

分类号 TU-9
密 级 公开

单位代码 10618
学 号 2130120013



重庆交通大学

硕士学位论文

房地产开发企业大数据应用能力成熟度评价研究

研究生姓名：蒋 鸿 鹏

导师姓名及职称：孟卫军 副教授

申请学位类别 工学 学位授予单位 重庆交通大学

一级学科名称 管理科学与工程 论文提交日期 2016 年 4 月 15 日

二级学科名称 管理科学与工程 论文答辩日期 2016 年 5 月 26 日

2016 年 5 月 7 日

Research on the Evaluation of Big Data Application Ability of Real Estate Based on the Maturity Model

A Dissertation Submitted for the Degree of Master

Candidate : Jiang Hongpeng

Supervisor : Prof. Meng Weijun

Chongqing Jiaotong University, Chongqing, China

重庆交通大学学位论文原创性声明


本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。


学位论文作者签名：

日期：2016年6月6日


重庆交通大学学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权重庆交通大学可以将本学位论文的全部内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。同时授权中国科学技术信息研究所将本人学位论文收录到《中国学位论文全文数据库》，并进行信息服务（包括但不限于汇编、复制、发行、信息网络传播等），同时本人保留在其他媒体发表论文的权利。

学位论文作者签名：
日期：2016年6月6日

指导教师签名：
日期：2016年6月6日

本人同意将本学位论文提交至中国学术期刊（光盘版）电子杂志社 CNKI 系列数据库中全文发布，并按《中国优秀博硕士学位论文全文数据库出版章程》规定享受相关权益。

学位论文作者签名：
日期：2016年6月6日

指导教师签名：
日期：2016年6月6日

摘 要

房地产企业目前面临去库存的问题和供给侧改革的要求，风险和压力不断加大，但与此同时，科技进步日新月异，大规模生产、分享和应用数据的现代信息技术深刻改变着人类的思维、生产、生活和学习方式，深刻展示了世界发展的前景。通过大数据，可能有全新的视角来发现新的商业机会和重构新的商业模式，使不太了解房地产的人可以在大数据时代应用新的载体和新的方式颠覆原有的商业模式。这对于房地产开发企业的升级换代和可持续发展有重要意义。因此，应用大数据的能力对当下的房地产企业非常关键和必要。本文将针对房企应用大数据的能力建立评价指标和方法，解决应具备什么能力，达到什么程度，从什么途径进行提升的问题。

对房地产开发企业大数据应用能力理论以及成熟度模型进行综述，重点分析了现有大数据应用能力方面的不足。

能力成熟度模型的建立：1) 成熟度模型构建分析；2) 能力成熟度模型对于房地产开发企业大数据应用能力适用性分析；3) 对成熟度模型维度进行分析，将房地产开发企业大数据应用能力进行等级分析，将其能力等级分为对过去（现状）的把握、发现模式、预测和优化四个等级；5) 建立最终的房地产开发企业大数据应用能力成熟度模型。

指标体系的建立：1) 成熟度特性分析；2) 在总结分析大量文献和前人研究成果的基础上初步构建了房地产开发企业大数据应用能力成熟度评价指标体系；3) 基于系统工程学理论进行指标体系优化；建立最终的指标体系。以此为基础，结合 4 个能力等级以及评价准则，运用灰色理论评价法量化大数据应用能力建立起评价模型，并与 4 个应用能力等级、11 个评价指标形成完整的成熟度模型。

以某房地产开发企业大数据应用能力作为案例，成熟度等级为发现模式级，表明该企业的大数据应用能力成熟度水平还不够高，与实际相符，据此佐证该模型的有效性。

关键词：大数据应用，能力成熟度模型，系统工程学理论，灰色理论评价法

ABSTRACT

Real estate enterprises are now facing the problem of inventory and supply side of the requirements of the reform, the risk and pressure to increase, but at the same time, the rapid progress in science and technology, large-scale production, sharing and application data of modern information technology of profound changes in the human thinking, production, lift, learning stely, shows the profound world development prospects . Through big data, there may be a new perspective to find new business opportunities and reconstruct a new business model, so do not know real estate people can be in the era of big data application new carriers and new way to subvert the original business model. This is of great signification for the upgrading and sustainable development of real estate development enterprises. Therefor, the ability to apply big data on the real estate business in the movement is very critical and necessary. This paper will be based on the ability of housing prices to establish evaluation indicators and methods to solve the problem should have what ability to achieve what level, from what approach to enhance the problem.

Firstly, Carries on the summary to the real estate development enterprises data application ability theory and maturity model, focusing on analysis of the deficiency of existing data application ability.

Secondly, Capability maturity model were estabilished: 1) maturity model analysis; 2) capability maturity model for real estate development enterprises data application ability for analysis; 3) the maturity model dimension analysis, the estabilishment of a real estate development enterprise data application ability maturity model is a two-dimensional structure; 4) to the real estate development enterprise data application ability of analytical grade, will be its ability to grade is divided into (status quo) in the past, grasp the pattern discovery, prediction and optimization of four grades; 5) establish the final real estate development enterprise data application ability maturity model.

Thirdly, The establishment of the index system: 1) maturity characteristic analysis; 2) after conclusion of the analysis of a large number of documents and experts interviews based on initial construction of the real estate development enterprise data application ability maturity evaluation index system; 3) based on the theory of system engineering were index system optimization; the final index system is established. This is based, combined with four capability levels and evaluation criteria, evaluation model

is established by using Grey theory evaluation criteria, evaluation model is established by using Grey theory evaluation method quantitative data application ability, and from a complete maturity model and four application ability level, 11 evaluation index.

Fifthly, On the case of a real estate development enterprise data application ability, maturity level in order to find out the schema level, indicating that ability in applications of the enterprise's data maturity level is not high enough, and in line with the actual situation. According to the evidence of the effectiveness of the model.

KEY WORDS: Big data application, Capability maturity model, System engineering theory, Grey theory evaluation method.

