ั , -ั , -ั , -ั , the inherent "a", the inherent "o" and all vowels shortened by the sign -็ over the consonant or by the addition of the vowel -ะ at the end. All the others are long vowels.

2.3. Ant Colony Optimization and Ant Miner

ACO [3] is a branch of a newly developed form of artificial intelligence called swarm intelligence. Swarm intelligence is a field which studies "the emergent collective intelligence of groups of simple agents" [1]. In groups of insects, which live in colonies, such as ants and bees, an individual can only do simple tasks on its own, while the colony's cooperative work is the main reason determining the intelligent behavior it shows. Most real ants are blind. However, each ant while it is walking, deposits a chemical substance on the ground called pheromone [4]. Pheromone encourages the following ants to stay close to previous moves. The pheromone evaporates over time to allow search exploration. In a number of experiments presented in Dorigo and Gambardella [4] illustrate the complex behavior of ant colonies. For example, a set of ants built a path to some food. An obstacle with two ends was then placed in their way such that one end of the obstacle was more distant than the other. In the beginning, equal numbers of ants spread around the two ends of the obstacle.

Since all ants have almost the same speed, the ants going around the nearer end of the obstacle return before the ants going around the farther end (differential path effect). With time, the amount of pheromone the ants deposit increases more rapidly on the shorter path and so more ants prefer this path. This positive effect is called autocatalysis. The difference between the two paths is called the preferential path effect; it is the result of the differential deposition of pheromone between the two sides of the obstacle, since the ants following the shorter path will make more visits to the source than those following the longer path. Because of pheromone evaporation, pheromone on the longer path vanishes with time. The goal of Ant-Miner is to extract classification rules from data [13]. The algorithm is presented in Figure 2.

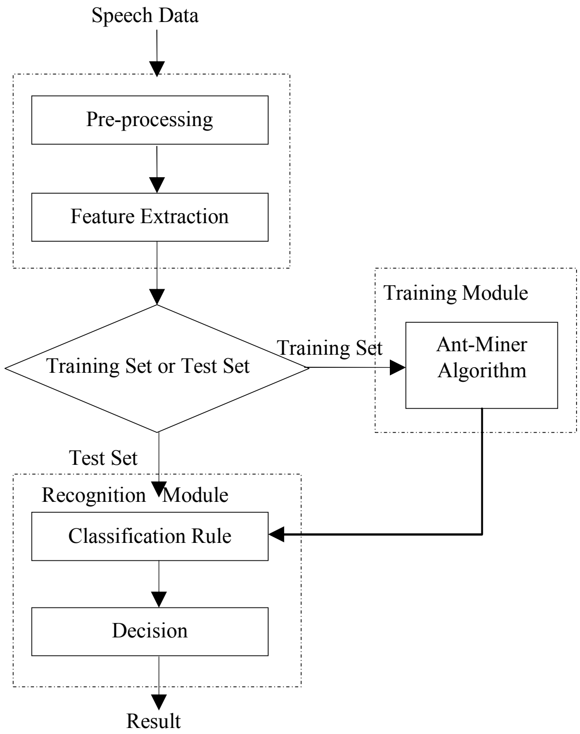

Figure 2. Block diagram of the system.

2.3.1. Pheromone Initialization

All cells in the pheromone table are initialized equally to the following equitation 2:

$$\tau_{ij(0)} = \frac{1}{\sum_{i=1}^{a} b_i} \tag{2}$$

where:

- a is the total number of attributes.
- b_i is the number of values in the domain of attributes A_i.

2.3.2. Rule Construction

Each rule in Ant-Miner contains a condition part as the antecedent and a predicted class. The condition part is a conjunction of attribute-operator-value tuples. The operator used in all experiments is "=", just as in Ant-Miner, all attributes are assumed to be categorical. Let us assume a rule condition such as term $_{ij} \approx A_i = V_{ij}$, where A_i is the ith attribute and V_{ij} is the jth value in the domain of A_i. The probability, that this condition is added to the current partial rule that the ant is constructing, is given by the following equation 3:

$$P_{ij}(t) = \frac{\tau_{ij}(t)\cdot\eta_{ij}}{\sum_i^a \sum_j^{b_i} \tau_{ij}(t)\cdot\eta_{ij}}\ ; \forall i \in I \tag{3}$$

where: η_{ij} is a problem-dependent heuristic value for term $_{ij}$, $\tau_{ij}(t)$ is the amount of pheromone currently, available (at time t) on the connection between attribute i and value j, I is the set of attributes that are not yet used by the ant.

