**关于现阶段国内外推荐系统发展现状的综述**

**1.推荐系统的概念和定义**

**2.推荐系统的分类**

**3.协同过滤的推荐及其改进**

**4.推荐算法面临的新形势及其改进**

**1.推荐系统的概念和定义**

随着互联网的迅速发展以及信息量的暴涨，网络上的信息严重“过载”，用户很难在大量的信息中找到自己真正需要的信息。推荐系统根据用户个人的习惯和偏好向用户推荐其有可能感兴趣的项目（item）是解决信息“过载”问题的主要工具【4】。比如，为用户提供主动推荐的功能越来越多地被应用到各种门户网站和电子商务系统中，这些系统通常会包括用户过去的点击日志、评分和评论等信息，通过这些历史的信息来分析用户的特征并提供推荐服务，给访问者主动推荐最可能感兴趣的信息【5】。

目前被广泛引用的推荐系统的非形式化概念是Resnick和varian在1997年给出的:“它是利用电子商务网站向客户提供商品信息和建议，帮助用户决定应该购买什么产品，模拟销售人员帮助客户完成购买过程”【10】。

典型的个性化推荐系统主要包括用户数据记录模块、偏好分析模块和推荐算法模块。用户数据主要包括关注、购买、浏览和评价等行为，以及相应的上下文信息、用户标签、社交网络数据等内容，通过对这些数据的分析，可以建立合适的偏好模型，以描述用户的需求、喜好等信息。用户的偏好模型表征方法包括显式与隐式两种。显式用户偏好模型包括基于本体的偏好模型、偏序偏好模型、向量空间模型和基于正反例的模型；隐式用户偏好模型间接描述用户的偏好，例如，隐语义模型、基于邻域的算法。推荐算法包括基于内容的推荐、协同过滤的推荐等【26】。

**2.推荐系统的分类**

目前，对推荐系统的分类并没有统一的标准，很多学者从不同角度对推荐方法进行了不同的划分。但主流的推荐方法基本包括以下几种：基于内容推荐、协同过滤推荐、基于链路分析的推荐、基于知识推荐、组合推荐【10】。

基于知识的推荐：在某种程度上可以看成是一种推理技术，它不是建立在用户需要和偏好基础上推荐的，而是利用针对特定领域制定规则来进行基于规则和实例的推理【10】。

组合推荐：没有一种推荐算法是完美的，协同过滤存在新产品问题，内容过滤对交易频繁的产品推荐质量不够好，等等。那么很容易就设想将这些推荐算法组合起来，取长补短，这也就是组合推荐的思想。理论上讲可以有很多种的推荐组合方法，但目前研究和应用最多的组合推荐是基于内容的推荐和协同过滤推荐的组合【24】。

协同过滤作为目前个性化推荐系统中广泛使用的、最成功的推荐算法，完全依赖于用户评分 来构建用户—项目评分矩阵，使用统计技术寻找与目标用户有相同或相似兴趣偏好的邻居用户，再根据邻居用户对商品项的评分来预测目标用户对其未评分项的评分值，进而选择预测评分最高的前N项商品作为推荐集反馈给用户，其基本思想是用户会对邻居用户所喜欢的商品产生兴趣，即基于用户的协同过滤【14】。

**3.协同过滤的推荐及其改进**

协同过滤推荐算法可以分为基于用户的协同过滤（UCF）与基于项目的协同过滤（ICF）。这两个算法的共同点在于二者都是根据用户—项目评分矩阵建立推荐系统模型，从而为用户提供个性化推荐服务的。不同之处在于UCF是根据用户之间的相似性找到目标用户的最近邻集，然后根据该集合中用户的评分情况确定目标用户的推荐结果。而ICF则是通过分析项目之间的相似性，最终将与目标用户评价较好的、项目相似度较高的作为推荐列表的结果。【http://www.xzbu.com/8/view-6084451.htm】

基于用户的协同过滤：其基本思想是，目标用户对未评分项的评分可以通过相似用户对这些项目的评分进行预测。首先度量用户之间的相似性，相似性高的被作为目标用户的最近邻，然后根据最近邻对项目的评分预测目标用户对项目的评分。度量用户间相似性的传统方法，目前主要有修正的余弦相似性和皮尔森相关系数法【29】。

