1 Introduction

Restatement of the problem

In order to expand its international business, a multinational service company needs to set up international offices at different places overseas to meet the needs of different customers and promote the development of the company.
We mainly face five questions as follows:
1. Simulate the distribution of speakers of various languages over time.
2. Predict changes of the number of native speakers, and second language speakers in the next 50 years, and whether the current top ten languages will be replaced by others in the future.
3. Predict the global population and migration patterns to determine the geographical distribution of languages.
4. Choose six places for the multinational service company to set up international offices.
5. Whether the number of international offices opened can be reduced?
To this end, we have established a geographic, economic and trade relations model to predict the distribution of language speakers. The model predicts native speakers and the total number of language speakers in the next 50 years by using the gray prediction model and compares it with the current situation. In the proposed model, population migration model is used to determine the geographical distribution of language. The Analytic Hierarchy Process (AHP) was applied to determine five major indicators as the basis for the selection of six major international offices. The level of communication in different countries was used to judge whether the number of international offices opened can be reduced.

2 Model One: Language Distribution Model

The model used to predict the change of different languages in the next 50 years. It simulates the distribution of language speakers over time and the geographical distribution of different languages in the same period.
Languages populations mainly consist of two parts: native speakers and second language speakers., as given in (1),

$$P_T=P_{L1}+P_{L2}, \tag{1}$$

where P_T is the total language population of a given language, P_{L1} and P_{L2} are the number of native speakers and second language speakers of the language. P_{L1} is obtained by the model described in section 2.2, and P_{L2} is calculated by the geographical language distribution model and economic and trade relation based language distribution model as introduced in section 2.3 and 2.4, respectively.

2.1 Geographical language distribution model

We mainly consider the establishment of a geographic distribution graphic model and a economic trade relations model to solve the problem of simulating speaker distribution over time in different languages.

2.1.1 Assumptions

Assume that the distance between countries is based on the distance of the capital.
Symbol Table 1 . s, and that capitals will not change in all countries in 2017-2025
Assume that all countries will not be let go by irresistible factors such as war during 2017-2025.
Symbol table1

Symbol	Definition
Constant	
R	Earth radius
$\mathrm{MLat_i}$	Latitude of language area
$\mathrm{MLon_i}$	Longitude of language area
C_{ij}	The angle between language area and language area j from the Earth's center of the globe
D_{ij}	The distance between language area and language area j
D_{max}	The maximum distance value of two language areas
α_{ij}	A coefficient that measures the impact ratio of language zone j using language as a second language.
Variables	
P_i	The population of the language area .
P_{ij}	The population of language area j use language i as the second language.

This model focuses on geographical distribution, and defines the principle that both the neighborhood is the minimum and the distance and the distance are the minimum. According to the calculation of the distance of two language interval, it examines whether each language area are adjacent or not, and uses the other language area in the X area as the second language ratio. And multiplied by the number of people in the language zone to obtain the number of second language learners who learned the language in other language zones and determine the distribution of the second language.

The model starts from the distance between capitals of different countries. Through Google Earth, we can accurately locate the longitude and latitude of each

capital.Acording to the latitude and longitude of two points and distance conversion formula (16) to get the distance D_{ij} between two countries:

$$C_{ij}=\sin(MLat_i)*\sin(MLat_j)*\cos(MLon_i-MLon_j)+\cos(MLat_i)*\cos(MLat_j), \quad (2)$$

$$D_{ij}=\frac{R*Arccos(C_{ij})*Pi}{180}, \quad (3)$$

Find the minimum distance between capitals of two different countries through the shortest path algorithm. If it is need to calculate the number of countries in which you want to learn j as a second language, then we need to define a second-language propagation factor (18) as a weight.

$$\alpha_{ij}=1-\frac{D_{ij}}{D_{max_{ij}}}, \quad (4)$$

The model imitates the population of language area in different period that as the base number of the learning language j multiplying the propagation factor (18) and use it as the second language population of language j at different periods.

$$P_{ij}=P_i\alpha_{ij}, \quad (5)$$

Finally, the number of second language learners who learn j language in other language regions is obtained to determine the distribution of j language as the second language.

2.1.2 Model testing

We chose countries in western Europe, southern European and Denmark (The reason for considering Denmark is that Denmark borders on Germany.) and excluding Greece. Since Britain, France, Portugal, Spain, Denmark, and Italy have their own distinct native languages, the capitals of each country are taken as six points on the graph model, , respectively. As the Netherlands and Belgium are both native Dutch speakers, Germany, Switzerland and Austria are German native speakers. Therefore, the Netherlands and Belgium are translated into a Dutch-speaking region and Germany, Switzerland and Austria into one German-speaking region. Capitals of Netherlands, Germany are treated as two points on the graph model. The graph model is shown in Figure 1.

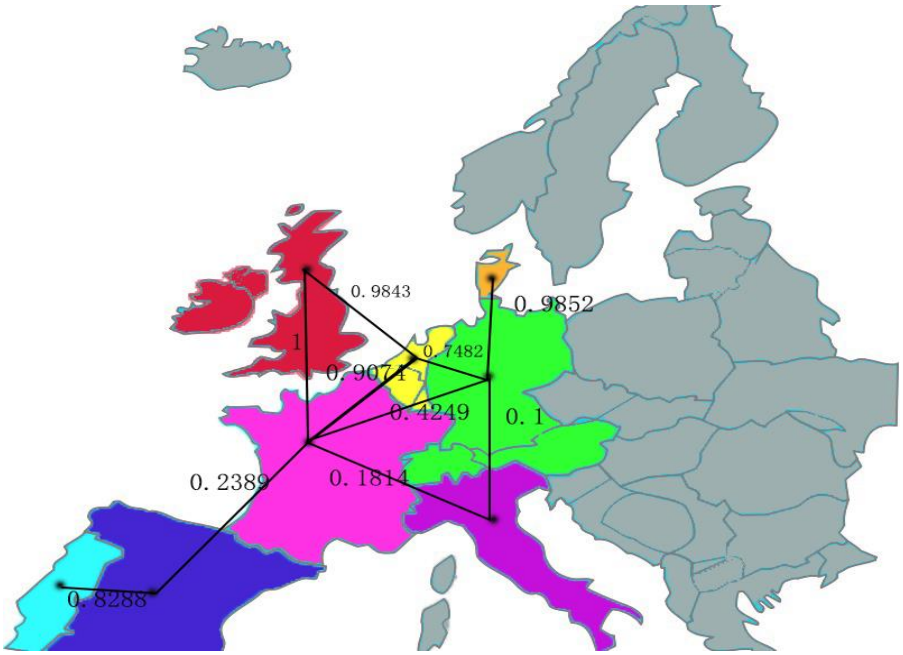

Figure 1 The influence of a language on other language areas

Portugal Spain Italy France

United Kingdom Denmark German area Dutch area

The computation process of the geographical language distribution model is as follows: Firstly, we use the formula (18) (19) to calculate the propagation factor of the second language as shown in the figure as a weight, which has been marked on the map. Secondly, calculate the population in the eight language regions from 2017 to 2050 on the website and obtained the total number of people who use them as a second language respectively according to the formula (19) (20). (On the following table shows the second language population.)

