

## **A study of telephone sales method based on SPIN model combined with intelligent speech analysis\***

Tingting Liu\*\*

Fudan University

Siyu Yan\*\*\*

Hunan Institute of Technology

Xuemei Zhou\*\*\*\*

Shanghai Billow Information  
Science and technology Co. Ltd.

Xingyun Dai\*\*\*\*\*

China Financial  
Futures Exchange

### *Abstract*

Compared with the widespread online sales patterns under modern e-business environment, telephone sales method can enable the salesmen to adjust their real-time dialogue strategies benefited from the voice perception of customer's attitudes and emotions during the telephone conversations, and stimulates the customers to participate, and make a buying decision quickly. This paper studied the application of SPIN model in telephone sales combined with the intelligent speech analysis technology according to the customers' dynamic feedbacks which included their interest and attitude information, and can lead the telephone sales teams to improve their sales efficiency, performances as well as the relationship with customers.

*Keywords:* Telephone sales, SPIN model, telephone speech, intelligent analysis

## **1. Introduction**

Telephone sales method firstly appeared in 1980s, and became a commonly used way for its convenience and intuitional experience in the communication with customers. Compared with the face-to-face sales, it has lower cost, but some limitations. For example, the salesman is not able to observe the customers' reactions through their facial expressions and body languages, and can hardly adjust the sales strategy timely. Those defects will certainly affect the sales efficiency and performance.

Actually, customers' speeches in the conversations of telephone sales usually contain rich information about their preferences, attitude, and emotions (Gong et al., 2015), and can provide very helpful implicit feedback of the customers' reactions beyond their word expressions. However, the above information can hardly be perceived and understood precisely by a salesman without keen awareness and rich experience. Owing to the fast development of artificial intelligence, it's possible that the customers' speeches can be analyzed intelligently through the advanced technologies such as emotion recognition and affective computing (Han et al., 2014; Jiang et al., 2015), and thus offer the indicators of customer psychological responses. Furthermore,

\* This research work was supported International Cooperative Program, Science and Technology Commission of Shanghai Municipality, China(No.13DZ2252200), Online Education Research Center, Ministry of Education of China (Qtone Education, No.2017YB115), and Shanghai Pujiang Program (No.16PJ007).

\*\* 15222010111@fudan.edu.cn

\*\*\* ysy18956@163.com

\*\*\*\* 1643593468@qq.com

\*\*\*\*\* Corresponding author: daixy@cffex.com.cn

the salesman may take the proper strategy and actions supported by the situational-coping knowledge base (Dai, 2012).

SPIN model (SPIN selling method) was developed by Neil Rackham in 1988, who was a sales consultant for Haswaite Company. SPIN is the abbreviation of the four English words: S(Situation Question) stands for background question; P(Problem Question) stands for difficult problem; I(Implied Question) stands for the suggestibility problem; and N(Need-pay off Question) represents the demand-benefit problem. SPIN model provides an effective tool to ask customer's questions and explore customer's potential needs. It achieved remarkable performance in practices, and has been adopted by more than half of the world's top 500 companies for their sales team training.

This paper aims to explore the telephone sales method based SPIN model which utilizes the intelligent technology to analyze customers' speeches, and therefore assists the salesman's work and training. When the salesman communicates with customers by telephone, the real-time intelligent speech analysis can help to better grasp and understand the customers' needs as well as their cognitive and psychological status, so as to facilitate the salesman on these aspects such as adjusting the questions and designing the follow-up marketing strategy timely, according to the process of SPIN model.

This paper is organized as follows. In Section 2, the background of the SPIN model, the intelligent technology use in the telephone selling and the related technical development are introduced as the basis of our research work; in Section 3, the SPIN model and the application in telephone sales are displayed; in Section 4, how to establish corpus, the speaker identification and the intelligent analysis of speeches are discussed; in Section 5, the telephone method based on SPIN model is given out; Section 6 is the conclusion and discussion of this paper.