基于项目的协同过滤：其基本思想是，用户对项目的评分可以通过项目之间的相似性以及用户已有的评分进行预测。在大多数个性化推荐系统中，相对于用户的更新，项目信息的更新较慢，项目之间的关系相对稳定。所以可以先离线计算项目之间的相似性，对相似性进行评分预测。评分矩阵稀疏对项目间的相似性影响比对用户之间相似性的影响小得多【29】。

数据稀疏：用户评分数据收集越多，协同过滤算法的推荐质量越高。但是，由于电子商务网站用户及商品项的数量庞大且不断增加，使得用户—项目评分矩阵成为高维矩阵；同时用户给予评分的商品项很少，通常在1%以下，导致用户—项目评分矩阵中的数据极端稀疏【14】。

冷启动问题：冷启动问题是稀疏性问题的极端情况，也称为第一评价人问题或早期评价人问题，具体包括新用户问题和新项目问题。即当新用户（新项目）进入推荐系统后，由于还未提供（接受）任何项目（用户）的评分，导致系统无法向新用户推荐其可能喜欢的项目（或将新项目推荐给可能喜欢它的用户）。在协同过滤推荐系统刚投入运行时，每个用户在每个项目上都将面临冷启动问题【14】。

**3.1协同过滤的改进方法——混合推荐**

没有一种推荐算法是完美的，每种算法都是既有优点，也有不足，于是就想到将这些推荐算法组合起来，取长补短，这就是组合推荐的思想。比如，基于内容的推荐和协同过滤的推荐相结合的混合推荐。一个最重要原则就是通过组合后能避免或弥补各自推荐技术的弱点。研究人员提出的组合思路大致分为后融合（融合两种或两种以上的推荐方法各自产生的推荐结果）、中融合（以一种推荐方法为框架，融合另一种推荐方法）、前融合（直接融合各种推荐方法）【10】。

崔春生等【24】在研究组合算法的基础上，提出一种基于泛函网络实现前融合推荐算法。通过推荐系统泛函拓扑结构的建立，利用泛函神经元的自学习、自组织和自适应能力，建立了内容推荐算法和协同过滤推荐算法的泛函拓扑结构，进一步优化推荐结果，极大地提高了系统的推荐准确度。他探讨的是推荐算法的“前融合”问题。

黄创光等【5】针对预测场景中的不确定性，提出了不确定近邻的协同过滤推荐算法UNCF。首先，因为传统的协同过滤使用kNN算法，当所在同一群体内的个数不足k个时，误差就会较大，因此通过定义相似度阈值，使推荐集的对象必须满足相似度大于阈值；然后，引入不确定近邻因子的概念。通过引入φ这个调和参数，用来协调基于用户以及基于产品的两者之间不同的影响；最后，不确定近邻的协同过滤推荐算法。结合用户的最近邻集和产品的最近邻集，对用户在产品上的评分进行预测，定义推荐公式。他用到的混合方法就属于“后融合”。

王桐【31】等注意到推荐系统一般忽略了用户兴趣随时间变化的特性，认为只有在同一时间段内爱好相似的用户才应该被称为相似用户，引入了时间因子，提出一种基于柯西分布量子粒子群的混合推荐算法。该算法首先构建基于时间因子的混合推荐模型，再利用柯西分布量子粒子群算法搜索模型中的最优参数组合。其中混合推荐模型通过把用户和项目的属性信息添加到协同过滤推荐算法中，并引入能够代表用户兴趣迁移特性的时间因子构建而成。

**3.2协同过滤的改进方法——基于模型的推荐**

黄震华等【39】详细研究了基于排序学习的推荐算法，它与传统的推荐算法不同，传统的推荐算法不需要训练，直接计算用户之间以及项目之间的相似度，产生推荐结果；基于排序学习的推荐算法是一种监督性学习，通过训练数据集得到排序模型，然后对测试数据集使用该排序模型产生最终的推荐结果。

候翠琴等【51】提出一种基于项目多类属概率潜在语义的协同过滤算法，首先将概率潜在语义分析法中的隐变量集固定为项目的多类属集；然后迭代学习隐变量的分布，即用户兴趣模型；最后用学到的兴趣模型度量用户的相似度，对目标用户做出推荐。