2.3.3. Heuristic Value

In traditional ACO, a heuristic value is usually used in conjunction with the pheromone value to decide on the transitions to be made. In Ant-Miner, the heuristic value is taken to be an information theoretic measure for the quality of the term to be added to the rule. The quality here is measured in terms of the entropy for preferring this term to the others, and is given by the following equations 4 and 5:

$$\eta_{ij} = \frac{\log_2(k) - InfoT_{ij}}{\sum_i^a \sum_j^{b_i} \log_2(k) - InfoT_{ij}} \tag{4}$$

$$InfoT_{ij} = -\sum_{w=1}^{k} \left[\frac{freqT_{ij}^{w}}{|T_{ij}|}\right] * \log_2 \left[\frac{freqT_{ij}^{w}}{|T_{ij}|}\right] \tag{5}$$

where:

- k is the number of classes.
- $|T_{ij}|$ is the total number of cases in partition T_{ij} (partition containing the cases where attribute A_i has value V_{ij}).
- freq T_{ij}^{w} is the number of cases in partition T_{ij} with class w.

The higher the value of info T_{ij}, the less likely that the ant will choose term $_{ij}$ to add to its partial rule.

2.3.4. Rule Pruning

Immediately after the ant completes the construction of a rule, rule pruning is undertaken to increase the comprehensibility and accuracy of the rule. After the pruning step, the rule may be assigned a different predicted class based on the majority class in the cases covered by the rule antecedent. The rule pruning procedure iteratively removes the term whose removal will cause a maximum increase in the quality of the rule. The quality of a rule is measured using the following equation 6:

$$Q = \left[\frac{TruePos}{TruePos + FalseNeg}\right] \times \left[\frac{TrueNeg}{FalsePos + TrueNeg}\right] \tag{6}$$

where:

- *TruePos:* Is the number of cases covered by the rule and having the same class as that predicted by the rule.

- *FalsePos:* Is the number of cases covered by the rule and having a different class from that predicted by the rule.
- *FalseNeg:* Is the number of cases that are not covered by the rule, while having the class predicted by the rule.
- *TrueNeg:* Is the number of cases that are not covered by the rule which have a different class from the class predicted by the rule.

2.3.5. Pheromone Update Rule

After each ant completes the construction of its rule, pheromone updating is carried out as follows equitation 7:

$$\tau_{ij}(t+1) = \tau_{ij}(t) + \tau_{ij}(t) \cdot Q, \forall\ term_{ij} \in the\ rule \tag{7}$$

To simulate the phenomenon of pheromone evaporation in real ant colony systems, the amount of pheromone associated with each term $_{ij}$ which does not occur in the constructed rule must be decreased.

The reduction of pheromone of an unused term is performed by dividing the value of each τ_{ij} by the summation of all τ_{ij}.

3. Speech Database

In this paper, the speech database is a part of NECTEC-ATR Thai speech database called the isolated word. The data was collected from 8 native Thai speakers (4 male and 4 female speakers), aged from 18-40 years old. The conditions of recording are as follows. All utterances were recorded in a quasi-quiet room. The qualities are around 20dB. And only dynamic microphone is used in recording. All utterance is recorded in reading style and is middle and official dialect that is spoken in the middle area of Thailand. The recorded speech is 8 bits and 16kHz sampling rate. All utterances (65 words, 416 wav files) were used as training data, and 104 wave files were use as test set.