目 录

第一章 绪论	1
1.1 研究背景与意义	1
1.1.1 大数据时代中国房地产业面临的挑战与机遇	1
1.1.2 研究意义	2
1.1.3 研究目的	3
1.2 国内外研究现状	3
1.2.1 大数据应用能力研究	3
1.2.2 能力成熟度模型评价	6
1.2.3 相关研究现状评述	7
1.3 研究内容、方法与技术路线图	8
1.3.1 研究的主要内容	8
1.3.2 研究的方法	8
1.3.3 技术路线图	9
第二章 房地产开发企业大数据应用成熟度理论概述.....	11
2.1 大数据的内涵与应用	11
2.1.1 大数据的概念与特征	11
2.1.2 大数据应用	13
2.1.3 房地产开发企业的大数据应用	19
2.2 能力成熟度模型	21
2.2.1 成熟度的概念	21
2.2.2 能力成熟度模型理论	21
2.2.3 几种成熟度模型的介绍	24
2.2.4 能力成熟度模型的应用	28
2.3 房地产开发企业大数据应用能力成熟度模型	29
2.3.1 CMM 构建的必要性	29
2.3.2 CMM 构建的意义.....	30
2.3.3 CMM 构建的难度及工作的特点	30
2.4 本章小结	31
第三章 房地产开发企业大数据应用能力成熟度模型的分析与构建..	33
3.1 成熟度模型的构建分析	33
3.2 能力成熟度模型(CMM) 对于房地产开发企业大数据应用能力的适用性分	

析	34
3.2.1 CMM 的特性	34
3.2.2 房地产开发企业大数据应用能力运用 CMM 的适用性	34
3.3 房地产开发企业大数据应用能力成熟度模型的构建	35
3.3.1 房地产开发企业大数据应用能力成熟度模型的范畴	35
3.3.2 房地产开发企业大数据应用能力成熟度模型的维度	36
3.3.3 房地产开发企业大数据应用能力成熟度模型的等级	37
3.3.4 房地产开发企业大数据应用能力成熟度模型的内部结构	40
3.4 本章小结	42
第四章 房地产开发企业大数据应用能力成熟度评价指标体系构建 ..	43
4.1 CMM 指标体系设置要求	43
4.2 CMM 指标体系的识别	43
4.3 房地产开发企业大数据应用能力成熟度评价指标体系的提取	45
4.4 房地产开发企业大数据应用能力成熟度评价指标体系的确定	50
4.5 本章小节	56
第五章 房地产开发企业大数据应用能力的评价方法	59
5.1 几种常用评价方法介绍	59
5.2 应用灰色综合评价法进行评价的适用性分析	61
5.3 基于灰色综合评价法的评价步骤	62
5.3.1 权重计算及权重集的确定	63
5.3.2 灰色评价的过程	64
5.4 本章小节	68
第六章 案例分析	71
6.1 某房地产开发企业的现状	71
6.2 房地产开发企业大数据应用能力成熟度模型的应用	72
6.2.1 评价指标集	73
6.2.2 权重计算	73
6.2.3 测评结果计算	76
6.3 改善房地产开发企业大数据应用能力的相应措施	82
6.4 本章小结	84
第七章 结论与展望	85
7.1 结论	85
7.2 不足与展望	85
7.2.1 本文研究的不足	85

7.2.2 未来研究的展望	86
致 谢.....	87
参考文献.....	89
附录 A.....	93
附录 B.....	97
附录 C.....	99
攻读学位期间取得的研究成果.....	101

第一章 绪论

1.1 研究背景与意义

1.1.1 大数据时代中国房地产业面临的挑战与机遇

伴随着“云计算”、“物联网”的浪潮，人类进入了“大数据时代”，并且已经波及到重多的领域。电子计算机和互联网技术的迅猛发展成为数据信息量几何级增长的引擎燃料，百度每天需要处理几十拍字节的数据量；线上购物平台淘宝每天产生的数据也达到了 20 太字节；youtube 上平均每秒钟就有一段时长超过一小时的视频发布；“脸书”上每天上传的照片大概有 1000 万张，点赞或者评论也有几十亿次。百度公司创始人李彦宏在为国家主席习近平讲解信息技术领域的前沿课题中提到，人类已经进入了大数据时代，在过去的两年中全球产生的数据达到人类总数据量的 90%^[1]。

在 2011 年麦肯锡全球研究院发布的《大数据：创新、竞争和生产力的下一个新领域》报告中，首次提出大数据的概念。紧接着又过了一年，越来越多的人也开始关注和了解“大数据”了，“现实世界中大数据的使用”来自于 IBM 和牛津大学萨伊德商学院的一项研究，其结果也表明，越来越多的企业也开始认为它们的竞争优势与大数据存在着密不可分的关系，大数据也不断被应用到各个行业与领域中^[2]。牛津大学著名科学家维克托·迈尔-舍恩伯格认为预测是大数据的核心，分析复杂繁多的数据并对其进行取舍，从而达到创造积极和安全未来的目的，于此同时，国际的顶级期刊 nature 和 science 也都对大数据进行了专刊，阐述了大数据在应用时会产生的技术困难以及其蕴含的潜在价值。我国“十二五”规划中也重点强调了与大数据相关的关键技术创新工程。人们对大数据的认同率也反映出其逐渐成为关注的热点甚至焦点话题，从 2010 年的 34%增长到 2012 年的 63%。2013 年也被定义为数据元年，数据指数级增长、分享和应用的时代就此拉开序幕。

当今时代，数据资源的战略价值显得尤为重要，从国家到组织、企业到个人各个层面都有广泛的应用。通过大数据挖掘出有价值的信息，辅助进行决策，提高了经济效益，所以大数据也被应用于各行各业之中。比如在房地产行业，目前正面临着行业的巨变；开发—建设—销售的模式正在转变为服务模式。为了在竞争白热化的房地产市场中求得生存，必须应用大数据，提高房企从产品设计开发生产、降低风险、节约成本、营销、增值服务等各个方面的大数据应用能力，而不是继续沿用传统开发销售模式。对于房地产企业来说，拥有“给客户画像”的能力，就可以掌握前端的产品定位；与此同时拥有大量目标客源，就可以掌握后

端的销售变现。通过应用大数据做深度的信息搜集与分析，对客户的需求做到洞若观火，有利于房地产企业实现精确营销，优化产品设计，提高服务品质，改善业务流程。

因此，大数据应用对于房地产开发企业来说不再是可有可无的附属产品了，转变成为了房企的重要资源，是房企商业能力的衡量标准，同时也是未来发展的方向。房地产企业应具备什么样的大数据应用能力，该如何改善显得非常重要，有针对性地对房地产企业的大数据应用能力进行评估并对其薄弱的地方提出相应的改善措施来提高房地产企业的盈利能力是十分必要的，这同时也为构建社区生态、城镇生态发展奠定了基础。