杨娟等【18】借鉴了基于模型的稀疏线性（SLIM）推荐算法，提出了融合用户集合关系的稀疏线性（UCSLIM）算法，首先将用户划分为用户集合，然后计算出用户之间的相关系数以及用户与用户集合的相关系数。在预测用户i对物品j的评分时，会同时将用户i和其他用户的相关性以及用户i和其他用户集合的相关性都考虑在内。

**3.3协同过滤的改进方法——矩阵分解**

近年来，矩阵分解算法（MF）已逐渐应用于推荐系统，传统的矩阵分解算法有：奇异值分解（SVD）、非负矩阵分解（NMF）、概率矩阵分解（PMF）等。这些算法的共同点是，通过将一个高维的矩阵分解为两个或多个低维矩阵的乘积实现维度规约，方便于在一个低维空间研究高维数据的性质【32】。对m行n列的用户打分矩阵R，m为用户数，n为产品数，rij为矩阵中第i行第j列的元素。由于用户评分矩阵十分稀疏，R中存在大量的缺失数据。基本的矩阵分解算法把打分矩阵R分解为m行k列的用户潜在属性因子矩阵U和n行k列的产品潜在属性因子矩阵V，并使UVT尽可能接近R【11】。

概率矩阵分解（PMF）算法是应用较多的矩阵分解算法。刘胜宗等【16】提出基于PMF进行潜在特征因子分解的标签推荐（TagRec-UPMF）。社会标签系统采用社会标签整合各类资源，这有助于用户组织、浏览、搜索自己感兴趣的资源，也能够更好地帮助用户之间进行沟通及共享，而标签推荐系统可将用户感兴趣的标签推荐给使用同一资源的用户。TagRec-UPMF将用户、标签、资源三方面的潜在特征因子进行联合分解，并将求得的特征向量两两之间的内积进行线性加权并产生推荐。

涂丹丹等【32】同样也使用了联合概率矩阵的算法，但分解时不是结合标签，而是关注广告，结合与广告点击相关的多方面信息进行广告推荐：用户浏览网页信息，用户点击广告信息，网页与广告内容相似性和广告投放在目标网页上的点击率。本文采用正则化的UPMF方法，同时最优化分解3个矩阵，求解用户、广告、网页内容的隐含特征向量，并以此进行广告点击率的预测。

在传统的矩阵分解算法中，一般使用潜在属性因子，赵长伟等【11】提出了混合因子矩阵分解（HMF）算法。显示属性能够表示用户或产品之间已知属性的相关性，隐式属性表示不能明确表示的推荐因子。与基本的矩阵分解算法相比，HMF算法使用了显式属性之间的相关性，对稀疏数据矩阵分解给出了合理约束，提高了推荐精度。同样，由于使用了显式属性因子，HMF算法能够推荐新产品并对推荐结果给出一定的解释。

**3.4协同过滤的改进方法——聚类**

聚类，是在没有人工标注的基础下，将具有相似属性的数据聚集在一起的无监督学习方法。它使具备一定相似性的数据实例组织成一些相似组，处于同组内的数据彼此相似，处于不同组的数据彼此不同【44】。基于聚类的协同推荐算法（一般采用K-means算法聚类）首先取定聚类簇数，利用聚类算法对用户——项目评分数据进行聚类；在得到聚类结果之后，需要找到离目标用户最近的簇（也可以是最近的若干个簇），将整个簇中用户作为目标用户的最近邻（或者计算目标用户与最近的一个或者若干个簇中用户的相似度，选取相似度最高的若干个用户作为目标用户的最近邻集合）；最后，加权平均最近邻集合中用户对目标项目的评分，从而可以预测目标用户对所有未知项目的评分【13】。

袁汉宁等【13】针对传统协同推荐不能有效考虑项目内容特征，引入了项目的内容特征，从包的角度重新表示一个用户的评分状况， 以每个示例表示用户对某个项目的评分，构建了基于多示例（MI）的用户评分信息表达模型。根据多示例学习模式具有一定容错性的特点，设计了基于多示例聚类的协同推荐算法，通过多示例聚类计算用户的最近邻集合，根据最近邻集合对用户评分进行预测。