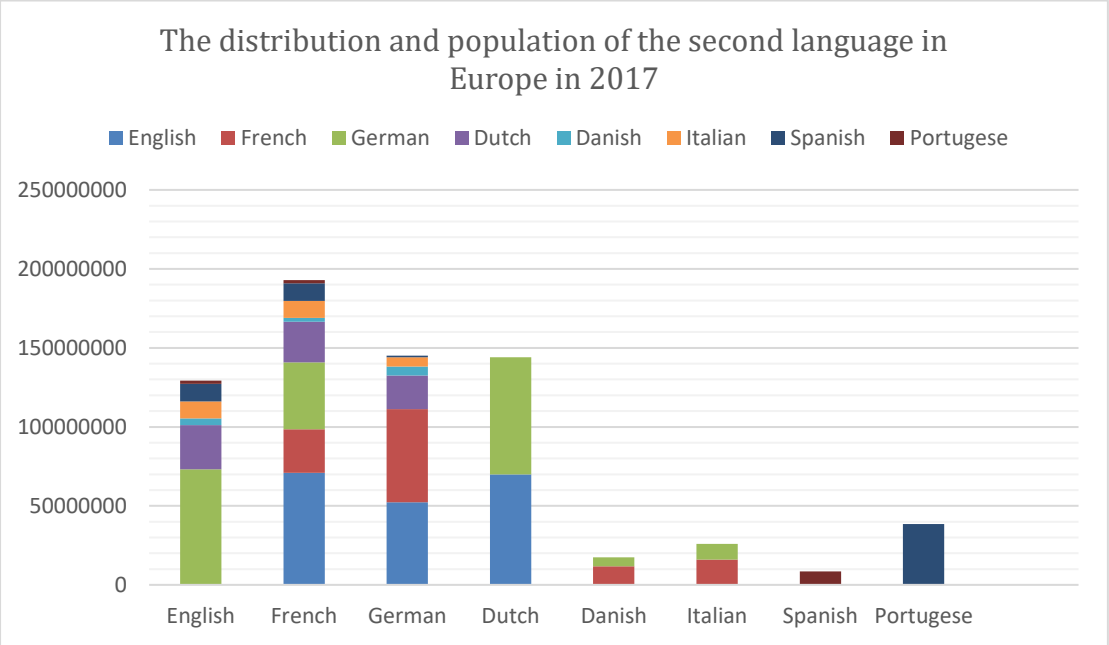

Figure 2 The distribution and population of the second language in Europe in 2017

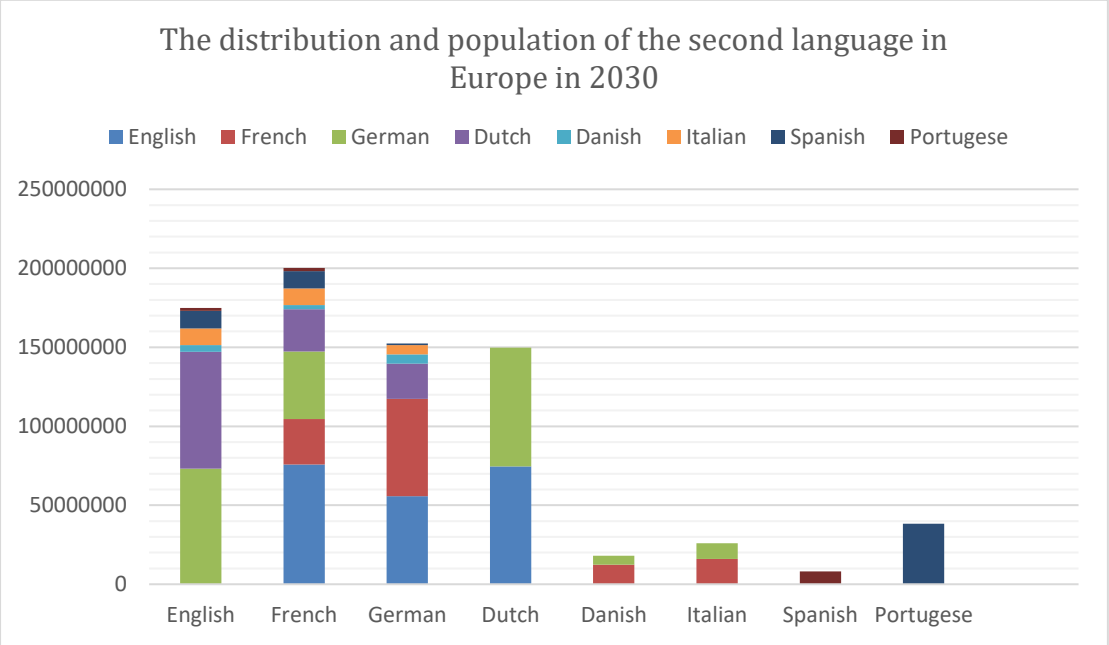

Figure 3 The distribution and population of the second language in Europe in 2030