## 2. Related works

Neil Rackhamin gave a comprehensive introduction to SPIN model in his book titled *SPIN Selling*, which included the origin of SPIN, its theoretical basis, and how to apply in practice. It has become a burgeoning area of research in selling during the past years. Shen (2010) think the sales can start selling process by way of a series of SPIN questions, then explore customers' needs continuously, and finally introduce the products through USP method, and achieve the deal fast. This article emphasizes the role of SPIN model in identifying customer needs and guiding customer purchases. Ding (2008) said the sales can find out sales opportunities quickly through SPIN model. His view provides a new way to use SPIN sales method in this paper. Sun et al. (2003) told us how we can make big deals through answering SPIN questions, he thought the SPIN model can describe a complex process into a simple and precise way, and help you see what you're doing, then find out what you need to achieve most. On the whole, it is important for the sales to collect information about customer's background, requirements, difficulties, and desire for solutions. So the sales can achieve the answers through SPIN questions in the telephone selling.

However, with the development of computer, we think the computers can help the sales to collect and analyze the speech length, volume, amplitude, phonetic features, and judge the attention degree, emotional state, the degree of understanding and attitude, so there are many intelligent analysis technologies in the SPIN model when the sales use it in telephone selling, such as dialogue knowledge corpus, dialogue model analysis, speaker identification, preprocessing of speech, preprocessing of speech feature parameter, emotion recognition and so on.

For this paper focuses on the customer's emotion analysis, because the emotions can well reflect the

customer's attitude, preferences, interests, etc. in a way, so we study mainly for emotional intelligence analysis technology here. Dai (2012) expounded out the machine can have "perception" about human emotions, and can offer services to satisfy human needs according to the human's emotional response. Li (2012) introduced the origin, basis research method about affective computing, and discuss the emotional measurement and the possibility of emotional calculation and theoretical basis, finally it is concluded that the emotional calculation is in order to meet the needs of the practice and development, must study not only the subject about emotion calculation but also psychology. Yu et al. (2007) introduced the current research situation of speech emotion recognition at home and abroad, illustrated the common method in classification of all feelings, and also summed up how to extract speech feature parameters, and how to model, and discussed the future development trend of speech emotion recognition. Zhao and Huang (2014) discussed on several key techniques in the practical speech emotion recognition, and also on emotional modeling, database establishment and emotion recognition algorithm, combined with analyzing the actual application requirements. Wang et al. (2013) proposed to apply speech emotion recognition to the call center work, it can help improve the recording quality inspection, customers' complaints calibration, improve service satisfaction. In addition, in 2014 Ji analyzed the emotional characteristics of customer complaints telephone speech in his master's degree thesis, also by speech emotion theory and PAC theory, and gave the customer complaint handling process on the aspect of application, which is based on the customer emotion and PAC dialogue model. These contents provide great help for the design of speech emotion analysis technology and process of telephone sales in this paper.

From the research, scholars have presented a lot of achievements in intelligence analysis technology area, but the intelligent speech analysis applied to explore in the field of sales is lack relatively, especially combined with the selling model, so we need further research on this field.

### 3. SPIN model and application in telephone sales

#### 3.1 SPIN model and operation procedures

SPIN model is advocated by a series of effective questions and guidance, help the customer express specific requirements, and know what interests and values they will get if they buy the products, then lead the sales staff to achieve sales target step by step.

The sales method based on SPIN model includes a series of typical sales techniques that help salespeople increase their sales effectiveness as shown in Table 1.

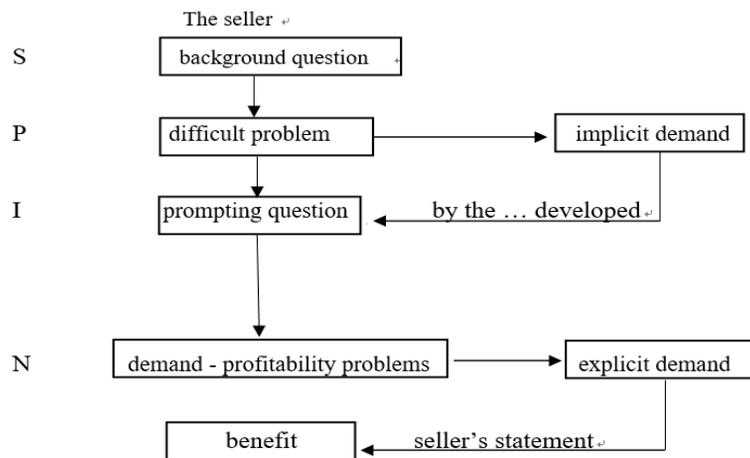
Table 1. The composition of SPIN model

SPIN Model	Purpose	Conclusion	Question Example
Situation Question	It is the background question, by using it for the collection of facts, information and background data of the customer's current situation	Whether the customer is a target customer?	What product is the customer currently using?
Problem Question	It is difficult problem, by asking about the difficulties and dissatisfaction of customers, and guiding customers to express implicit requirements	Find out if the customer has purchase demand	What are the defects in the current product?