严隽薇等【20】针对目前个性化服务中用户模型稳定性低、推荐结果不尽人意的现状，在建立基于本体的用户兴趣模型（UIM）基础上，通过模型更新提高稳定性，建立用户群实现用户模型管理，其中UIM以兴趣树的形式呈现。在传统协同过滤算法基础上，提出利用矩阵聚类降维分解技术的个性化推荐算法，并引入偏好方差的概念计算用户最近邻。避免了传统协同过滤算法的数据稀疏性缺陷，提高了推荐质量。

郁雪等【15】通过结合主成分分析（PCA）和SOM聚类，提出一种基于PCA-SOM的混合协同过滤模型。首先利用PCA的维数约简技术对高维数据进行降维处理，使原始数据转换到主成分空间上，缓解了数据稀疏性的特点，并且为后面的聚类减少了计算复杂度；随后在低维的向量空间上进行SOM聚类，得到用户的偏好模式{C1,C2…Cj}，由于上述算法的时间复杂度较高，因此可以作为系统的离线模型，改善系统的伸缩性。

韦素云等【37】提出一种基于联合聚类平滑的协同过滤推荐算法。首先对原始矩阵中的评分模式进行用户和项目两个维度的联合聚类；然后采用联合聚类平滑的方法预测用户对未评分项目的评分值，分别从用户聚类簇、项目聚类簇和联合聚类簇多方面对评分矩阵空缺项进行平滑填充；最后结合基于项目的协同过滤算法查找项目最近邻并进行推荐。

吴泓辰等【44】提出了一种基于协同过滤与划分聚类的改进推荐算法。建立了协同过滤算法的推荐矩阵，并给出评价数值和更新系数，使推荐算法能综合所有用户的评价，为用户提供满意有效的推荐；还给出划分聚类算法，使用户分组更加准确，进一步提高了推荐算法的准确性和实时性。

吴毅涛等【45】针对离散评分不能合理表达用户观点和过滤算法存在稀疏性问题，借鉴年龄模糊模型，提出了梯形模糊评分模型。该模型将离散评分模糊化为梯形模糊数，考虑了评分模糊性和信息量，通过梯形模糊数来计算用户相似度，据此设计了协同过滤算法，并证明了该算法是传统协同过滤算法在模糊域的扩展。

**3.5协同过滤的改进方法——标签**

标签是一种对网络资源进行标记、分享及识别的工具，标签信息相对稳定。通过将商品抽象为标签，建立用户对标签的偏爱或兴趣分布，一方面可以了解用户的喜好特征，另一方面可以减弱原始数据集稀疏程度对推荐算法的影响【21】。

蔡国永等【21】提出了一种基于标签和因子分析的协同推荐方法。首先把用户在物品上历史行为的统计，转化为对用户在物品标签上的统计，从而缓解数据稀疏的情况；为了降低标签维度过高导致的计算复杂性问题，提出了采用因子分析法，抽取出潜在重要且稳定的特征因子向量来最终表示用户的历史行为，并据此度量用户行为在特征因子向量上的相似性。最后采用协同过滤的思想给出了一种新的协同推荐方法。

张晓敏等【27】提出了基于概念层次树的用户模型。在改进算法中，将用户对商品的评分转化为对每个商品种类的评分，建立起基于概念层次树的新用户模型；利用新的用户模型计算用户在不同商品种类上的相似性，寻找该商品种类上的最近邻居；综合各商品种类上的邻居推荐产生top-N推荐。

传统的协同过滤推荐算法都是基于单一指标评分，张付志等【19】认为要综合考虑项目的多个指标评分才能知道用户的具体偏好，引入多指标评分的概念，提出一种基于Widrow-Hoff神经网络的多指标推荐算法。提出一种用户偏好特征向量的计算方法，以挖掘更精确的用户偏好并并给出一种基于用户偏好特征向量的用户相似度度量方法。在此基础上，提出一种基于Widrow-Hoff神经网络的多指标推荐算法。

考虑到标签可能存在语义歧义，陈毅波等【36】认为许多研究人员可能使用不同的标签来描述同一个对象，但表达同样的意义，导致推荐的效果不理想。本文首先根据标签共生分布计算标签相似度，然后根据相似度分组同义标签，最后通过组标签计算用户的相似度，将其混合到经典的协同过滤算法中。