4. Experiments and Results

The platform used for conducting following experiments is a PC with Pentium 4 1.8GHz CPU, 1G memory, and runs under Windows XP2. The Algorithm is implemented by Java; the experimental parameters are confirmed by large number of experiments. Our system has the following four user-defined parameters:

- *Number of ants():* This is also the maximum number of complete candidate rules constructed and pruned during an iteration of the WHILE loop of Ant-Miner's algorithm, since each ant is associated with a single rule. In each iteration, the best candidate rule found is considered a discovered rule. The larger, the more candidate rules are evaluated per iteration, but the slower the system is.
- *Minimum number of cases per rule():* Each rule must cover at least cases to enforce at least a certain degree of generality in the discovered rules. This helps to avoid an over fitting of the training data.
- *Maximum number of uncovered cases in the training set():* The process of rule discovery is iteratively performed until the number of training cases that are not covered by any discovered rule is smaller than this threshold.
- *Number of rules used to test convergence of the ants():* If the current ant has constructed a rule that is exactly the same as the rule constructed by the previous ants, then the system concludes that the ants have converged to a single rule (path). The current iteration of the WHILE loop of Ant-Miner's Algorithm is therefore stopped and another iteration is started.

We assess the performance of the improved Algorithm, and compare its test results to Ant-Miner under the same experimental environment. All the experiments reported in this paper these parameters were set as follows:

No_of_ants=50,100 and 250 respectively.
Min_cases_per_rule=5.
Max_uncovere cases=10.
No_rules_convergence=10.
Evaporation factor (ρ)=0.95.
α and β=1.

Table 6 shows the main characteristic of the datasets, which were used to evaluate in our system.

Table 6. DataSet characteristics.

DataSet	Number of Examples	Number of Attributes	Number of Classes
65 words	520	63	65

Table 7. Results of testing with the Ant-Miner algorithm.

No_of_Ant	Accuracy Rate	Rules Number	Conditions Number
50	88.65% ± 1.59	62.6 ± 0.4	212.4 ± 3.14
100	87.69% ± 1.27	63.2 ± 0.2	218.2 ± 2.52
250	91.54% ±2.86	63.2 ± 0.66	219 ± 3.48

For all datasets, a 10-fold cross validation was used. In this procedure, all cases are used only once as testing and training. The final accuracy rate is simply the average of the accuracy rate of the iterations. All the data partitions are randomly generated considering all available cases.

Table 7 summarizes the results of testing obtained by the proposed Ant-Miner algorithm in our datasets. The table shows the accuracy rate, the number of rules found and the number of terms (the shown values are the average values of the cross-validation procedure followed by the corresponding standard deviation).

The result shows that the Ant-Miner parameter is able to produce minimum rules for recognition with more than 89% accuracy. The average number of rules used is 63.

5. Conclusions

We have described an Ant Colony System called Ant-Miner for the discovery of classification rules in databases [4]. We have also shown results indication that Ant-Miner had a good classification performance on the datasets used in our experiments. These results also show that the proposed algorithm is able to achieve both good predictive accuracy and a reduced number of rules at the same time. This facilitates the practical use the system, since it usually generates comprehensible rules. The main drawback is still its computational cost, especially when the search space (number of predicting attributes) is too large. However, to develop the understanding of parameters and effects of each parameter of every system needs a very detailed experimentation. The sole purpose of this paper is to help the researchers to select the one according to their need.

As future research direction, it would be interesting to investigate the performance of different feature selection methods. Also, it would be interesting to compare the proposed classifier with other classifiers and evaluate its performance on different data sets.

Finally, it would be promising to automate the procedure by embedding the feature selection algorithm to the ACO classification algorithm.

References

[1] Bonabeau E., Dorigo M., and Theraulaz G., *Swarm Intelligence: From Natural to Artificial System*, Oxford University Press, London, 1999.

[2] Debyeche M., Haton J., and Houacine A., "Improved Vector Quantization Approach for Discrete HMM Speech Recognition System," *The International Arab Journal of Information Technology*, vol. 4, no. 4, pp. 338-344, 2007.

[3] Dorigo M., Caro G., and Gambardella L., "Ant Algorithms for Discrete Optimization," *Artificial Life*, vol. 5, no. 2, pp. 137-172, 1999.

[4] Dorigo M. and Gambardella L., "Ant System: Optimization by a Colony of Cooperating Agents," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 26, no. 1, pp. 29-41, 1999.

[5] Gandour J. and Dechongkit S., "Tonal Coarticulation in Thai," *Journal of Phonetics*, vol. 22, no. 4, pp. 477-492, 1994.

[6] Kasuriya S., Kanokphara S., Thatphilhakkul N., Cotsomrong P., and Sunpethniyom T., "Context-Independent Acoustic Models for Thai Speech Recognition," *in Proceedings of IEEE International Symposium on Communications and Information Technologies*, Japan, vol. 2, pp. 991-994, 2004.