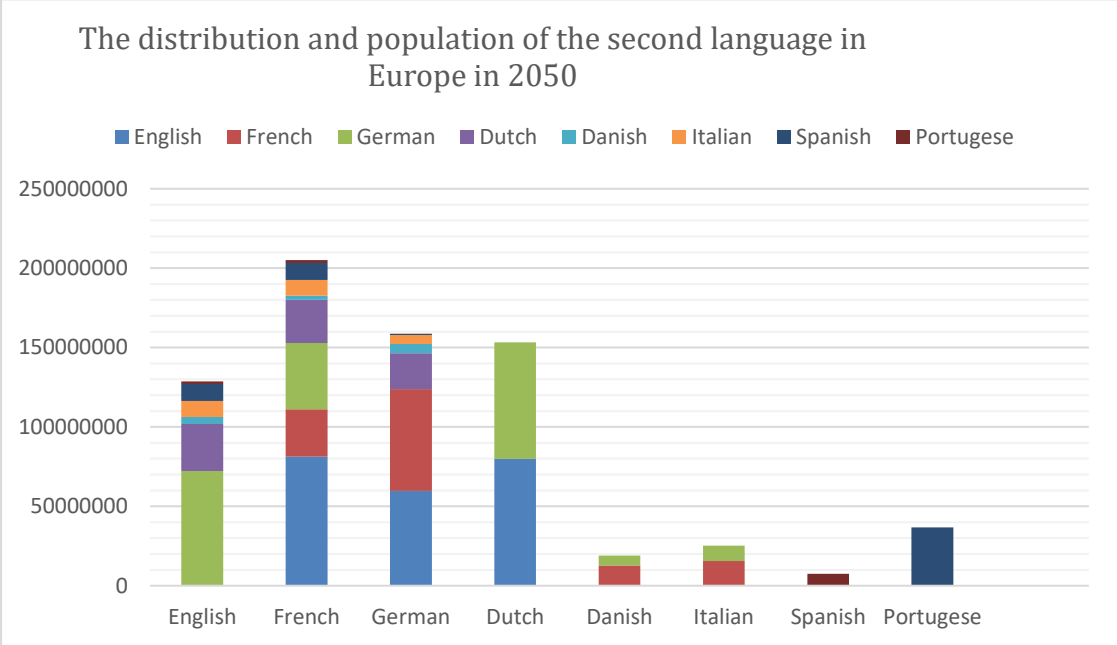

Figure 4 The distribution and population of the second language in Europe in 2050

Finally, we can get the distribution of language areas in other language areas. Let's take English as an example. (The most deep color is the English as the native language area; The lighter color is the area as a second language where the propagation factor is greater than 0.5； The lightest color is the area as a second language where the propagation factor is less than 0.5.)

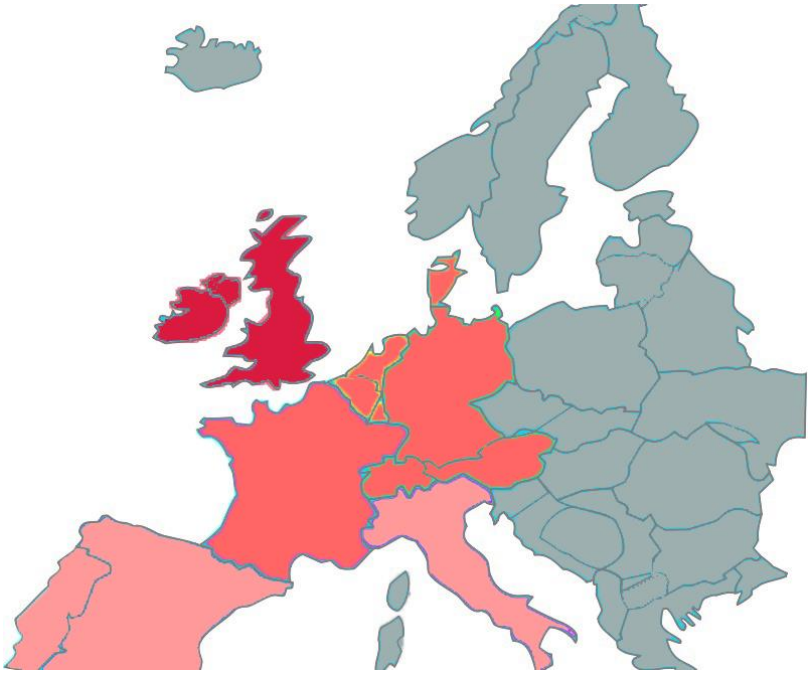

Figure 5 Distribution of English in this area

Thus, we can see the distribution of English, the distribution of other languages can be derived from the analysis process in the same way.

2.2 Language Distribution Model Based on Economic and Trade Relations

The model ranks the top nine countries in the global economic power as nine economies, and each economy has its own mother tongue. Through the economic and trade volume between different economies, the influence factors of the economic bodies are defined, and then the distribution of the second languages in different economies is obtained.

2.2.1 Definition and formula

Symbol Table2.

Symbol	Definition
Constant	
ω_{mn}	Language area m 's total imports of language area n's share
Variables	
P_m	The population of language area m
P_{mn}	Number of native speakers who use the language m as the native language of the language area n as the second language

Economic development is a strong pillar of the country. Economic and trade relations are important measurements of the economic influence of a country or region on other countries. Our model focuses on economic and trade relations, synthesizes rankings on different websites, and measures the share of one economy's exports to other economies and the population of the economy. Improve maximum density subgraphs, build our strongest economic impact model, and measure the total number of second language learners through this model.

First determine the population of n economies, logarithm it and define it as $\log P_{n_{econ}}$. The data obtained easily eliminate heteroscedasticity problems ; at the same time, take the logarithm, the variable flexibility. It is obtained by multiplying the share of exports of the n economies in m economies

$$P_{mn}=\log P_n * \omega_{mn}, \tag{6}$$

Through the proportion of different languages in the region , and then get the distribution of speakers in various languages in the region.

2.2.2 Model testing

After our careful choice, there are mainly six major economies, namely the United States, the European Union, China, Russia, Japan and India. Among them, the four major countries in the EU have the higher economic strength, and in the model they are regarded as independent economies, namely Britain, Germany, France and Italy. In other words, there are nine major economies in our model.

Figure 6 shows predictions of the total population of each country's mother tongue as the second language of other economies. (2017, 2030 and 2050)

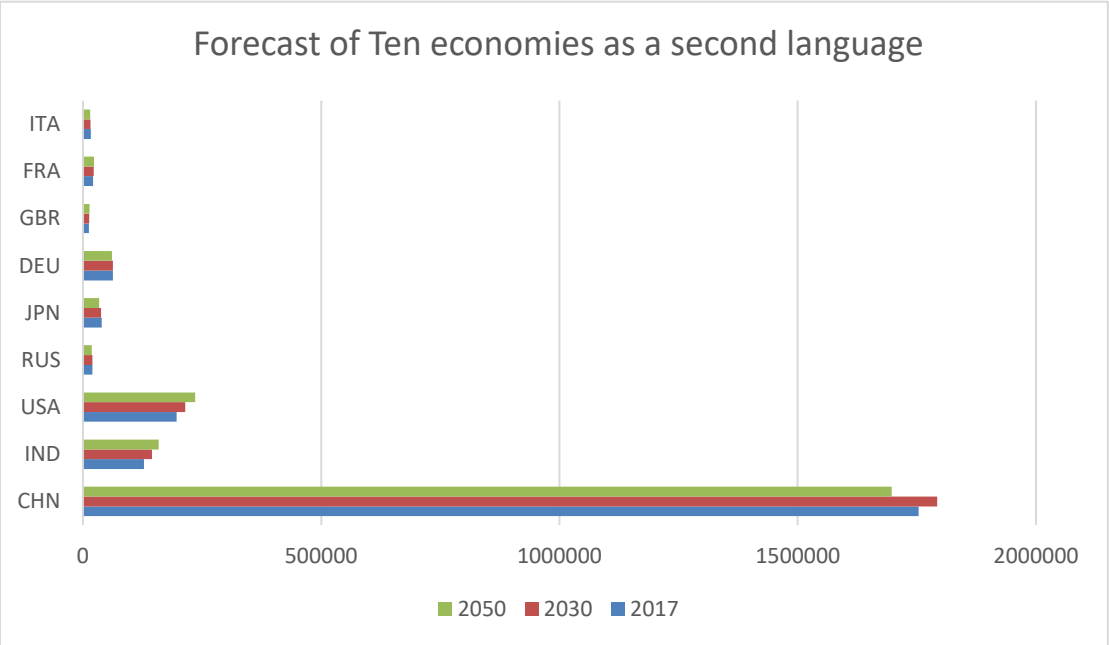

Figure 6 Forecast of Ten economies as a second language

To show the distribution of second languages more clearly, let's take Germany as an example. The figure below shows the distribution of second languages in Germany. Obviously, China and the United States, as the two largest foreign trade partners in Germany, have a very strong influence on their native languages in Germany. The foreign trade relations between EU countries are very steady.