SPIN Model	Purpose	Conclusion	Question Example
Implied Question	It is prompting question, by creating the crisis for the customer, and making them to think it is urgent to solve the problem, and the cost is worthy.	Whether the customer buy product immediately?	If you still use the dissatisfaction product, how can you achieve sales goal this year?
Need-pay off Question	It is demand-profitability problems, through these positive and constructive questions, making the customer to know the benefits and interests if they buy the service.	Our products can help the customers solve difficulties, so the customers need to buy our products	If we can help you solve this problem, will you give the chance?

In addition, due to the SPIN sales model query sequence will directly affect the customer desire to purchase, which contains the implicit demand to explicit demand and eventually buying process. Therefore, you need to follow the order which is shown as Figure 1 when using SPIN model.

Figure 1. The operation procedures of SPIN model



According to the above SPIN logic diagram, the sales process needs to be used according to the actual situation. Different sales phases require targeted adjustment and communication based on customer feedback. This feedback includes information feedback and emotional feedback. Adjust the order and content of the SPIN questions according to the information feedback, and decide whether to continue asking questions and using the way of questioning based on emotional feedback. If the client is happy and calm, you can continue to ask background questions and suggestive questions. If customers show anger, surprise, and other emotions, you'd better ask difficult questions and demand benefits.

### 3.2 Application in telephone sales

Telephone sales has many different selling processes. If the customer is an individual consumer, the sales process needs to pass the introduction, expound the product value, deal with opposite opinion from the customer, and promote the transaction. If the customer is the enterprise consumer, because the product is more complex, and the deal amount is larger, so the buyer needs to spend longer time to make the purchase decision, and the telephone sales process also will be complicated relatively.

In this paper, we discuss about how the telephone salesmen can sell products to enterprise consumers. In the present management of telephone sales team, more attention is put on the number of phone dial, telephone duration, the customers hang up rate, sales volume and so on. So people put less attention on customer's emotion as well as the way of communication in the sales process. This is not conducive to sales team management. Therefore, it is necessary to adjust the communication strategy timely based on the SPIN model by perceiving the customer's speech emotion. We designed the sales process as in Table 2.

Table 2. The sales process depend on SPIN model

Sales step	Specify	Sales stage	SPIN model
Step1: Contact non-critical people	Non-critical people are people who can provide basic information, and create sales opportunities. The sales can reach key person by contacting non-critical people.	Preliminary contact	Situation Question
Step2: Touch key person	A key person is someone who can decide whether to cooperate or not, usually the general manager.	Make business with target person	Situation Question Problem Question
Step3: Stimulate customer interest	To understand customer demand, demonstrate products service value, inspire customers to understand and purchase products.	The stage of substantial sales	Problem Question Implied Question
Step4: Recommend solution	When the customers show the cooperation intention, the sales need to recommend solution for the customer.	The selling stage	Need-pay off Question
Step5: Sign the contract	When the customer accepts the service plan recommended by the sales, the sales can sign the contract.	Clinch a deal	Need-pay off Question
Step6: Collection	After the customer and the sales person sign the contract, the customer should pay at least 50% or more, and then the salesman will start to provide the service.	Receive payments	Need-pay off Question
Step7: After-sales service	Provide service to customers, and maintain customer relationship, then conduct secondary sales.	Lay the groundwork for the next sale	Problem Question Implied Question

In the above steps, the sales need make effective communication with customers in each step, so the SPIN model can help the sales people adjust communication strategy timely. At the same time, the intelligent speech analysis can analyze customer purchase attention and mentality, then learn the customers accept, reject or wait-and-see attitude for the product. In the SPIN model, we can use intelligent speech analysis for dialogue knowledge corpus, dialogue model analysis, speaker identification, emotion recognition or others. Because the

emotion can reflect the customer’s attitude, preferences and interests better, so we make research on emotional intelligence analysis.