1.1.2 研究意义

通过对房地产开发企业大数据应用能力评价指标体系和成熟度的研究，将其应用于房地产开发企业的成熟度模型中，有助于房地产开发企业了解自身企业大数据应用的成熟度等级，找到不足之处，从而改善其薄弱环节，达到最终提升企业大数据应用的能力。在理论和实践中都有比较重要的意义。

理论意义

国外大数据应用发展顺利，涉及不同领域不同行业的企业都已经引入了大数据应用，但在国内，这方面的研究相对还比较少，且大部分研究集中在大数据的概念和特征上，很多企业还是采用传统的技术手段和工作模式，没有真正将大数据应用作为一个必备因素贯穿项目的全过程，忽略了大数据应用对于企业发展的重要性，缺乏大数据应用的战略思考。针对这些不足，论文从大数据应用能力的重要性出发，构建房地产开发企业大数据应用能力指标体系并确定指标权重，最后得出企业成熟度等级。理论上丰富了成熟度模型和成熟度理论，对房地产开发企业大数据应用的不断发展也起到了一定的推动作用。

实践意义

国外非常注重大数据应用，美国投入巨资启动“大数据发展计划”，把大数据应用上升到国家战略高度，不仅加速了大数据的普及率，同时帮助企业取得了较好的成绩。在我国，应用大数据的企业比重还比较少，房地产开发企业主要以万科、世茂、绿城等为主。构建房地产开发企业大数据应用能力成熟度评价模型，使房地产开发企业能正确认识自身的实力以及不足，对其薄弱的地方提出相应的改善措施来提高房地产开发企业的盈利能力，以达到增强企业竞争力的目的，同时也为企业的发展奠定了基础。

1.1.3 研究目的

随着房地产行业的迅猛发展，出现了房地产泡沫，一二线城市主要表现在价格泡沫上，三四线城市则体现在规模泡沫上。大数据的出现让房地产业看到了曙光，研究房地产开发企业大数据应用的目的主要体现在以下两个方面：

为房地产企业大数据应用能力调查出重要标准及最佳实践。如何确定房地产企业大数据应用能力的影响因素指标，如何构建科学合理的评价指标体系，为大数据应用提供合理的依据。

构建适用于房地产企业的大数据应用能力成熟度模型。如何从众多大数据应用指标中选择适用于房地产企业的指标，并构建能力成熟度模型。

论文将分别对上述两个关键问题进行初步研究讨论，旨在为房地产开发企业大数据应用能力提供一套科学合理的评价指标体系，有针对性地对房地产开发企业的大数据应用能力进行评估，并对其薄弱环节提出相应的改善措施来提高房地产开发企业的盈利能力，同时也为房地产的良性发展奠定了基础。

1.2 国内外研究现状

1.2.1 大数据应用能力研究

国外对大数据的研究相对较早，也比较系统。相比于某些企业在实践中已经开始应用而言，学术方面的研究较为滞后。未曾发现在房地产开发企业大数据应用能力成熟度评价的引入。

其主要研究如下表所示：

表 1.1 国外大数据主要研究成果

序号	研究名称	发表单位	研究时间
1	特殊在线收集：数据处理 (Special online college: Dealing with data ^[3])	科学杂志	2011.02
2	大数据：创新、竞争和生产力的下一个新领域 (The next frontier for innovation , competition , and productivity ^[4])	麦肯锡公司 报告	2012.10
3	自然：大数据 (Nature:Big data ^[5])	自然杂志	2012.10

4	大数据的机遇与挑战 (Chanllge and Opportunities with Big Data ^[6])	美国数据管理白皮书	2012 . 10
5	大数据 ,大影响 :国际发展的新可能(Big data for impact New possibilities with Big Data ^[7])	世界经济论坛	2012.10
6	发展大数据：挑战与机遇(Big data for development:chanllenge&opportunities ^[8])	UN Global pulse	2012.10
7	大数据研究和发展倡议 (Big data research and development initiative ^[9])	奥巴马政府计划	2012.10
8	大数据计算：对商业，科学和社会产生革命性变革 (Big data computing: Creating revolutionary in Commerce,science,and society ^[10])	计算社区联盟	2012.10
9	现实世界中大数据的使用 (The use of big data in the real world ^[2])	IBM 和牛津大学萨伊德商学院	2012

这些相关研究取得了一定的成果，其中全球著名麦肯锡咨询公司的报告，对大数据的影响、核心技术和应用领域进行了全面的分析，代表了大数据前沿的研究方向和成果。

国外研究，可以归纳为以下几个方面：

大数据的概念

介绍大数据的概念是国外研究大数据的主要成果，即什么是大数据，怎样理解大数据，介绍其与传统数据库的差别，从表面上看，从数据库演进到大数据只是一个简单的技术，但其实大数据的出现，颠覆了传统意义上的数据管理方式，对数据来源、数据处理方式以及数据思维模式都带来了革命性的变化。

大数据时代的挑战与机遇

大数据的发展离不开云计算技术，云计算和大数据是紧密联系在一起，如何存储，保存大数据并运用云计算进行提炼、分析，都是应用大数据所面临的挑战，云计算的发展水平直接关系着大数据的应用能力，是应用大数据不可或缺的工具。同时，大数据背后也蕴含着无限的商机，大数据是资源，是一种特殊的资源，它不像其它资源一样会消耗枯竭，它会越用越多，越用越有价值。