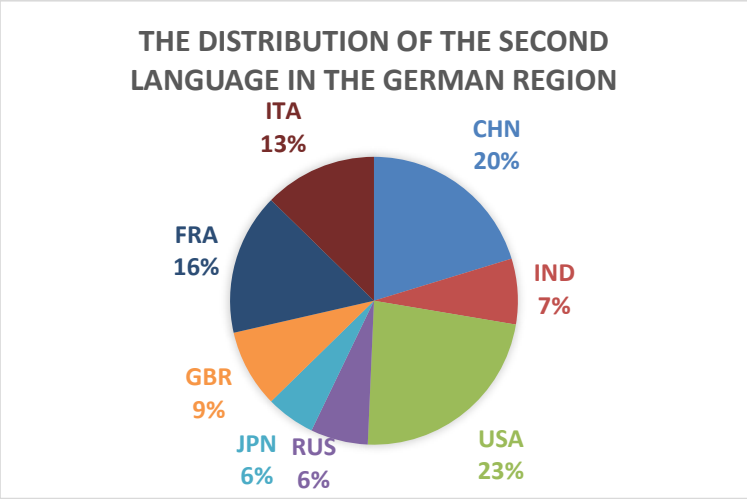

Figure 7 The distribution of the second language in the German region

2.3 Native language population prediction model

After analyzing native-speaker population data from the 20 most spoken languages of the past 50 years, a gray prediction model is used to estimate the number of native speakers of the 20 languages predicted over the next 50 years.

2.3.1 Reasons for choosing the gray prediction model

We have obtained data from native-speaking population in the top 20 languages for the past 50 years (these statistics are available every 5 years). Since only the data with the statistical years are separated, there is no strong connection between the resettlement data, and thereby Linear Regression model is not suitable for this problem. Since complexity and huge data collection work required by the neural network model, we neglect it for predicting native language speakers. In the view of current situation, we devise a Grey Forecasting Model (GFM) to obtain data with higher reliability. The advantage of using Grey Forecasting Model is that it allow us to obtain more reliable results when lacking accessible data, which is perfectly fitted with our current situation.

2.3.2 Correlation Degree Analysis

In the grey forecasting, we try to find and grasp the law of development of the native

language population data in 20 languages over the past 50 years data and at last make a scientific quantitative prediction for the future condition of the system by raw data processing and grey model building. Currently, the gray forecasting model GM (1,1) is the main application of grey forecasting, but GM (1,1) model is applicable to sequences with strong exponentially, and can only describe the monotonous process of change. The numbers of people prediction model is a dynamic time series system with some random volatility, and therefore it is more suitable for us to use Verhulst model for non-monotonic swing development sequence.

2.3.3 Model Solution

We define $x^{(0)}$ as the original data sequence of the population of top 20 languages' native language in year 1962-2016:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)), \tag{7}$$

Then we can get the whitened equation of Verhulst model,

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(t) = b, \tag{8}$$

where the $x^{(1)}$ is the accumulated generating operation sequence of $x^{(0)}$. After that, we use the least square methods (LSM) to get the parameter a and b as:

$$\hat{u} = (a, b)^T = (B^T B)^{-1} B^T Y, \tag{9}$$

where

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad and \; Y = \begin{bmatrix} x_2^{(0)} \\ x_3^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix}, \tag{10}$$

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1), \tag{11}$$

The respective time response sequence of Verhulst model is:

$$x^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}, \; k = 1,2,3,\ldots,n-1, \tag{12}$$

And we can get the reduced $\hat{x}^{(0)}$ by repeated decreasing:

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), \; k = 1,2,\ldots,n, \tag{13}$$

2.3.4 Model analysis

We test our model by using residual analysis. Define gray forecast sequence as,

$$\hat{x}^{(0)} = (\hat{x}_1^{(0)}, \hat{x}_2^{(0)}, \ldots, \hat{x}_n^{(0)}), \tag{14}$$

and residual sequence as

$$\varepsilon^{(0)} = (X_1^{(0)} - \hat{x}_1^{(0)}, X_2^{(0)} - \hat{x}_2^{(0)}, \ldots, X_n^{(0)} - \hat{x}_n^{(0)}), \tag{15}$$

Then, we get relative error sequence:

$$\Delta = (|\frac{E_1}{X_1^{(0)}}|, |\frac{E_1}{X_2^{(0)}}|, \dots, |\frac{E_1}{X_n^{(0)}}|), \quad (16)$$

Finally, we get the Average relative error sequence as shown below:

$$\bar{\Delta} = \frac{1}{n}\sum_{k=1}^{n}\Delta_k, \quad (17)$$

2.3.5 Results and Conclusions

In the first 20 languages, the distribution of second language in English, French and Spanish is fragmented and has a wide range of influence, mainly because other countries study them as a second language and are native speakers of other countries, so we use the geographic model to learn their second language population. Other languages are basically only used by their own countries and they have a relatively small impact on other countries. Therefore, the economic and trade relations model is adopted[1].

The table below lists the prediction of native language speakers after 50 years（sorted by the number of native speakers), ranking Turkish from No. 16 to No. 9, the rest of the language sequence remained unchanged. The reason for this analysis is that Turkey is geographically close to the Middle East, has a large number of refugees and war immigrants, and has a relatively developed economy in the Middle East[2].