Generally speaking, when the sales use the SPIN model, they should consider the sales phase and customer emotion. Speech emotion analysis will be introduced in the next section. It is very meaningful to make intelligent analysis of the telephone speech of customers.

### 4. Intelligent analysis of telephone speeches

#### 4.1 Collection of telephone speeches

Because there is a recording system in every Telephone Selling Department, so each phone-call recording can be found in the system, then we have selected 500 speech records randomly from the recording system , and screened the records according to the process as shown in Figure 2 and Table 3.

Figure 2. The screening process of the records

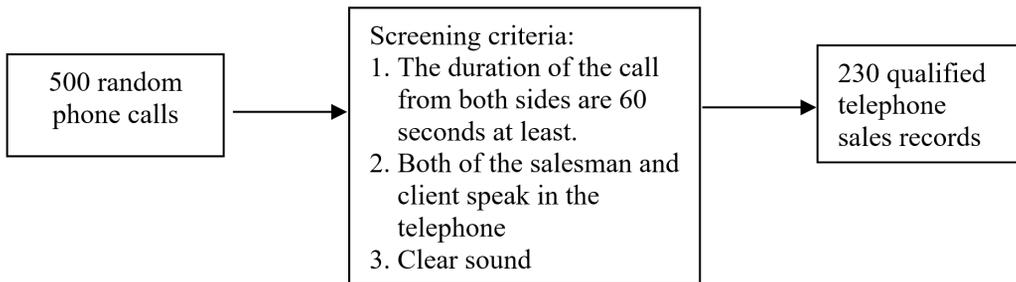


Table 3. The screening process of the records

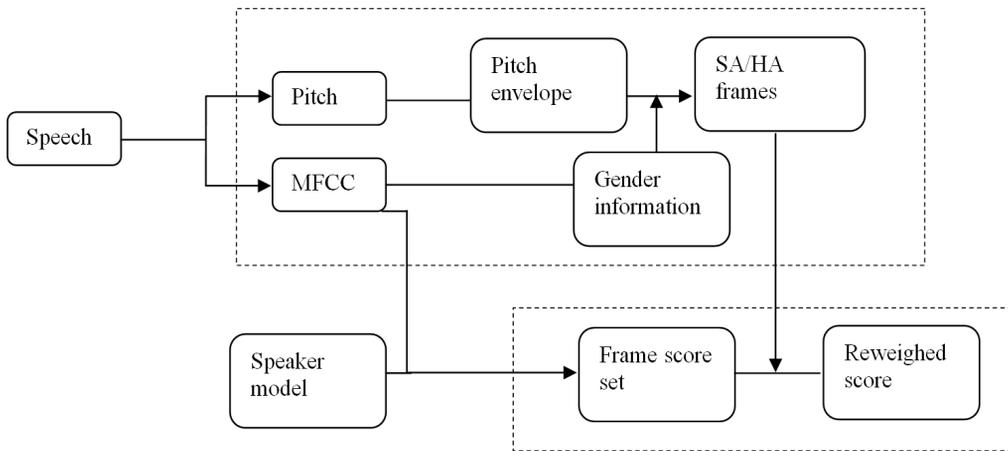
Step number	Sales process	The tape quantity
Step 1	Contact non-critical people	73
Step 2	Touch key person	15
Step 3	Stimulate customer interest	49
Step 4	Recommend solution	28
Step 5	Sign the contract	14
Step 6	Collection	17
Step 7	After-sales service	34

Afterwards, we listen to each call recording carefully, and then analyze the records according to the sales process. We analyzed the talk time, duration of the recording, the customer area, position of the person who answers the call, telephone sales staff working years, customer attitudes, speech emotion and so on , and the basis of this research provides the data for this article.

## 4.2 Speaker identification from conversations

In order to distinguish the salesmen's speech from the conversations, then give the effective selling sentences to salesmen, we should adopt a robust speaker identification algorithm. The cost-sensitive learning technology can finish this task, which can reweight the probability in the pitch envelope level. As shown in the following figure.