大数据的生产与应用

大数据所包含的海量数据信息超过了历史上所有时期数据信息的总和。数据信息的产生也已经不再受到时间、地点的限制。电子计算机与互联网的应用为大

数据的出现提供了富饶的土壤，大数据也会越来越多的在现实世界中被广泛使用，越来越多的企业为了提高自身的竞争力领跑于行业，也开始加大在大数据方面的应用。

随着国外大数据理论的深化研究和广泛应用，国内也逐渐重视，在国内基本是从大数据的特征出发，通过阐述和归纳特征给出其定义。例如，比较有代表性的4V定义，即规模性（volume）、多样性（variety）、高速性（velocity）和价值性（value）共同构成了大数据的四个特点。维基百科对大数据的定义是：利用常规的软件工具去捕捉、管理、处理数据需要花费不可容忍的时间集合。最具代表性的是中国工程院院士李国杰和中科院计算所副总工程师在中国科学院院刊上发表的大数据的研究现状与科学思考^[11]。系统的介绍了何谓大数据，大数据在国际已经引发高度的关注和我国对于研究大数据的重大战略意义。孟小峰、慈祥^[12]等对大数据处理的基本框架进行了阐述，并对云计算和大数据时代数据管理之间的关系进行了分析。在房地产开发企业大数据应用的研究相对重要的有：杜丹阳、李爱华^[13]从房地产开发企业的角度阐述了大数据的应用情况，系统的分析了近年来我国房地产企业应用大数据的案例，结合国外经典案例分析了大数据在房地产企业开发和营销方面的积极作用。张玉娟^[14]等分析研究了我国房地产业在大数据时代面临的新形势，大数据给房地产企业在支持产业决策、增加盈利空间、缓解当前困境等方面都带来了新的机遇，但同时建立数据库、辨别数据噪音和应对激烈的市场竞争方面也带来了新的挑战。严娟^[15]通过大数据信息的挖掘，对潜在的目标客户进行精准的细分，从而实现精确营销，实施资源投放效益的评估闭环，实现营销模式创新。李鹏、吕欣^[16]提出应用大数据技术辅助政府房地产宏观调控，为制定正确可行的政策提供依据。陈大川、张宝山^[17]阐述了如何建设大数据库，分析住房监测、监管住房公积金、监管住房保障等。

这些研究取得了一定的成果，从现有的文献上看，主要有以下几个方面：

对大数据的概念进行解释，以及在大数据时代的机遇与挑战。如孟小峰^[12]的《大数据管理：概念、技术与挑战》，张玉娟^[14]的《大数据时代：房地产业的机遇与挑战》，王珊^[18]的《构架大数据：挑战、现状与展望》，胡雄伟^[19]的《大数据研究与应用综述》，李芬、朱志祥^[20]的《大数据发展现状及面临的问题》，刘智伟，赵静^[21]的《大数据及其应用技术探讨》等。

房地产企业在大数据时代营销模式的转变。如严娟^[15]的《基于大数据的房地产企业精确营销研究》，陈志^[22]的《战略性新兴产业发展中的商业模式创新研究》，程明书、彭晓愈^[23]的《未来房地产网络营销在线交易模式》，肖剑^[24]的《房地产营销新热点—体验式营销》等。

房地产企业应用大数据的研究。如杜丹华、李爱华^[13]的《大数据在我国房

地产企业中的应用研究》，国家统计局浙江调查总队课题组^[25]的《大数据在浙江房价统计中的应用研究》，陈大川、张宝山^[17]的《大数据技术在住房信息统计中的应用》等。

1.2.2 能力成熟度模型评价

能力成熟度模型（Capability Maturity Model）起源于计算机软件行业，具体描述某个软件的概念、量化、运行、测试以及升级软件等的所有过程。CMM 的目的是结合软件某阶段的综合运行情况，并发现其升级发展的关键因素，最后通过对这些识别的因素进行分析，采取相应的措施自我纠偏改进后提高软件的综合能力。常见的 CMM 模型主要有初始级、可重复级、已定义级、可管理级和优化级五个等级。

1987 年卡内基梅隆大学研究所(SEI)在公开发布 CMM 模型后，Bill Curtis^[26]和 William E Hefley 等人将其引入到人力管理体系中，并建立了人力资源方面的能力成熟度模型(People Capability Maturity Model)^[27]，这个模型在一定程度上为人力资源部门管理提供了参考，通过对所有等级中 KAP 的关键实践运作，从而有针对性的提升企业的人力资源管理能力。随着能力成熟度模型的不断发展和演化，其应用领域也变得越来越广阔，从最开始的软件工程领域，到如今更多的行业和领域都可以用能力成熟度来评价。例如市场营销、产业信息和建筑工程项目领域，如表 1.2 所示。

表 1.2 能力成熟度的应用领域

模型名称	适用范围
能力成熟度模型（CMM）	大型系统集成组织
项目管理成熟度模型（OPM3）	项目管理
知识项目管理成熟度模型（K-PMMM）	知识项目的过程改进
软件能力成熟度模型（SW-CMM）	软件过程改进
服务集成成熟度模型（SIMM）	服务工程管理
安全成熟度模型（SMM）	系统安全工程管理
人力资源能力成熟度模型（PCMM）	改进人力资源实践

国内也有一些学者对能力成熟度模型有一定的研究，吕潇潇^[28]结合了我国房地产企业项目管理的特点，构建了一个适合我国房地产项目管理成熟度模型与评价体系，通过对管理能力成熟度的分析，来提升企业的管理能力。冉秀阶^[29]以房地产企业项目管理为研究对象，一方面对国外的研究成果和应用实践进行消化吸收，另一方面与我国现有的房地产企业项目管理特点和现状联系起来，运用层次分析法、模糊综合评价法、规范研究和实证研究等多种方法构建了房地产企业项

目管理成熟度模型，并进行评价研究。

CMM 凭借有助于提高团队各方面管理能力而广泛运用于多个行业领域，针对不同行业领域自身的特点又改良演化出具有独特特征和运用范围的管理能力成熟度模型，在初始阶段，团队的管理杂乱无序、没有原则，但随着团队管理成熟度等级的提高直至最终达到优化级，团队的管理能力都会增强，当团队的管理能力成熟度处于优化级时，团队管理就会进入一个自身不断优化的良性循环。

1.2.3 相关研究现状评述

纵观上述文献，国内外学者在大数据应用方面达成了一些共识，但侧重点也有所不同，国外研究多以报告为主，其主要研究成果涉及大数据的概念、大数据时代所面临的机遇和挑战以及大数据的生产与应用等方面。其成果对于国内的相关研究有很好的借鉴价值。在企业层面有一定的大数据应用研究，但房地产开发企业的大数据应用却是凤毛麟角。这和国外研究报告多从宏观层面把握理解大数据时代有关。国内研究多以国外研究报告定义的大数据概念为主，总结出大数据的特征，开始探索在各行各业的大数据应用，如大数据在我国房地产企业中的应用研究等，对房地产开发企业大数据应用能力成熟度的研究则基本没有，也没有涉及灰色理论评价法的引入。这和市场经济更注重盈利有关，很多企业（包括房地产开发企业）在大数据应用方面已经取得了一定的成果。和已经开始使用实践应用的房地产开发企业相比较，我国的学术理论研究则相对滞后。