Table 1 Sort by native language in the next fifty years

Ranking	Country	L1	L2	Total(Million)
1	Chinese	889.32	205.35	1094.67
2	Spanish	675.00	490.88	1165.88
3	English	378.07	662.2	1040.27
4	Indic	677.98	406.62	1084.60
5	Portuguese	236.22	27.26	263.48
6	Bengali	224.34	4.27	228.61
7	Russian	160.66	28.6	189.26
8	Japanese	131.45	33.83	165.28
16	Turkish	123.48	6.09	129.57
9	Javanese	100.28	29.88	130.16
12	Korean	90.33	8.43	98.76
13	French	89.64	412.4	502.04
10	German, Standard	74.19	134.6	208.79
20	Italian	65.66	34.21	99.87

Table 2 is based on the speaker's total ranking, French ranked fifth from No. 13. The reason is that France's central hub in Europe is widely used in Europe. And because early colonial French was quite influential in Africa, with many native languages in Africa, French was very popular as a common language

Spanish overtook Chinese as number one. The reason is that the population in China is declining due to the influence of the family planning policy. And Spanish as a common

language is very popular in Latin America.

Table 2 Sort by total language in the next fifty years

Ranking	Country	L1	L2	Total(Millon)
2	Spanish	675.00	490.88	1165.88
1	Chinese	889.32	205.35	1094.67
4	Indic	677.98	406.62	1084.60
3	English	378.07	662.2	1040.27
13	French	89.64	412.4	502.04
5	Portuguese	236.22	27.26	263.48
6	Bengali	224.34	4.27	228.61
10	German, Standard	74.19	134.6	208.79
7	Russian	160.66	28.6	189.26
8	Japanese	131.45	33.83	165.28
9	Javanese	100.28	29.88	130.16
16	Turkish	123.48	6.09	129.57
20	Italian	65.66	34.21	99.87
12	Korean	90.33	8.43	98.76

2.4 Population migration model

2.4.1 Assumption

Assume a positive inflow which indicates that the language area is the destination of immigrants. Assuming the hypothesis is negative inflow, it indicates that the language zone is the source of the immigrants, so that the migration pattern of the population can be analyzed.

2.4.2 Data and formula

We have found the global population of the United Nations for the next 50 years.

Table 3 The origin of immigrants

Country	2030	2040	2050
Spain	-188679407	-506374081	-1363004344
Bengalese	-57773179	-83922632	-122146396
Hindi	-55004383	-80693486	-118380360
Mandarin	-2663441	-3423167	-4399599
Javanese	-1495159	-1954538	-2555060
Language	2030	2040	2050
Japanese	6474	4026	2504
Portugese	1050739	1520537	2175086

Russi	1991541	2347963	2758025
German	2165123	2523628	2983475
English	17563763	21778180	27465015
Turkey	14716642	21269438	30737424

Table 4 Immigration destination

2.The World Bank has provided net emigration of all countries over the past 50 years, and we use the gray prediction model to predict the net immigration of the population over the next 50 years[7].

Symbol Table 3 .

Symbol	Definition
m_i	Net migration of a country with i language as the mother tongue.
M_i	Summation of net migration of all countries in which language is the native language.
W_{ij}	The total pull factor is the direction of immigration.
α_{ij}	Geographical Rally Factor
ω_{ij}	Economic Rally Factor
β_α	The weight of influencing geographical rally factor.
β_ω	The weight of the economic pull factor.
P_{ij}	The migrated population of Language area i to language area j .

In table 3, m_i represents the net migration of a country in which i language is the native language. The displaced people do not change their mother tongue, but learn to move into the foreign language as a second language, thereby affecting the language of the country they move to. M_i is used to determine the migration pattern (source or destination) of the speaker in the language.

$$M_i = m1 + m2 + \ldots\ldots + m_n, \quad (18)$$

where n represents the number of countries in which language is the mother tongue. The total tension factor(W_{ij}) represents the direction of immigration, from a geopolitical and economic point of view. Including geo-tension factor and economic tension factor. Among them, the weights β_α and β_ω determine the two factors in W_{ij} share.

$$W_{ij} = \beta_\alpha \times \alpha_{ij} + \beta_\omega \times \omega_{ij}, \quad (19)$$

Language area represents the source of immigrants. The language zone j is the destination of immigrants. P_{ij} denotes the migrated population from the language zone i to the language zone j.

$$P_{ij} = W_{ij} \times M_i, \quad (20)$$

Through the statistics to move out of the language area, into the language area, only consider the direction of moving into the language zone, calculated top14 languages.

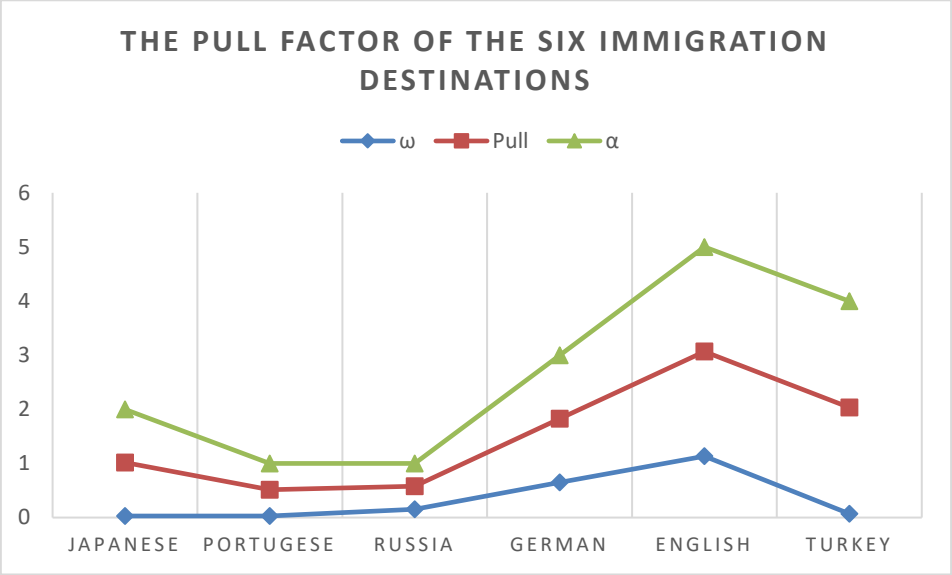

Figure 8 The pull factor of the six immigration destinations

3 Language Influence Assessment Model

3.1 Office location model

3.1.1 Overview

In a multilingual and globalized society, language is the prerequisite for us to be able to communicate with others and to allow us to participate in the social, cultural and economic activities. Which languages are most useful? Now a large multinational service company with offices in New York City, the United States and Shanghai, China, is expanding to establish six international offices in different countries around the world. Which languages will you use in these six offices? This will be problems we will solve in this model.

The model uses analytic hierarchy to measure the "usefulness" of a language, and we define a language proficiency index that assigns weights. Broadly speaking, there are five main factors that affect the usefulness of the language. They are: geography, economy, exchange, knowledge and media, and diplomacy.

In a narrow sense, the five factors are each measured by different indicators (see the following table). In other words, the Language Proficiency Index (PLI) uses 17 indicators to evaluate the language's influence, which is called weight. The index assesses the usefulness of language for the whole of humanity and not for a single individual affected by any geographical environment, human circumstances or personal preferences. This index also fails to measure the beauty and value of language and its associated culture.