现有的用户建模方法只能分别计算单个标签之间的匹配程度，忽略了多个标签作为一个整体对用户兴趣产生的影响。为此，蔡毅等【47】提出了一种基于标签组合效应的推荐算法（TGER）。该算法利用用户对资源的评分筛选出对用户兴趣有重要影响的标签组合，通过高维标签组合优先匹配的方法计算用户与资源之间的相关度。

**3.6协同过滤的改进方法——考虑情境**

该方法引入了目前电子商务推荐领域的最新工具——情境。情境信息是指人的行为或事件发生影响的上下文信息或者场景信息，如时间、地点等。情境信息在很大程度上影响用户的最终选择。在一个系统中，情境包括多个因素，个性化情境是指最影响目标用户做出选择决策的因素。例如，当用户看电影的时候，不同的时间、地点、和不同的人看等情境都会影响其选择【29】。

张佳琳等【22】针对个性化推荐系统中普遍存在着信息共享程度低、资源复用不足等问题，提出基于多场景融合的分布式推荐模型。该模型采用分布式的双向刻画的方法，通过多场景融合算法，进行客户特征需求与服务场景的互生成，并最终生成推荐列表。

高旻等【29】提出了一种结合用户个性化情境和基于项目的协同过滤推荐方法。该方法首先计算各情境下的项目相异度矩阵，根据相异度矩阵和基于项目的协同过滤算法对每个用户进行预推荐。由于该方法先判断对用户推荐影响最大的情境因素，更符合人类在不同情境下有不同需求的事实。

环境感知推荐模型通过引入用户所处环境信息向用户提供适合当前环境的推荐结果，现有的一些研究采取环境信息硬分解或者硬聚类的方法，使得环境信息只能属于其中的一个聚类，并且聚类之间无相关性。为解决这一问题，张付志等【40】采用模糊C均值聚类方法，使环境信息同时属于多个聚类，从而解决环境信息硬分解和硬聚类带来的问题。

**3.7协同过滤的改进方法——信任度**

张维玉等【41】提出了一种基于协同矩阵分解的信息评分与信任预测联合模型。该模型在将评分矩阵与信任关系矩阵进行协同分解时，既能保证被分解的两个矩阵分解过程共享用户潜在变量，又能兼顾两个矩阵分解过程中能够各自获得反映本领域知识相关性的表达。使用分解得到的多个相关低维潜在变量矩阵乘积即可做出评分与信任两个问题的预测。

印桂生等【42】提出一种基于受限信任关系和概率分解矩阵的推荐方法，由不信任关系约束信任关系的传播，得到准确且覆盖全面的用户信任关系矩阵，并通过对用户信任关系矩阵和用户商品矩阵的概率分解联合用户信任关系和用户商品矩阵信息，为用户给出推荐。

刘贵松等【46】针对社交网络中协同过滤推荐算法的推荐速度计算问题，提出了一种基于最近邻方法的改进计算方法，并对算法有效性进行了分析。该算法对用户的相似性度量采用基于最短路径的信任关系，用分层图和动态规划的方法进行计算，并在社交网络的应用中对关系链的深度进行限制。

肖志宇等【9】针对稀疏性和冷启动问题提出了基于信任网络和随机游走策略的评分预测模型ReferentialUserWalker。该模型通过随机游走策略，利用信任网络中的信任朋友对目标物品或与目标物品相似的物品的评分进行评分预测，并在信任网络中找到最可信的Top-N评分参考用户，同时引入信任度权重，降低了噪声数据的影响。

针对传统协同过滤推荐算法存在推荐精度低和抗攻击能力差的缺陷，贾冬艳等【35】提出了一种基于双重邻居选取策略的协同过滤推荐算法。首先基于用户相似度计算的结果，动态选取目标用户的兴趣相似用户集。然后提出了一种用户信任计算模型，根据用户的评分信息，计算得到目标用户对兴趣相似用户的信任度，并以此作为选取可信邻居用户的依据。最后，利用双重邻居选取策略，完成对目标用户的推荐。

现有的基于信任的推荐模型中，交互请求者都会选择当前信任度最高的对象进行交互，可信度高的对象会因为超负荷而造成响应延时，引起信任度的下降。为此，郭晶晶等【50】提出了一个考虑社区节点信任度稳定性与社区利益的推荐算法。社区控制中心根据目前各成员节点的状态与请求者的特征做出满足请求者的约束条件，同时保证社区节点信任度稳定性，并使得社区所得利益最大的决策。