在能力成熟度的研究领域里，大多数成熟度模型应用于软件工程、安全工程以及项目管理，目前尚未发现将其应用于房地产开发企业大数据应用能力评价指标体系的构建中。成熟度模型是一个持续改进不断提升的模型。而在房地产开发企业大数据应用能力管理的过程中，主要是通过对大数据应用能力的评估从而识别房地产开发企业大数据应用存在的问题，继而采用切实可行的解决方法，以达到提高房地产开发企业盈利能力，实现可持续发展的目的，所以，在构建房地产开发企业大数据应用能力评价模型时可以融入 CMM 的思想，能够帮助企业决策者了解提高企业竞争力的方法并预先制定出企业在运营过程中可能出现问题的改善措施，以达到可持续化发展的目的。

因此，本文试着提出基于能力成熟度模型（CMM）的房地产开发企业大数据应用能力评价体系，通过对大数据应用能力成熟度等级的定量评价，将有助于房地产开发企业了解提高大数据应用能力的方法，并为企业在未来的投资运营过程中可能出现的问题提出相应的解决方法，这样即提高了企业的竞争力，同时也丰富了能力成熟度的理论体系。

1.3 研究内容、方法与技术路线图

1.3.1 研究的主要内容

本文是站在从事与房地产开发相关工作者的角度，针对房地产开发企业以及大数据应用的现状，分析目前我国房地产开发企业在大数据应用领域内的研究成果，并在能力成熟度模型（Capability Maturity Model, CMM）的基础上，提出了房地产开发企业大数据应用能力成熟度评价模型，确定应用能力的若干层次关系，在对大数据应用能力的关键影响因素进行了探讨后构建出合理的成熟度评价体系，从而帮助房地产开发企业明确分析、利用大数据的因素层次，针对评价指标因素提出改进措施，从而达到提高房地产开发企业的大数据应用能力。本文的主要研究内容有以下几点：

相关理论研究述评。包括房地产开发企业内涵、大数据应用以及能力成熟度模型进行理论综述，分析房地产开发企业与大数据应用的关系特点，并指出现有的评价模型在房地产开发企业应用中的不足之处。

在阅读参考现有文献，借鉴已有的研究成果、以及专家问卷调查的基础上，建立了初步的房地产开发企业大数据应用能力评价指标，运用系统工程理论的相关理论对指标体系进行了优化，得到最后的指标体系。

构建房地产开发企业大数据应用能力的评价模型。运用 AHP 对指标体系进行赋值，主要介绍 AHP 的运算步骤。

根据建立的成熟度评价模型，以某房地产开发企业为例进行实例分析。发放问卷得到最终的评价数据，运用灰色理论评价得出成熟度等级，针对其薄弱环节提出相应的改善措施，以达到提高房企大数据运用的目的。

研究总结。指出本文的研究不足以及后续的研究方向。

1.3.2 研究的方法

根据研究内容，本文主要采用文献调查法、访谈法、问卷调查、层次分析法、统计分析法以及灰色理论评价法作为主要的研究方法。

文献调查法：全面系统的调查国内外有关大数据以及房地产开发企业大数据应用能力研究的相关文献，了解该研究领域的现状和发展趋势。

访谈法：通过专家访谈以及从事与房地产开发相关工作者的观点，构建出房地产开发企业对大数据应用的能力成熟度评价指标体系。

层次分析法：运用专家打分和层次分析法确定评价指标体系各标准指标的权重。

问卷调查法：根据构建好的成熟度标准指标体系，设计问卷调查。问卷调查主要分两步进行。第一步对设计的调查问卷进行进一步的完善与修改，增强问卷的可读性、全面性以及准确度，同时使被调查者能够便于正确理解调查的目的与意义；第二步则搜集尽可能多的样本为成熟度评价提供数据。