Then we conduct sensitivity analysis on our model, analyze different results and find better parameters for the desired result. In summary, our model is a viable and

reasonable model with technical and data support. Due to its subjectivity, the data can be flexibly applied after training.

3.1.2 Reasons for choosing seventeen indicators

The Language Proficiency Index (PLI) is the measure of linguistic influence. In this model we use five factors to measure PLI: geography, economics, communication, knowledge and media, and diplomacy. But why use these five factors?

First of all, the correlation between the five factors is very low, and each factor can be separated as the influence index of PLI.

(a) Geography: Geospatial distances play its role of spatial isolation and are more likely to be spoken by the same or neighboring countries in the same language. Three indicators represent geography, namely geographical area, number of neighboring countries and number of languages used respectively.

(b) Economy: The influence of economies through their own economic power affects the setting of national language policies and the establishment of language training systems in other language regions. At the same time, the economy affects people's choice of language learning, so as to promote the speed and scope of the spread of a certain language. There are four indicators on behalf of the economy, namely, national GDP, GDP per capita, exports, foreign exchange.

(c) Exchange: The exchange of languages promotes the development of languages, the number of native speakers and the number of second language users, and they are important indicators of linguistic influence. At the same time, with the rapid development of tourism, language exchange and learning are promoted, and the influence of a certain language is also expanded. In other words, the number of native speakers, the number of second language users and the number of people who travel abroad are indicators of the exchange factor.

(d) Knowledge and media: with the rapid development of the Internet, the Internet has made the Earth a global village. Internet users are increasing day by day, which increases the possibility of people in different countries in their communication and exchange, and thereby expands the influence of languages in different countries. At the same time, the level of education is a measure of the people's likelihood of learning different languages. The model uses the global top200 universities rankings and pedagogical investment as an indicator to measure the level of education and the "usefulness" of a language. The diplomacy of a country influences the learning of the second language of the country, in which we judge the languages of each language based on the countries joining the International Monetary Fund, the permanent members of the UN, the countries that have joined the World Bank, and the ten international coalitions "Usefulness."

3.1.3 Solution steps

When we try to obtain the weight of the five aspects of the first-level evaluation

and the weight of 17 second-level evaluation criteria, subjective judgment is ill-considered. So we choose the Analytic Hierarchy Process (AHP) as the way to combine the weighting coefficients of all the indicators in the evaluation system[3].
Obtain the index weights

• Determine the judging matrix. We use the pairwise-comparison method and 1–9 method of AHP to construct the judging matrix A = (a_{ij}):

$$a_{ij} = a_{ik}a_{kj}, \tag{21}$$

where a_{ij} is set according to the 1–9 method.

• Calculate the eigenvalues and eigenvectors. The greatest eigenvalue λ_{max} of matrix A has corresponding eigenvector u = $(u_1, \ldots, u_n)^T$. Then we normalize u by

$$x_i = \frac{u_i}{\sum_j u_j}, \tag{22}$$

• Do a consistency check. The indicator of consistency is

$$CI = \frac{\lambda_{max} - n}{n-1}, \tag{23}$$

where n is the dimension of the matrix.

The expression of consistency ratio is

$$CR = \frac{CI}{RI}, \tag{24}$$

Having confirmed the weighting coefficients of all the indicators in the evaluation system, now we quantify the importance of CW_i denotes the weight of criteria level factor i, where AW_j is the weight of secondary critical level factor j for the ith critical level, mi denotes the total number of secondary critical factors, and F_j denotes the secondary critical level factor.

The evaluation grade Y_1 should be

$$Y_1 = \sum_{i=1}^{5} CW_i \sum_{j=1}^{mi} AW_j F_j, \tag{25}$$

3.1.4 Results and analysis

We obtain the following results:

• Judging matrix:

$$A = \begin{bmatrix} 1 & 1/4 & 1/2 & 1 & 3 \\ 4 & 1 & 3 & 4 & 5 \\ 2 & 1/3 & 1 & 2 & 3 \\ 1 & 1/4 & 1/2 & 1 & 1 \\ 1/3 & 1/5 & 1/3 & 1 & 1 \end{bmatrix}, \tag{26}$$

• Weight vector of criteria level:

CW = {5,9,7,3,3}

• Weight vector of components level:

AW_1= {0.5,0.25,0.25}

AW_2= {0.45,0.1,0.25,0.2}

AW_3= {0.2,0.4,0.3}

AW_4= {0.5,0.25,0.25}

AW_5= {0.2,0.4,0.2,0.2}

For this level, CI =0.0304, CR =1.12, satisfying the criterion for consistency of CI/RI < 0.1.

The rankings of language proficiency in the top ten countries by AHP are shown in the following table:

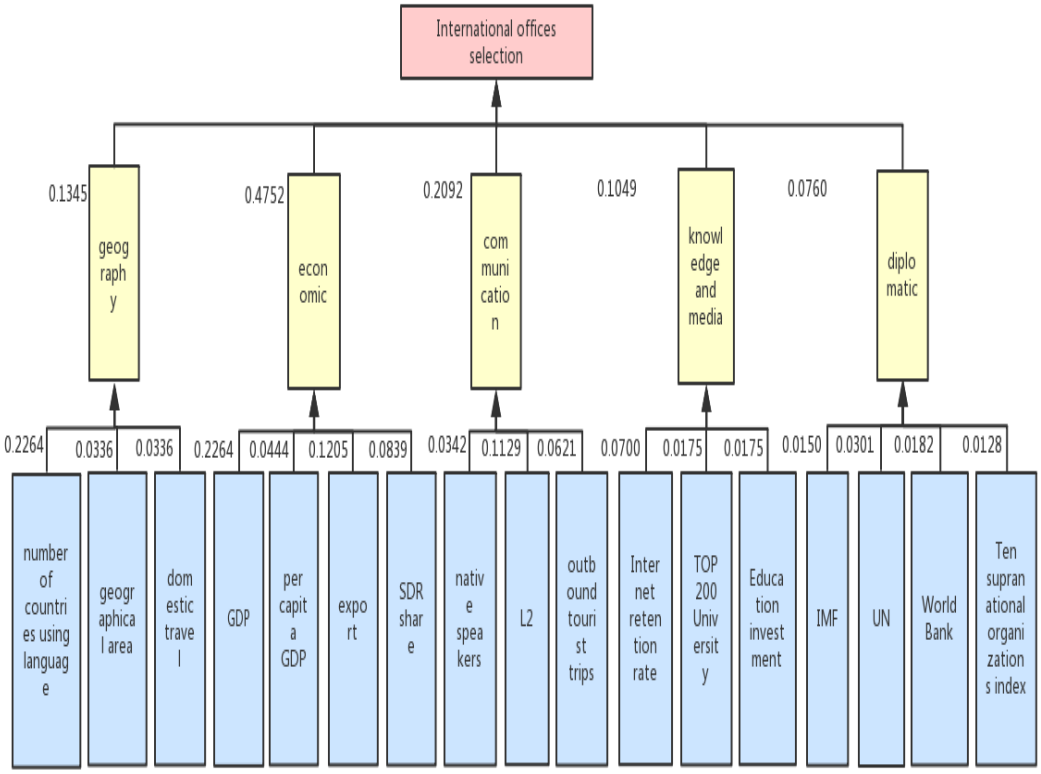

Figure 9 The weight of indexes

Table 4 Language area rankings

Ranking	Language	L1(a hundred million)	Score
1	**English**	**44.6**	**0.889**
2	**Mandarin Chinese**	**96.0**	**0.411**
3	**French**	**8.1**	**0.337**
4	**Spanish**	**47.3**	**0.329**
5	**Arabic**	**29.5**	**0.273**
6	**Russian**	**14.9**	**0.244**
7	**German**	**9.3**	**0.191**

8	**Japanese**	**12.5**	**0.133**
9	**Portuguese**	**21.5**	**0.119**
10	**Hindustani**	**29.8**	**0.117**

3.1.5 Conclusion

From the above analysis, we can draw the following conclusion: the establishment of international offices no longer exists in the US and china due to the existence of offices in these two countries. We only consider the establishment of international office in the language area ranked 3-8 in the table above, and the specific locations will be established as topical in the center of the language area: Paris, Madrid, Jeddah, Moscow, Munich and Tokyo.

3.2 Communication level assessment model

3.2.1 Introduction

Considering the communication level of the countries where hotspots are located in each center separately and comparing with the selected office locations to determine whether there is a need to reduce the number of offices.

3.2.2 Assumption and formula

Symbol Table 4 .

Symbol	Definition
N_{host}	Internet hosts
$\mathbf{N_p}$	Internet users
$\mathbf{L_{internet}}$	Internet level
$\mathbf{N_{Telepthone}}$	Telephone users
M_{max}	Maximum telephone users
$L_{\mathbf{Telepthone}}$	Mobile communication level
$\mathbf{w_{internet}}$	The weight of Internet level
$\mathbf{w_{Telepthone}}$	The weight of mobile communication level
$L_{communication}$	Communication level

We use two evaluation methods, the first used to measure the level of Internet：

$$L_{internet}=\frac{N_{host}}{N_p}, \quad (27)$$

The second one is used to measure the level of mobile communications:

$$M=N_{Telepthone}, \quad (28)$$

$$L_{mobile}=\frac{M}{M_{max}}, \quad (29)$$

Comparing the level of communications in country where the international office is located is measured by the following formula:

$$L_{communication} = w_{internet}L_{internet}+w_{Telepthone}L_{Telepthone}, \quad (30)$$

Where $w_{internet}$=0.8, w_{mobile}=0.2.

Using the above formula, we get the rankings of the top ten languages only considering the level of communications, as shown in the following table:

Table 5 Language area rankings

Ranking	Language	Internet	Mobile	Communication
1	**Arabic**	**0.020**	**0.048**	**0.022**
2	**Spanish**	**0.150**	**0.046**	**0.130**
3	**German**	**0.308**	**0.098**	**0.266**
4	**French**	**0.381**	**0.057**	**0.316**
5	**Russian**	**0.364**	**0.238**	**0.339**
6	**Japanese**	**0.650**	**0.126**	**0.545**

3.2.3 Conclusion and analysis

In the above table, the communication levels of the six selected international offices in Model 5 are arranged. The higher the level is, the more network communication is made. From the perspective of saving the cost of a company, it can be deleted.

4 Conclusions

In order to solve this transnational corporation's problem, the team needs to describe the model that investigates the development trend of the global language and the choice of international offices. We will resolve each sub-issue one by one.

To solve the problem of simulating the distribution of speakers in various languages over time, the team established a model that measures the distribution of a language in different regions and the distribution of a second language in a region based on two factors of geography and economic and trade relations. To solve the problem of predicting the total number of native speakers and speakers over the next 50 years, the team used gray prediction models to obtain results. In order to solve the problem of the change of geographical distribution of language, the team established a model of population migration and obtained the change of geographic distribution

in different languages. MNCs need our design model to establish six international offices. To do this, we choose AHP method to model the language proficiency index of the top ten languages as a factor to determine the influence of language, select the office and take the national communication Based on the level of development, in order to save company costs, minimize the number of international offices.

5 Sensitivity Analysis

In sensitivity analysis, we introduce a disturbance variable Δa_{tk} and a function

$$\overline{a_{tk}} = a_{tk} + \Delta a_{tk}$$

We affect the weights by modifying the values of the disturbance variable as shown in the figure 9. The intersection of the black vertical line and the x axis indicates the weight of the attribute when the actual decision result is obtained. The intersection of the blue vertical line and the x axis is the change point, indicating that the decision result will change after the attribute weight value exceeds this point. If there is no blue Colored vertical bars, indicating that the weights of various alternatives (order) do not change with the weight of this property and are stable. There is no blue line in the figure 9, no matter how the weight of the attribute changes will not affect the decision-making result, so our model is stable.