Figure 3. Speaker identification algorithm based on cost-sensitive learning technology



The speaker identification algorithm includes three parts: gender identification, the mean value of pitch envelope (PEM) based pitch envelope selection, and frame-level probability reweighted.

First, we need set different PEM threshold for female speakers and male speakers, then make out the likelihood scores based on the MFCC feature, and test them with male and female GMM models. Then match the frames selection thresholds. The PEM can be calculated as follows,  $j$  is a pitch sequence of an utterance:

$$j = \frac{1}{n - m + 1} \sum_{i=m}^n j_i$$

The  $m$  and  $n$  are the numbers of the start and the end frame of a pitch envelope, and  $j \neq 0$ ,  $0 \leq n \leq i \leq m$ .

Second, we analyze the pitch distribution under different emotion, and compare with PEM threshold. The voiced part that is slightly affected (SA) together with the unvoiced should be chosen for reweighting. The process as follows:

$$\text{score}(X | \lambda_i) = \sum_{t=1}^{T1} f(\log p(x | \lambda_i)) + \sum_{T=1}^{T2} \log p(x | \lambda_i)$$

$T1$  is the number of frames in the part that is slightly affected (SA) and unvoiced part,  $T2$  is the number of frames in HA part.

Third, we strengthen the confidence of the selected speech segments and make the final frame scores better through the score reweight.

### 4.3 Intelligent analysis of speeches

In Section 3, we said the emotion can reflect the customer's attitude, preferences and interests, so we mainly make research on emotional intelligence analysis here. The emotional intelligence analysis of speeches includes three parts: preprocessing of speech voice, processing of speech feature parameter, emotion recognition.

#### 4.3.1 The basic classification of emotional types

However, the premise of speech emotion recognition by computer is must clear emotional categories, what's the feature of different types of emotional, then the emotion can be identified. The basic classification of emotional types is as shown in Table 4.

Table 4. Basic emotion classification

The researchers	Basic emotion
Plutchik	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise
Arnold	Anger, disgust, courage, depression, desire, despair, fear, hope
Frijda	Desire, joy, interest, surprise, sadness
Gray	Anger, fear, anxiety, joy
Izard	Anger, contempt, disgust, pain, fear, guilt, joy, surprise
James	Fear, sadness, love, anger
Mowrer	Pain, pleasure
Oatley/Johnson-Laird	Anger, disgust, anxiety, sadness, happiness

#### 4.3.2 Preprocessing of speech voice

In order to get more accurate identification results, the speech emotion signal need be preprocessed to make it easier to identify. The preprocessing of speech emotion signal includes pre-accentuation of signal, framing, add window, signal endpoint detection and so on.

##### 4.3.2.1 Pre-accentuation of signal

The purpose of preprocessing the signal is to improve the high frequency resolution of the speech signal, so it is mainly to aggravate the high frequency part of the signal and eliminate some low frequency interference. In general, the preweighted filter is used to improve the high frequency part, in which the first-order high-pass digital filter is used more, as shown in formula 1.

$$H(z) = 1 - \alpha z^{-1}$$

In the above formula,  $\alpha$  is a preemphasis coefficient, its value is close to 1.

#### 4.3.2.2 Framing and add window of signal

In order to identify speech data more clearly, the speech signal need be processed in frame. By dividing a piece of the speech signal into a number of short periods, the smallest unit is called a frame, and the two frames that have no overlap are called frame shifts. The speech signal frame is realized through the window function, and more commonly used are rectangular Windows, hamming Windows and hanning Windows.

(1) Rectangular window, function expression is as follows:

$$w(n) = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{others} \end{cases}$$

(2) hamming window, function expression is as follows:

$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{others} \end{cases}$$

(3) hanning window, function expression is as follows:

$$w(n) = \begin{cases} 0.5 \left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right), & 0 \leq n \leq N-1 \\ 0, & \text{others} \end{cases}$$

In the formulae above,  $w(n)$  represents the window function, and  $n$  is the frame length. The window function of different shape and length has obvious influence on the analysis parameters. In this paper, we use the hamming window to analyze the speech signal of telephone sales.

#### 4.3.2.3 Signal endpoint detection

The purpose of endpoint detection of speech signals is to determine the effective information in speech, where the starting point and ending point are, to reduce the computation of computer inefficiency and improve the identification efficiency.

