据地理模型考虑的分布模型

我们主要考虑以地理分布和经贸关系为主要因素分别构建两个图模型，解决随时间推移模拟各种语言说话者的分布问题。此问题的模型为后续预测说话总人数中的第二语言人数做下准备工作。

假设国家间的距离以首都的距离为标准。

假设各个国家在2017-2025年首都不会变化。

假设各个国家在2017-2025年不会因战争等不可抗拒因素灭亡。

常量

地球半径R

中心热点城市（首都或母语源头国家的首都）经度

中心热点城市（首都或母语源头国家的首都）纬度

I、j两个国家距离地球球心的夹角

I、j两个地区间的距离

两个连线地区间距离的最大值

X地区总人数

X地区使用j语言的比例

变量

时间t

I国家的人口

使用语言i为母语的国家用语言j作为第二语言的比例因子

使用语言i为母语的国家用语言j作为第二语言的本国人数

语言j受地理分布影响作为第二语言的使用人数总和

X地区使用j语言的总人数

公式

=sin()\*sin()\*cos(-)+cos()\*cos()  
=

=

=

= (4)

=+ (5)

= (6)

= (7)

大部分研究可以分为宏观的和微观的，由于宏观方法很难适应于我们的问题，我们采用微观技术来考虑问题。学习第二语言使各个国家打开国门，增加了与外界的沟通，促进经济、政治、文化的发展。而我们的基础模型I以地理分布为重点，把第二语言的学习归为临近国家语言的学习，因此我们建立了一个最短距离的图模型，通过该模型衡量第二语言学习者的总人数。

该模型首先从各个国家的首都间的距离出发，我们通过谷歌地图可以精准定位各个国家首都的经度、纬度，根据两点经纬度与与距离转换公式(1)得到不同国家间的距离:

=sin()\*sin()\*cos(-)+cos()\*cos()  
= (1)

通过最短路径算法找到不同国家首都之间的最小距离，如果需要计算i国家的学习j国语言为第二语言的人数则需定义一个第二语言的传播因子(2)，作为权重。

= (2)

模型以模拟i国人民的不同时期的人口数量作为学习j国语言的基数，通过与传播因子(2)的乘积得到不同时期i国学习j国语言并把它作为第二语言的人数。

= (3)

X地区有N个国家，用上述方法分别算出N个国家使用j国语言作为第二语言的人数，通过累加求和得到X地区把j国语言作为第二语言学习的总人数。

= (4)

(N=1,2,…,n)

以人口数量为依据，我们可以得到X地区不同国家的母语数量和第二语言数量，通过计算j语言的母语人数(本模型中的母语人数以该国人数为依据)和第二语言人数，得到X地区j语言的总使用人数。

=+ (5)

与X地区人口总数

= (6)

相除，得到j语言的比例

= (7)

通过不同语言在该地区的所占比例，进而得到该地区的各种语言说话者的分布。

据经贸关系模型考虑的分布模型

常量

I经济进口总量中j经济体的份额（衡量i对j的语言影响力）

变量

I经济体的人口数

i经济体中使用j经济体语言作为第二语言的人口数

使用经济体中使用j经济体语言作为第二语言的人口数

公式：

=\*

=

经过我们的认真筛选，主要有六大经济体，分别为美国，欧盟，中国，俄罗斯，日本和印度。其中，欧盟中有四大国家的经济实力排名靠前，在本模型中将他们作为独立的经济体，分别为英国，德国，法国以及意大利。也就是说，本模型中共有9大经济体。

标准引力模型

经济发展是国家强大的支柱，而经贸关系是衡量一个国家或地区对于他国经济影响力的重要指标，我们的模型II以经贸关系为重点，综合不同网站上的排名，通过将某个经济体对其他经济体的外贸出口份额以及该经济体地区的人口为衡量指标，改进最大密度子图，建立我们的最强经济影响模型，通过该模型衡量第二语言学习者的总人数。

首先确定n经济体的人口数，对其取对数，为，所得到的数据易消除异方差问题；同时，取对数以后，变量具有弹性。通过与m经济体中占n经济体的外贸出口份额作乘积，得到

=\* (8)

进而通过将该经济体与其他所有经济体一一对比，得到

= (9)

(N=1,2,…,n)

以人口数量为依据，我们可以得到某地区不同经济体的母语数量和第二语言数量，通过计算某语言的母语人数(本模型中的母语人数以该经济体人数为依据)和第二语言人数，得到某语言的总使用人数。