**3.8协同过滤的改进方法——其他**

赵琴琴等【4】提出一种基于内存的传播式协同过滤推荐算法SPCF，在每位用户找到了自己的邻居之后，可以利用相似度传播寻找每位用户更多的邻居。第一，邻居的生成。我们计算用户u和其他用户v的相似度，考虑迭代h次时的相似度，引入衰减系数C，以模拟实际情况下相似度随着迭代次数增加而衰减；第二，推荐的生成。从用户和项目两方面考虑目标用户对目标项目的评分，用给定的公式来计算目标用户a对目标项目t的预测评分值。推荐给目标用户预测评分最高的一个项目或N个项目作为结果。

王玙等【34】为了高效准确地预测朋友关系，本文首先提出一个基于关系的社交圈检测算法，快速聚类自我网络，完成社交圈的自动分类。其次充分考虑公共邻居之间的关系，给出基于社交圈的用户相似性定义，并利用该相似性设计朋友推荐算法。

冷亚军等【38】提出一种有效的针对稀疏评分的最近邻选择方法——两阶段最近邻选择算法（TPNS），分为两个步骤。首先计算用户间的近邻倾向性，选择近邻倾向性较高的用户组成初始近邻集合；然后根据初始近邻集合计算目标用户与其他用户间的等价关系相似性，使用等价关系相似性对目标用户的初始近邻集合进行修正，得到最近邻集合。

李聪等【33】提出了基于领域最近邻的协同过滤推荐算法，以用户评分项集作为用户相似性计算基础，并将并集中的非目标用户区分为无推荐能力和有推荐能力两种类型；对于前一类用户不再计算用户相似性以改善推荐实时性，对于后一类用户则提出“领域最近邻”方法并对并集中的未评分项进行评分预测，从而降低数据稀疏性和提高最近邻寻找准确性。

李聪等【14】提出了一种基于n序访问解析逻辑的冷启动消除方法。考虑某位用户访问Web页面，由于冷启动，传统的协同过滤算法无法为其提供推荐服务。我们转而对用户访问路径进行研究，即我们不仅考虑用户访问了哪些Web页面，也要考虑用户访问Web页面的顺序。算法步骤是：首先通过Web日志来获取用户访问项序，进而定义了n序访问解析逻辑将其分解为用户访问子序集；在此基础上设计了用户访问项序的相似性计算方法来搜寻新用户的最近邻集合；进而提出了改进最频繁项提取算法IMIEA来生成面向新用户的top-N推荐。

李宏涛等【28】本文采用概念格从数据中挖掘知识，利用用户特征属性和社交网络图建立概念格，提出了弹性随机游走方法SRWR，并在此基础上用概念格知识指导随机游走，提出了概念融合概念格和随机游走的FCASRWR方法，度量了用户之间的相似性，算法最终根据相似度进行朋友推荐。

张磊等【12】在项目评价相对充足的情况下，在基本协同过滤的基础上，探讨新的有效推荐的研究思路。提出了一种新的TMNN-CFRA，采用两层面多个并行人工神经网络（ANN）完成项目推荐的全过程。两层面的多个BP网络协同工作，高层面BP网反向误差传播，直至低层面多ANN进行网络权值修正，以此为基础借助用户评价等特征前向给出项目推荐。

**4.推荐算法面临的新形势及其改进**

**4.1面向服务的推荐**

云计算是基于互联网的计算，它依赖于对互联网上的计算资源进行共享来处理应用程序。云计算中的资源都是以服务的形式通过互联网来提供给用户共享使用的。服务是云计算中不可或缺的重要元素之一。Web服务作为云计算中使用的重要的服务类型之一，可以被看作是一个小型的软件系统。该系统支持计算机与计算机之间通过网络进行协同交互【23】。由于互联网上用户的需求呈现出多样化、个性化的特点，为了给用户提供高体验质量（QoE）的服务，需要以用户为中心进行服务的有效组织与推荐，实现按需服务【3】。大多数功能相对单一的服务是无法满足用户需求的，必须将共享的服务组合起来才能达到目的。因此，如何将多个服务组合起来形成功能更为强大的服务，就成为目前服务计算领域重要的研究方向和研究热点【25】。典型的服务网络包括以亚马逊为代表的商品交易平台、以大众点评网为代表的生活服务平台、以programmableWeb为代表的应用程序接口管理平台等【26】。