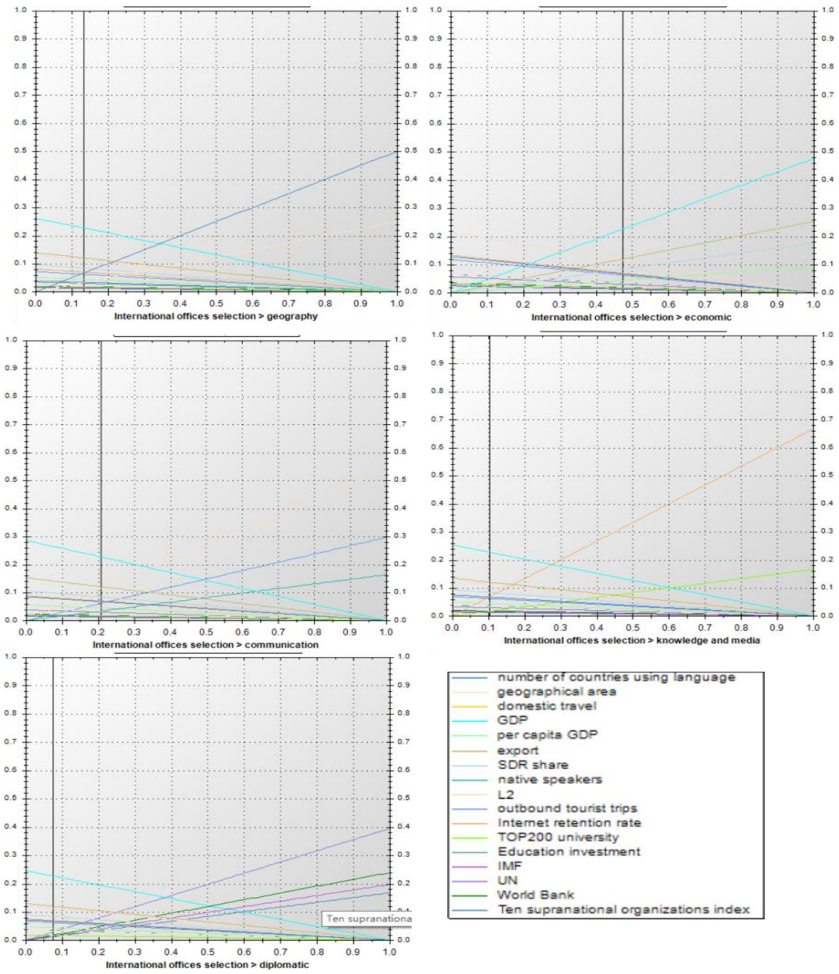

Figure 9 Sensitivity analysis results

6 Strengths and Weaknesses

6.1 Strengths

We propose six models for different issues.
In Model 1, we reasonably quantify the effect of language geographically on the periphery and analyze the case of Europe. Compared with the real data, the results of our model are accurate.
In Model 2, we consider the nine major economies in the proportion of bilateral trade, our model is scientific and rational.
In Model 3, according to the characteristics of the first 14 languages, we use different models to separately calculate the population as the second language because our forecast result has certain reliability.
In Model 4, we only consider the two dominant and important factors, geo-economy, but the invisible factors such as history and culture are not considered.
We directly predict population data in the future world in the UN database and inevitably the errors in the data.
In Model 5, we use AHP to analyze more influencing factors, and our model results are comprehensive.
Gray prediction model is more suitable for such a small sample of world immigration data, less relevant data, relative to the linear programming, neural network and other prediction models.

6.2 Weaknesses

We directly predict population data in the future world in the UN database and inevitably the errors in the data.
In Model 6, we only consider the host and cell phone communication in these two factors, ignoring the use of other less communication methods, the model results have some errors.

7 References

[1]List_of_languages_by_total_number_of_speakers,Wikipedia,https://en.wikipedia.org/wiki/List_of_languages_by_total_number_of_speakers

[2]ethnologue_longitudinal_dataset_and_report_for_mcm,Ethnologue,https://ethnologue.podia.com/ethnologue-longitudinal-dataset-and-report-for-mcm?mc_cid=f5ce474ce7&mc_eid=41e1d6fad5

[3]rank_order_guide,CIA,https://www.cia.gov/library/publications/resources/the-world-

factbook/rankorder/rankorderguide.html

[4]The world's languages, in 7 maps and charts, Washington Post,
https://www.washingtonpost.com/news/worldviews/wp/2015/04/23/the-worlds-languages-in-7-maps-and-charts/?utm_term=.dfe79cd3c032

[5] The countries' import data,OEC,
https://atlas.media.mit.edu/en/visualize/tree_map/hs92/import/deu/show/all/2016

[6] World population prospects 2017,UN, https://esa.un.org/unpd/wpp/Publications/
[7] World Bank Open Data,World Bank, https://data.worldbank.org/
[8] List_of_languages_by_number_of_native_speakers, Wikipedia,
https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers
[9] UNESCO Special Group of Experts on Endangered Languages.
Language vitality and language are endangered. National Language,2006,NO.3

8 Memo

To: Chief Operating Officer of the company
From:88596
Data：February 12，2018
Subject/Re: Advice for opening additional overseas offices

According to the survey of global languages, there are over 6,000 different languages in the world today, but one third of the languages have fewer than 1,000 speakers. In addition, speakers of the top 15 languages are more than half of the global population. In a multilingual and globalized society, language is the prerequisite for us to be able to communicate with others and to allow us to participate in the social, cultural and economic activities. The establishment of additional international offices for the company's long-term development is necessary for companies which want to be truly international companies.

However, where additional international offices should be set up and what language would be spoken? The answer to these two questions is the one I will give advice. Our team, evaluates the global language and finally lists the **top 10 most powerful languages** in the world based on the combination of geography, economics, communication, knowledge and media, and diplomatic skills,. They are **English, Mandarin, French, Spanish, Arabic, Russian, German, Japanese, Portuguese and Hindi** respectively.

English is now the de facto common language throughout the world: it is the dominant language in three G7 countries (United States, United Kingdom, Canada), and the legacy of the British Empire gives English a global reach. Mandarin Chinese is ranked second, but its value is only half. French ranks third with its outstanding position in

international politics. The fourth and fifth respectively are Spanish and Arabic. The top six languages are exactly the six official languages of the United Nations, and they are still in the top six even excluding the diplomatic impact. The four remaining languages of the top ten, two from the BRICS (Portuguese-Brazil, Hindi-India) and two from the official languages (German and Japanese) of two economic powers.

So it is wise to decide the location of international offices based on the global distribution of strong languages. Because the company already has offices in New York and Shanghai, we recommend setting up an office in the language area ranked 3-8. Because some language zones are made up of several countries with native languages, the city in which the office is located should be the central city of the language zone. In German, for example, Germany, Austria and Switzerland are all German-speaking countries. Germany, as one of the most influential countries, the capital of Germany is selected as the hotspot city in the German-speaking area. **So these six offices are located in Paris, Madrid, Jeddah, Moscow, Munich and Tokyo.**

The above six cities except Jeddah not the capital of the host country, the other five are all. The capital is the center of the country's economy, culture and politics, which facilitates the distribution of information. Although Jeddah is not the political center of the host country, it is a diplomatic center. Jeddah adjoins the Red Sea is one of the most economically open areas in the Arab world. Therefore, the selection of these six cities is very much in line with the company's long-standing development here.

In addition, we separately assessed the communications level of countries in which the above six offices are located, from low to high respectively to Saudi Arabia, Spain, Germany, France, Russia and Japan. At present, only Japan's communication level is above average. Forecasts show that communication levels in Germany, France and Russia will be above average over the next 10 years. So in the short term, five other offices, except the one in Japan, will be given priority in the short term to saving client company resources; in the long run, Russia, France and Germany can be gradually phased out.

At the same time, in order to ensure the business level in these areas, it is necessary to establish a sound network office system and set up a host internet in hot spots to directly communicate with the headquarters. This will greatly improve the company's operational efficiency and reduce the company's operating costs.

Above all is the result and suggestion of our team on the issue of opening an additional international office.