=+ (10)

与该地区人口总数

= (11)

相除，得到该语言的比例

pro= (12)

通过不同语言在该地区的所占比例，进而得到该地区的各种语言说话者的分布。

6个办事处模型

概述

在一个多语言的全球化社会中，语言是让我们能够与他人交流、允许我们参与社会的文化、经济等活动的首要前提。那么哪些语言是最有用的？现在有一家在美国纽约市和中国上海设有办事处的大型跨国服务公司正在不断拓展，在全球不同国家分别建设6个办事处，你会在这6个办事处中使用哪种语言？这将是本模型中我们要解决的问题。

该模型采用层次分析法衡量语言的“有用性”，我们定义了一个语言能力指数，分配权重。广义来说，主要有5种因素影响语言的有用性。分别为：

1.地理

2.经济

3.交流

4.知识与媒体

5.外交

从狭义上说，5种因素分别由不同指标衡量（参见下表）。也就是说，语言能力指数（PLI）采用17个指标评价语言的影响力，即权重。该指数评价的是语言对人类整体而言的有用性，而并非针对受到任何地理环境、人文情况和个人偏好影响的单一个体。这一指数也无法衡量语言和其相关的文化的优美程度和价值。

（此处插入17个指标的表）

进而对我们的模型进行敏感性分析得到，分析不同的结果并为理想的结果找到更好的参数。综上所述，我们的模型是一个具有技术和数据支持的可行和合理的模型。由于其主观性，数据训练后可以灵活运用。

选取十七大指标的原因

语言能力指数（PLI）是衡量语言影响力的标准，在这里，我们用五个因素来衡量PLI：地理，经济，交流，知识与媒体，外交。但为什么使用这五个因素？

首先，五个因素之间的相关性很低，每个因素都可作为PLI的影响指标分离出来。

其次，它们与PLI密切相关：

（a）地理：地理空间距离发挥着它的空间隔离作用，相同国家或相邻国家说同一种语言的可能性更大。有三个指标代表地理，分别为地理面积，临近国家数和语言使用国数。

（b）经济：经济体通过自身经济实力产生的影响力，影响着其他语言区域国家语言政策的设定和语言培养体系的建立，同时，经济影响着人们对语言学习的选择，从而推动某种语言的传播速度和范围。有四个指标代表经济，分别为国家GDP,人均GDP，出口，外汇。

（c）交流：语言的交流促进语言的发展，母语使用人数与第二语言使用人数是衡量语言影响力的重要指标，同时，随着旅游业的迅速发展，促进了语言的交流与学习，也扩大了某种语言的影响力。也就是说，母语使用人数，第二语言使用人数和出境旅游的人次是衡量交流因素的指标。

（d）知识与媒体：随着互联网的迅速发展，网络使地球称为了一个地球村，互联网用户与日俱增，增大了不同国家人民沟通交流的可能性，进而扩大了不同国家语言的影响力。同时，教育水平是衡量人民对于学习不同种类语言可能性的因素，该模型用全球top200大学数与教育学投资数作为指标衡量教育水平进而判断某种语言的“有用性”。

（e）外交：一个国家的外交程度影响着该国第二语言的学习，其中，我们以加入国际货币基金组织的国家，联合国常任理事国国家，加入世界银行的国家和十个国际联盟为参考指标判断各个语言的“有用性”。

解决步骤：

当我们试图获得一级的五个方面的权重评估和十七个二级评估标准的权重，主观判断是不恰当的。 所以我们选择Analytic Hierarchy过程（AHP）作为组合所有权重系数的方式评估系统中的指标。

确定判断矩阵。我们采用成对比较法和层次分析法1-9构造判断矩阵

计算特征值和特征向量。矩阵A的最大特征值λ\_max具有相应的特征向量u

做一致性检查。一致性的指标是

其中n是矩阵的维数。一致性比率的表达是

确定了评价体系中所有指标的权重系数后，我们量化了…的重要性.

CWi表示标准水平因子i的权重，其中AWj是第i个临界水平的次级临界水平因子j的权重，mi表示次级临界因子的总数，Fj表示次级临界水平因子。

至于对...的荣誉和贡献，我们考虑了这么多的奖项和因素，很难确定一个权重向量。 我们通过给这些因素赋予相等的权重来得到一个近似的解决方案。

最后，我们可以得到...使用AHP模型。

结论

•分析标准级别的权重矢量，最高权重用于...性能。

•...在确定...方面起着不太重要的作用，...的重量是最低的。