In the endpoint detection, commonly used are short-term energy and short-term average pass zero rate, and then the critical point of the sound is found through a first-order judgment and second-order judgment. Level decision depends on the short-time energy signal waveform and the average energy of noise, as well as their high threshold and low threshold, when the signal reaches a certain strength, as long as it is more than a high threshold, which contains the speech signal. Thus, the location of the speech signal can be determined.

#### 4.3.3 Processing of speech feature parameter

Telephone speech for emotional feature extraction is a very important step, only to extract the appropriate emotional characteristics, it will be effective for emotion recognition.

##### 1. Amplitude energy

The amplitude energy is what we usually call the intensity or volume of speech, and when people's emotions are high or low, the amplitude energy behaves differently.

## 2. Pitch frequency

The base tone frequency is the reciprocal of the periodic vibration of speech, and is also one of the most important characteristics of speech emotion.

## 3. The formant

When the voice from the vocal tract can cause the vibration of the vocal tract, that is a frequency of the vibration, and the change of frequency may change as the channel change, due to uneven sound energy, strong vibration will be like a mountain, so the vibration frequency is called resonance peak.

## 4. MFCC

MFCC is the characteristic parameter proposed by Davies and Mermelstein based on the correlation between human auditory principle and inverted spectrum, which is called the Mel frequency inverted spectral coefficient. The Mel frequency is expressed as follows:

$$f_{meL} = 2595 * \log (1 + f/700)$$

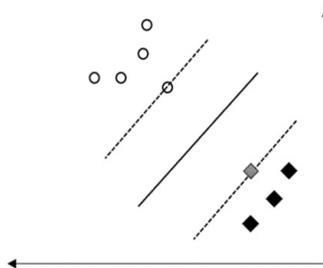
It can be seen from the above formula that the human ear has logarithmic relationship to the speech frequency.

### 4.3.4 Speech emotion recognition

As we discussed above, we should firstly identify the speakers through the algorithm based on the cost-sensitive learning technology. Then we adopted the affective computing technology based on SVM methods and recognized the dynamic changes of the clients' emotion from the extracted feature parameters.

Support vector machine is a nonlinear way to map input vectors to higher latitudes, and then classify the samples after calculation. As shown in Figure 4.

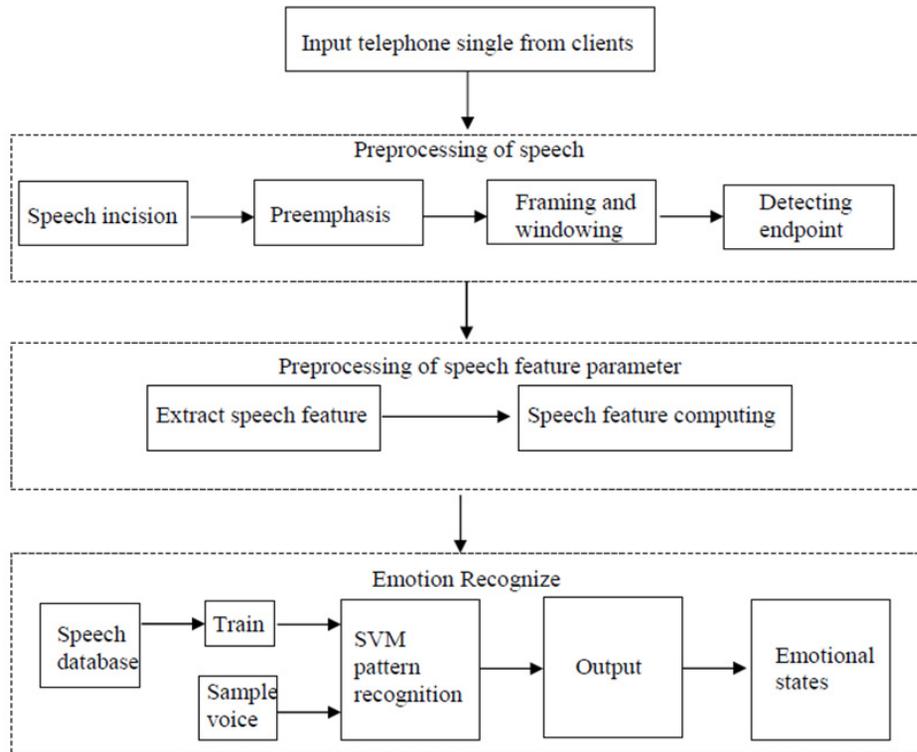
Figure 4. Support vector machine model schematic diagram



Support vector machines (SVM) are separated by two-dimensional plane to obtain the optimal solution of the existing information, thus obtaining a higher recognition rate.