灰色理论评价法：这是本文运用的核心方法，对调查结果进行处理以得到最后的评价结果。

1.3.3 技术路线图

依据主要研究内容，制定论文的研究思路如图 1-1 所示：

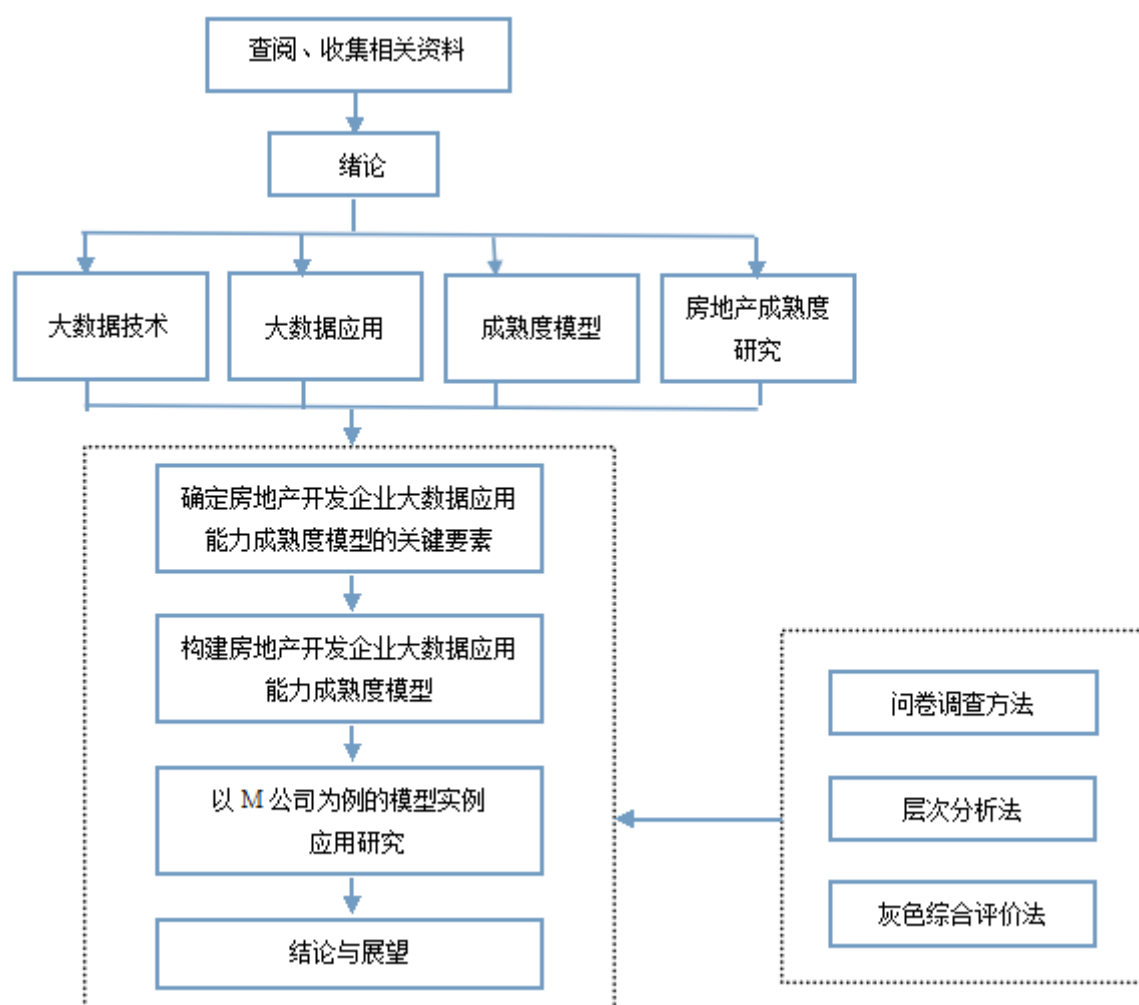


图1-1研究思路图

第二章 房地产开发企业大数据应用成熟度理论概述

2.1 大数据的内涵与应用

2.1.1 大数据的概念与特征

大数据的概念

大数据单从其字面意思来理解代表的是数据在规模上的庞大，但其本身是一个抽象化的概念。大数据不仅仅体现在数据规模上，若仅这样理解就不能分辨出大数据与“海量数据”(massive data)、“超大规模数据”(very large data)等概念的区别。截止到目前，还没有一个真正确定和统一的定义。

1) 大数据在维基百科中被定义为^[30]：使用常用软件工具对在一定区间的数据进行捕获、管理以及处理所消耗的时间超过可容忍的时间。这个定义也不够确切，首先哪些软件工具属于常用软件工具的范畴没有确定，同时对可容忍时间也只是一个概略上的描述，没有定量化，比如可以规定超出常用软件工具处理常规化数据时间几倍以上即为不可容忍时间。

2) IDC 则给出了大数据量化标准的定义^[31]：包含两种及其以上的数据形式，且数据规模在 100TB 以上的高速、实时数据流；或者是开始时数据量非常小，但年增长率在 60% 以上的数据叫作大数据。这个定义中尽管量化标准比较明确，但也仅仅是强调了数据的规模，种类以及增长迅速等数据自身的特征。

3) 研究机构 Gartner 给大数据的定义则是描述性的^[32]：大数据要发挥其对决策的指导、洞察发现以及集成优化海量、多样性和持续增长的信息资产的能力必须要在新的处理模式下。这个定义是在描述性的基础上加入了一些处理此类数据的特征，并用此来描述大数据。

大数据的特征

从以上对大数据定义的概念可以看出，体现了大数据本身具有数据量大、多样性、高速性和价值性等特征，和常规数据比较起来，大数据既包含了一般数据的特征，同时又有其不同的特点，为了更好的理解大数据，首先人们要明确大数据的特征，大数据的基本特征可以归纳为 4 个 V。具体特征如下：

1) 数据量大 (Volume)。全球的数据信息量呈现出爆炸性的指数级增长，数据资源是目前社会增长和需求最快的资源之一。IDC (国际数据公司) 的监测统计表明^[33]，即使在 2009 年遭遇全球金融危机的时候，数据信息量同比增长也达到了惊人的 62%，超过了 80 万 PB (1PB 等于 10 亿 GB)，到 2011 年末，全世界所有的数据信息量已经超过了 1.8ZB (1ZB 等于 1 万亿 GB)，平均会以两年一个周期

而成倍数增长，预计到 2020 年，全球的信息数据量将会突破 40ZB，十年时间增长了将近 20 倍，届时，全球人均数据资源也将到达 5247GB。

2) 多样性 (Variety)。数据在急速增长的同时，数据信息也表现出多样化的趋势。不仅仅体现在各个领域都呈现出海量的增长，数据的类型以及传播方式也更加充实，包括结构化的数据、半结构化的数据以及非结构化的数据类型，这就导致传统的数据处理方式已经很难满足数据多样化的需求，提炼处理数据的手段也必须丰富和完善。

3) 高速性 (Velocity)。大数据的高速性主要体现在数据的分析和处理速度上，数据的处理速率伴随着高性能计算应用市场的发展和深入也在不断提高，增强大数据处理能力，提高数据处理速度同时也是满足数据快速增长的必要条件，只有大力发展云计算技术能力，才能为大数据的应用提供发展的土壤，让大数据的发展和应用有更广阔的空间。

4) 价值性 (Value)。大数据是资源毋庸置疑，同大油田、大煤矿一样，是孕育大财富的摇篮，同时它还有一个与其他资源所不具备的独特优势，它是可以再生的，是越挖越多、越挖越值钱的，如同一座富矿一样蕴藏着无限商机，具有巨大的潜在价值。但同时数据的密度低又是其一个显著的特征，正是由于数据量大，数据信息增长快，数据之间彼此的关联性又比较少，数据的价值往往会比较分散，难以挖掘，传统数据挖掘很难提炼出真正有价值的信息，人工智能将是提炼低密度数据的发展方向。

谷歌的首席经济学家 Hal Varian 曾说过^[34]，数据是丰富且能够被人们所使用的，而目前人们缺乏的就是从越来越广泛的数据中提炼出有价值的，且将其转化为知识的能力。搜集数据、研究大数据的目的就是要提炼出满足于人们的不同需要，并将其运用到各个不同的领域中去，不同领域的大数据应用特点也会有所不同，下表列举了应用在不同领域中具有典型代表性的大数据特征：

表 2.1 典型大数据应用的比较

应用	实例	用户规模	相应时间	数据容量	可靠度	精准度
科学计算	生物信息学	小	慢	TB	中等	非常高
金融	高频交易	大	非常快	GB	非常高	非常高
社会网络	脸书	非常大	快	PB	高	高
移动数据	移动电话	非常大	快	TB	高	高
物联网	传感器网络	大	快	TB	高	高
微博数据	新闻网站	非常大	快	PB	高	高
多媒体	视频	非常大	快	PB	高	中等

2.1.2 大数据应用

大数据应用是指利用大数据资源收集、挖掘、提炼出对人们生产生活有益的海量信息，以帮助人们提高生产效率、产生增值效益。随着大数据的概念逐渐被人们所理解和接受，大数据应用也会慢慢普及到各行各业，这也将是未来企业发展的趋势。

大数据处理框架与运作流程

大数据有多种应用类型，目前较为常见和主流的处理模式是流处理^[35]（stream processing）和批处理(batch processing)^[36]。流处理则是直接进行处理，即 straight-through processing，而批处理的运行方式是先进行存储，然后再进行处理，即 store-then-process。