刘建晓等【3】对用户共性需求进行建模，从涉及的角色、目标、流程、服务几个方面进行考虑，设计了RGPS制导的按需服务组织框架，该框架由领域建模、按需服务组织、服务聚类、服务推荐等几个方面构成。设计了三种服务查找和推荐方法，分别是根据用户需求目标查找服务、根据用户角色查找服务、根据服务描述信息查找服务，来满足用户提出的不同需求表达形式，更好地指导用户利用服务间的交互与协作来查找合适的服务以满足其需求。

姜波等【23】为解决推荐领域存在的不足，提出一个基于二部图的服务推荐方法（BIGSIR）。不同于其他的基于历史信息进行推荐的Web服务推荐方法，BIGSIR将历史信息抽象成一个二部图，并充分利用二部图中的节点和边的信息为用户进行高效的推荐。同时，特别考虑了Web服务的兼容性特性。

潘伟丰等【25】提出了一种基于服务网络的服务组合推荐方法，为实现不具有WSDL文档的服务的组合问题提供了一种解决方案。该方法：基于服务组合历史，构建服务组合模型，抽象服务间的协作和竞争关系；利用复杂网络方法挖掘服务使用模式；提出了基于服务使用场景的服务组合推荐算法。

辛乐等【26】提出了基于服务信誉的偏好分析与推荐模型，它是可兼容数字评分、标签评价、自然语言评价的多维异构服务质量指标模型，实现自动、定量地提取用户对各质量指标维度的偏好权重，对备选服务进行效用预测。在服务推荐系统中应用多属性决策方法分析用户的偏好信息，可以有效避免协同过滤中用户打分稀疏性，且考虑了服务的多维信誉指标，从而增强了推荐的可信度。

刘轶等【49】考虑到当前的服务推荐策略大多缺乏对系统整体发展需求的关注，结合服务新增、失效等动态要素构建了服务生态系统仿真模型；从全局发展的视角出发，提出了服务系统综合发展指标体系；设计了能够保障系统中服务多样性、减小服务失效率的推荐策略，该策略将服务分为新服务、保护服务和普通服务并分别设计推荐机制。

互联网上经常会出现大量的冗余服务资源，使得用户常常受限于大量冗余的资源列表，而无法从中找出合适的资源。马建威等【30】提出一种基于混合推荐算法和隐马尔科夫模型的信息推荐方法。首先提出一种混合协同过滤方法，利用随机游走策略进行云环境下的信息推荐；然后，基于受欢迎度和信息熵的方法提出新用户判定的方法；最后，基于隐马尔科夫模型提出服务冗余的判定方法，并进行实验验证。

**4.2大数据环境下的推荐**

大数据管理及处理能力已经成为引领网络时代IT发展的关键，获取大量运行数据并建立对其进行动态高效处理的能力，已经成为产业竞争力的体现。然而，如何对大数据进行有效的管理分析仍面临着巨大挑战【7】。与传统的推荐系统不同，大数据环境下的推荐系统具有以下特征：需要处理的数据量更大、采集数据以用户隐式反馈数据为主、数据更新速度更快、对推荐项目的准确性要求更高、对推荐项目的实时性要求更高【8】。

马帅等【7】在《大数据管理技术专刊前言》中总结：第一，大数据分析处理系统和平台及其关键技术是大数据管理技术的研究重点之一；第二，分布式与云环境中大数据分发与任务调度是对大数据管理技术的有效支撑；第三，网络大数据尤其是社会网络数据的分析与应用得到了广泛关注。

孟祥武等【8】在《大数据环境下的推荐系统》中提出了大数据环境下推荐系统整体框架，由源数据采集层、数据预处理层、推荐生成层、效用评价层四层组成。数据预处理层的主要工作为用户偏好获取、社会化网络构建、上下文用户偏好获取。推荐生成层是推荐系统的核心，目前主要的推荐技术有基于矩阵分解的推荐系统、基于隐式反馈的推荐系统、基于社会化推荐系统、组推荐系统。