### 4.3.5 The process of intelligent analysis of speeches

The process of emotion intelligent analysis of speeches as shown in Figure 5.



## 5. Telephone sales method based on SPIN model

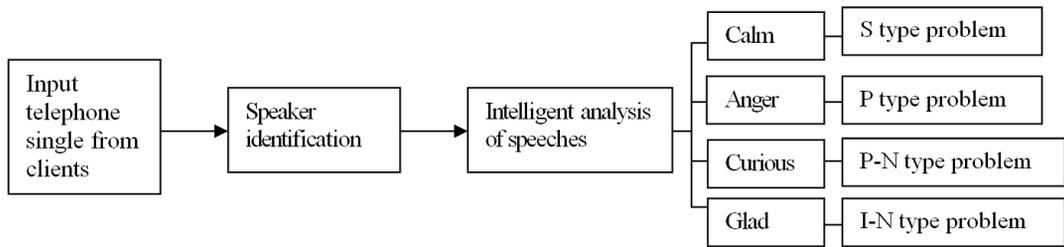
### 5.1 Telephone sales method based on SPIN model

As shown in the above, the salesperson can promote customers to buy the product with sales of sentences according to the customer's emotional responses and SPIN dialogue. Therefore, the telephone sales method based on SPIN Model needs to complete two levels of analysis results: the emotion intelligent analysis of the speeches and the suitable SPIN sentence match to the speech emotion. The types of SPIN selling script are detailed in the following Table 5 for example.

Table 5. The type of SPIN selling script

The type of SPIN selling script	Example
S type problem	How many employees are there in your company?
P type problem	What's your dissatisfaction with the present service?
P-N type problem	I have a solution, do you want to know?
I-N type problem	If you don't increase your AD spending, how can you finish this year's sales target?

Figure 6. The process of the telephone sales method based on SPIN Model combined with emotional intelligent analysis



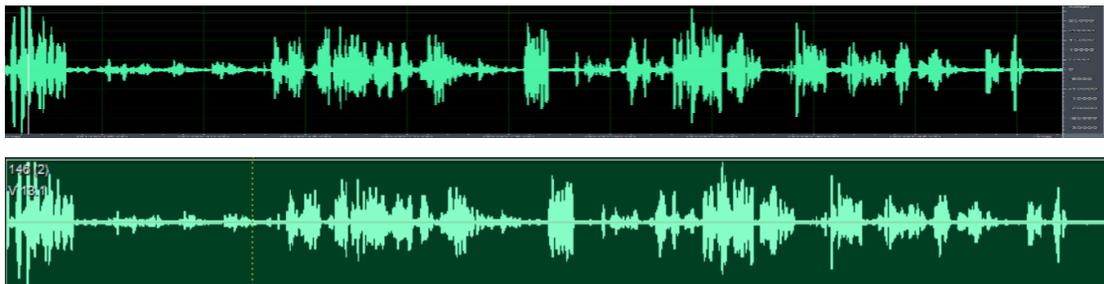
As shown in the above, the salesperson can promote customers to buy the product with sales of sentences according to the customer's emotional responses and SPIN dialogue.

### 5.2 Application of the telephone sales method

In the telephone selling department which in the company give services to enterprise consumers, the sales should call a lot of clients every day. Now, we can take their speech database and the records as experiment data. The records from the telephone selling department are real speech samples saved as wav files. Each sample contains a whole telephone selling conversation process between the salesman and the customer. We choose one record randomly, then use emotion intelligent technology to analyze it.