1) 流处理。流处理的理念是数据本身的价值和时间会呈现出反比的关系，即随着时间的不断流逝，数据的价值将逐渐变小，因此，也正是因为这种理念，所有流处理模式的目标都是要尽可能快速对最新的数据进行分析并得出分析结果，以获取最大的价值。流数据处理大数据主要应用在网页点击、浏览评论的实时统计，网络传感器和金融市场的高频交易中。

可以将流数据处理模式中的数据形象的比作流，而实时更新，不断涌现的数据就成为了动态的数据流，当最新的数据出现时就同步对其进行处理，分析并输出最终的结果。图 2.1 是模拟流数据处理的模型图。

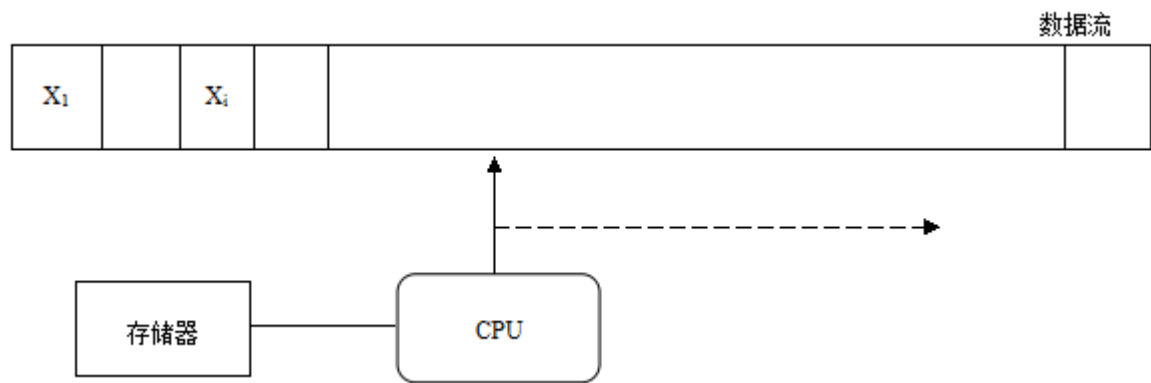


图 2.1 基本数据流模型

即时处理大量的数据本身就是一项非常有难度的工作，数据流具有持续不间断、速度快、规模庞大等特点，不可能也不需要所有的数据都进行永久性的储存，而且整个数据流以及网络处理环境也都处于不断波动的复杂不确定性状态下，很难对整个数据的全貌精准把握。随着数据流理论与技术的不断发展，于此同时例如 Twitter 的 Storm^[37]、Yahoo 的 S4^[38]以及 linkedin 的 Kafka^[39]等实际开源系统

也得了开发和应用，见表 2.2 大数据工具列表所示。

2) 批处理。2004 年谷歌公司提出最具有代表性的批处理模型 Mapreduce 编程模型^[40]。首先 Mapreduce 模型将用户的基本初始数据进行分类分块处理，然后由不同的 Map 任务区来处理划分好的块体，Map 任务从输入中解析出值对集合，进而根据用户自定义的 Map 函数得到中间结果，并存入本地硬盘。Reduce 任务从硬盘中读取数据后，再根据关键因素进行排序，然后将具有相同关键因素的值组合在一起，最后再由用户已经自定义好的 Reduce 函数对这些值进行计算，得到最终的结果。图 2.2 是 Mapreduce 执行流程图。

大数据的应用有很多种类型，在实际处理大数据中会根据流处理或是批处理各自的特点来相互配合应用，而不仅仅单独采用哪一种处理方式，比如根据处理时间的要求不同，有一些互联网公司会将业务分为在线、近线和离线，Linkedin^[41] 社交网站就是这样划分的。在线处理的时间一般都会比较短，用秒甚至毫秒来计算，就可以采用流处理，离线的处理时间要求就没有那么高，可以用天为单位，基本上就是采用批处理的方式。