[7] Kasuriya S., Sornlertlamvanich V., Cotsomrong P., Kanokphara S., and Thatphithakkul N., "Thai Speech Corpus for Thai Speech Recognition," *in Proceedings of Oriental COCOSDA*, Thiland, pp. 54-61, 2003.

[8] Kasuriya S., Jitsuhiro T., Kikui G., and Sagisaka Y., "Thai Speech Recognition by Acoustic Models Mapped from Japanese," *in Proceedings of the Symposium on Natural Language Processing Oriental COCOSDA*, Thailand, pp. 211-216, 2002.

[9] Kongkachandra R., Pansang S., Sripramong T., and Kimpan C., "Thai Intonation Analysis in Harmonic-Frequency Domain," *in Proceedings of IEEE Asia-Pacific Conference on Circuits and Systems*, Chiangmai, pp. 165-168, 1998.

[10] Luksaneeyanawin S., *Intonation in Thai*, University of Edinburgh, Scotland, 1983.

[11] Luksaneeyanawin S., "Speech Computing and Speech Technology in Thailand," *in Proceedings of the Symposim on Natural Language*, Thailand, pp. 276-321, 1993.

[12] Maneenoi E., Jitapunkul S., Wutiwuwatchai C., and Ahkuputra V., "Modification of BP Algorithm for Thai Speech Recognition," *in Proceedings of Natural Language Processing Pacific Rim Symposium*, Thailand, pp. 287-291, 1997.

[13] Parpinelli R., Lopes H., and Freitas A., *An Ant Colony Algorithm for Classification Rule Discovery*, Idea Group Publishing, London, 2002.

[14] Pornsukjantra W., *Speaker-Independent Thai Numeral Speech Recognition Using LPC and the Back Propagation Neural Network*, Chulalongkorn University, Thailand, 1996.

[15] Potisuk S., Gandour J., and Harper M., "Acoustic Correlates of Stress in Thai," *International Journal of Phonetic Science*, vol. 53, no. 4, pp. 200-220, 1996.

[16] Potisuk S., Harper M., and Gandour J., "Classification of Thai Tone Sequences in Syllable-Segmented Speech using the Analysis by Synthesis Method," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 1, pp. 95-102, 1999.

[17] Predawan S., Kimpan C., and Wutiwiwatchai C., "Thai Tone Recognition using Ant Colony Algorithm," *in Proceedings of Information Management and Engineering*, Malaysia, pp. 181-185, 2009.

[18] Ramalingam H., *Extraction of Tones of Speech: An Application to the Thai Language*, Asian Institute of Technology, Thailand, 1995.

[19] Thubthong N. and Kijsirikul B., "Tone Recognition of Continuous Thai Speech under Tonal Assimilation and Declination Effects using Half-Tone Model," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, no. 6, pp. 815-825, 2001.

Saritchai Predawan received his BSc in computer science, Faculty of Science, Burapha University, Thailand, and MSc in science education (computer), Faculty of Graduate Studies, king mongkult's Institute of Technology ladkrabang, Bangkok, Thailand. His interests include artificial intelligence, swarm intelligence, pattern recognition and ant colony optimization.

Chom Kimpan received his PhD in electrical and computer engineering from King Mongkult's Institute of Technology Ladkrabang, MSc in electrical engineering from Nihon University, Japan, and BSc in electrical engineering from King Mongkult's Institute of Technology ladkrabang, Bangkok, Thailand. Now he is an associate professor at the Department of Informatics, Faculty of Information Technology, Rangsit University, Thailand. His interests include pattern recognition, image retrieval, speech recognition and swarm intelligence.

Chai Wutiwiwatchai received his the PhD degree of speech technology, Department of Computer Science, Tokyo Institute of Technology, Japan, M.Eng in digital signal processing, and electrical engineering from Chulalongkorn University in 1998 and his B.Eng (electrical engineering with the 1st honour) from Thammasat University in 1994. He joined the software and language engineering laboratory, national electronics and computer technology center (nectec) as an assistant researcher since july 1998. His expertise is on speech processing especially speech and speaker recognition, which are currently researched at nectec. His interest is on pattern recognition and speech technology.