将传统的推荐技术直接应用到云计算环境时会面临推荐精度低、推荐时延长以及网络开销大等问题。针对上述问题，朱夏等【48】提出一种云计算环境下基于协同过滤的个性化推荐机制RAC。首先通过定义候选邻居（CN）的概念筛选对推荐结果影响较大的项目集，快速准确地定位候选邻居；在此基础上提出基于候选邻居的协同过滤推荐算法（CN-DCFA），在候选邻居中搜索目标用户已评分项目的k近邻，预测目标用户的推荐集top-N。

刁赢龙等【6】在《大规模配电网负荷数据在线清洗与修复方法研究》中指出，配电网负荷数据的数据规模已达到“大数据”级别，大规模配电网负荷数据在线清洗与修复方法是：首先分析配电网负荷数据的来源、典型异常数据的原因和特征，然后提出基于密度的负荷数据流异常辨识方法和基于协同过滤推荐算法的负荷数据修复方法。所提的算法和框架能够有效预处理配电网负荷数据，具有实际应用价值。

**4.3LBSN与POI推荐**

近年来，卫星通信、 GPS设备、无线传感器网络等移动互联网技术不断进步，基于位置的社交网络（Location-based Social Networks, LBSN）服务得到迅速发展，且受到广大用户的喜爱，如国外的Foursquare、Gowalla，国内的街旁、嘀咕等。LBSN与传统在线社交网络（Online Social Networks, OSN）的主要区别在于其增加了地理位置信息，用户可以对当前访问的兴趣点（餐厅、电影院、旅游景点等）签到，并与好友分享自己的签到信息【2】。为用户推荐用户可能感兴趣的地理位置的任务称为兴趣点（point-of-interest, POI）推荐。兴趣点推荐一方面满足了用户的个性化需求，减轻用户面临的信息过载问题，另一方面，帮助LBSN服务提供商实现智能化的位置服务，如位置感知的广告服务等，从而增加LBSN服务提供商的营业收入。因此，兴趣点推荐在基于位置的社交网络中扮演越来越重要的角色【17】。

曹玖新等【2】把LBSN看作异构网络，它包括用户节点、兴趣点节点，用四元组G<U,L,E,W>表示。源点与终点之间的路径可以用元路径来表示，与元路径的节点类型、边类型一致的真实路径为实例路径。基于元路径的兴趣点推荐算法为：首先，确定始于用户节点类型、终于兴趣点节点类型且路径长度不大于3的元路径，把它添加到元路径集中；然后计算元路径特征值，元路径P的特征值为所有实例路径体现的关联程度之和，实例路径p体现的节点关联度cor(p)基于随机游走思想计算；最后计算用户u去兴趣点l签到的概率，选择top-N推荐给用户。

李鑫等【1】针对用户冷启动，提出一种基于相同兴趣圈的推荐系统，兴趣圈是由所有访问某一类别商品的用户群及他们之间的社会关系构成的社交联系。主要思想是：先根据不同的兴趣圈把用户评分矩阵Rc划分，然后对分块后的Rc，结合矩阵分解给出目标函数；对于冷启动情况，利用朋友关系设计正则化项具有一定的合理性；人们有时会收到一个领袖人物（专家）的影响，容易接受他们的观点并对用户自身行为产生影响。所以我们在朋友关系的基础上提出了基于专家关系的规则化项。至此，我们结合以上因素，得出修改后的目标函数及其约束项，从而解决冷启动问题。

余永红等【17】提出了基于Ranking 的泊松矩阵分解兴趣点推荐算法。首先，为了更加准确地捕获用户对兴趣点的偏好，使用泊松分布来建模用户的签到行为；其次，为解决兴趣点推荐中的隐式反馈问题，使用BPR标准来优化泊松矩阵分解的损失函数；最后，为了进一步改进推荐算法的性能，利用包含地域影响力的正则化因子约束泊松矩阵分解过程。

刘树栋等【43】总结了基于位置的社会化网络推荐系统研究与应用的特点和难点：位置轨迹挖掘及用户位置活动特征的提取；用户的位置移动轨迹特征与社会关系的关联性；用户位置隐私保护和安全问题；推荐结果的解释与呈现；可扩展性和主动性。