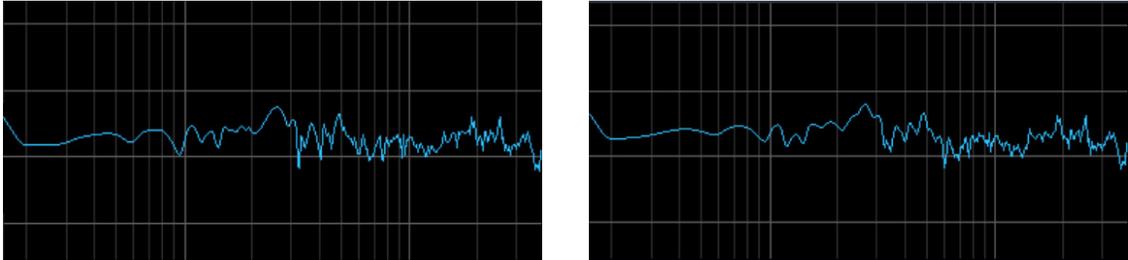
In the speaker identification, male and female model are built with using the speech from development subset, the statistical F0 thresholds of PEM for the female and male speakers are set out. Then we make preprocessing of the speech, in the following Figure 7, the expected emotional aggravation of the before and after comparison diagram.

Figure 7. Anticipated emotional aggravation before and after comparison chart



Since the frame length is a relatively smooth part between 10ms and 30ms, the frame length is 245 and the frame is 1/2. Then use the hamming window to add window processing. As shown in Figure 8.

Figure 8. The text is added to the picture after the window



In this paper, MFCC and SVM are determined to be used for pattern recognition, and the following identification results are obtained by combining the characteristics parameters of MFCC, as shown in Table 6.

Table 6. The recognition rate of MFCC characteristic parameters on speech emotion

The telephone sells voice emotion	calm	anger	curious	glad
MFCC parameter recognition rate	80.1%	81.4%	72.7%	79.5%

Therefore, we know when the speech from the client has anger emotion, the salesman should ask P type problem, such as “Why don’t you want to realize my product? Do you use it before?”, then if the client becomes calm, the salesman can say S type problem, such as “I know your company is very good, how many products in your company?”

## 6. Discussion and conclusion

In conclusion, this paper studied the telephone sales method based on the SPIN model combined with intelligent speech analysis. It provided an intelligent method to analyze different customers’ speeches to better understand their needs, cognitive, psychological feedback and attentions through the calculation of machine learning, and therefore help the sales to adjust the order of SPIN questions and judge the purchase possibility of the customer. The proposed method is worthy of further research, such as how to classify the speech emotion more accurately, how to make the sales strategies more effectively, and so on. We’ll carry out more improved researches in these areas.

## References

- Dai, W. H. (2012). Situational awareness and emotional intelligence: the wisdom gate to smart city. *Shanghai City Management*, 2012(4).
- Ding, X. L. (2008, July 7). The five principles of SPIN use. *Electromechanical Business Daily*, A08.

- Gong, S. P., Dai, Y. H., Ji, J., Wang, J. Z. and Sun, H. (2015). Emotion analysis of telephone complaints from customer based on affective computing. *Computational Intelligence and Neuroscience*, 2015, 1-9.
- Han, W. J., Li, H. F. and Ma, L. (2014). Progress review on speech emotion recognition research. *Software Journal*, 25 (1).
- Jiang, H. H. and Hu, B. (2015). Speech emotion recognition based on PCA and SVM in mandarin. *Computer Science*, 42(11).
- Li, J. Y. (2012). Research status and cognitive dilemma of affective computing. *Natural Dialectics Communication*, 34(2).
- Rackham, N. (1988). *SPIN Selling*. McGraw-Hill Education.
- Shen, H. B. (2010). Sales skills integration: using the USP technology and the FAB rule. *China Business*, The 12th issue of 2010.
- Sun, Y. P., Shen, M. H. and Yu, L. X. (2003) SPIN teach you ask the big deal. *The Way of Marketing*.
- Wang, W., Yang, L. P., Wei, L. and Liu, Y. (2013). Extraction and analysis of speech emotion characteristics. *Laboratory Research and Exploration*, 2013(7).
- Yu, L. L., Cai, Z. X. and Chen, M. Y. (2007). An overview of the analysis and recognition of emotional characteristics of speech signals. *Journal of Electrical Circuits and Systems*, 12.
- Zhao, L. and Huang, C. W. (2014). A number of key techniques in practical speech emotion recognition. *Data Acquisition and Processing*, 2(29).