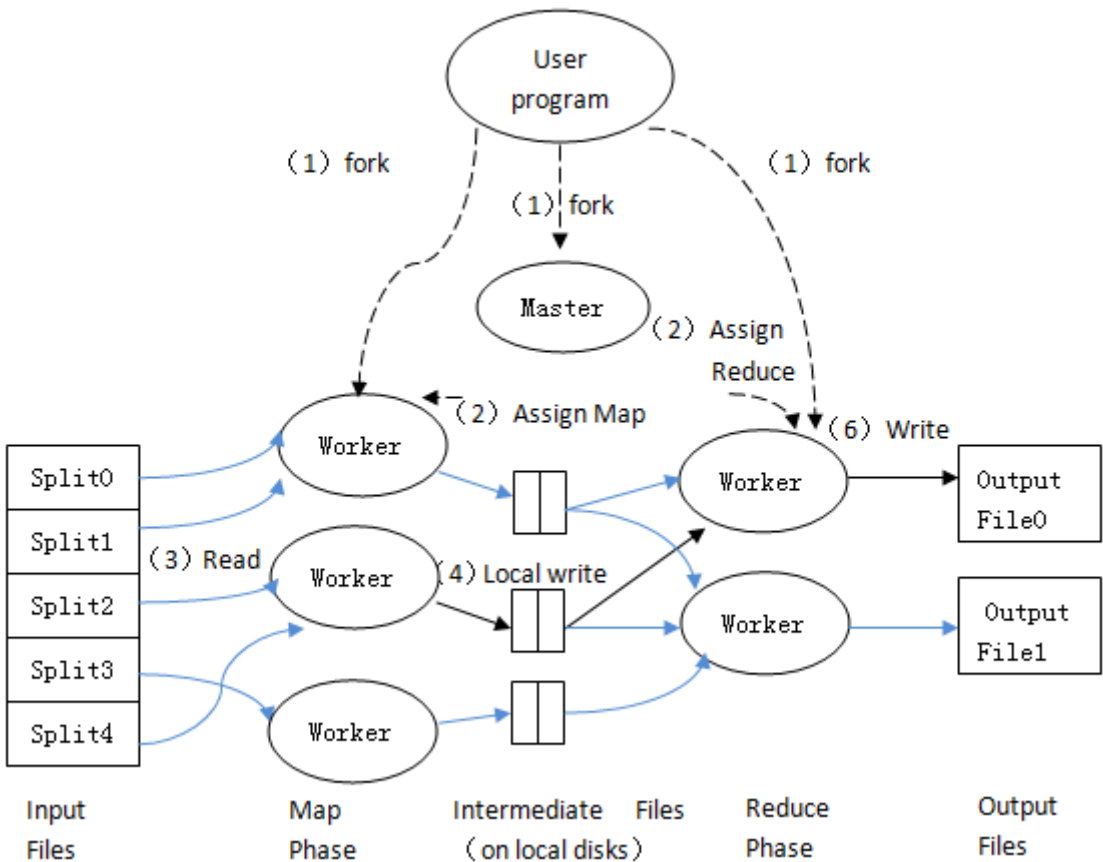


图 2.2 Mapreduce 执行流程图

大数据的来源十分广泛，虽然数据类型不尽相同，其应用的领域也是涵盖包

括理解客户、满足客户需求，业务流程优化，提高医疗与研发，提高体育成绩，优化机器和设备性能，改善安全与执法，金融交易等各个方面，但是其处理流程基本一致。由中国人民大学网络与数据管理实验室开发的 ScholarSpace^[42]系统已经在众多领域提取数据、集成并储存，在经过数据分析将储存的数据进行挖掘和提炼，最终显示出数据解释结果，完整的体现出大数据处理的一般流程，如图 2.3 所示。

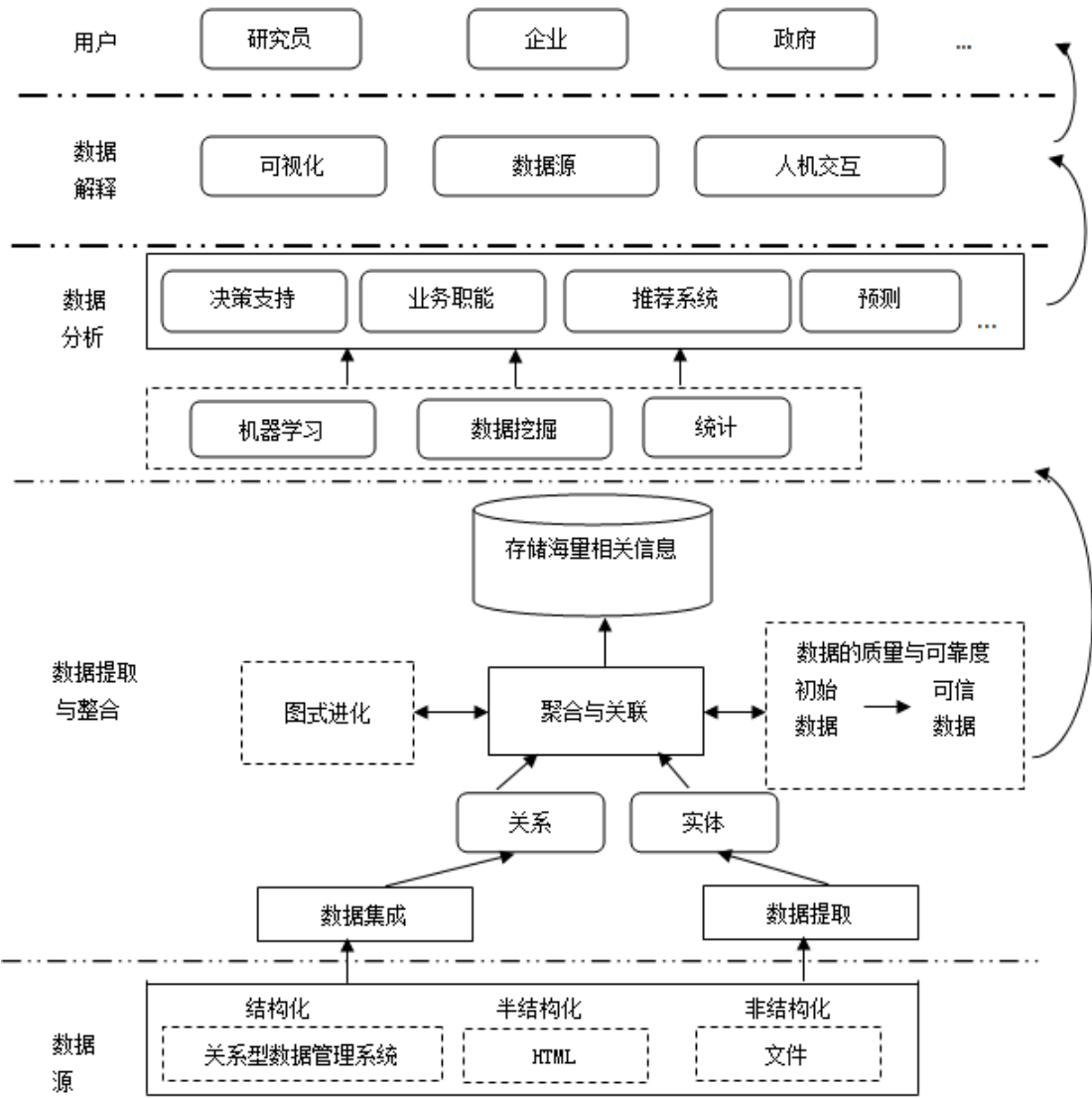


图 2.3 大数据处理基本流程

1) 数据抽取与集成。处理大数据的第一步工作就是对数据源的数据进行抽取和集成，从中提炼出关系和实体，对和其相关的实体再采用自定义的结构来存储，为了保证数据的质量和真实准确性，同时还要对抽取和集成的数据进行清洗。随着数据源的不断更新和扩大，集成数据的方法也在不断的提高和发展，数据抽取

与集成的方式按照数据集成模型可以划分为基于物化或 ETL 方法的引擎 (materialization or ETL engine)、基于联邦数据库或中间件方法的引擎(federation engine or mediator)、基于数据流方法的引擎(stream engine)和基于搜索引擎的方法 (search engine)^[43]。

2) 数据分析。在整个大数据处理流程中最为重要的部分是数据分析能力, 大数据能否发挥其价值与数据分析能力直接相关, 可以根据不同的需求将抽取与集成的原始数据整体选择或者部分进行分析。与传统意义上的数据分析不同, 大数据时代的数据分析还要面对一些新的挑战: 随着数据源的不断更新和扩大, 有价值的信息往往密度很低, 大量的数据噪音夹杂在成倍增长的数据中, 因此, 在进行数据分析前必须对数据源进行清洗, 这就对计算硬件以及算法提出了更高的要求。大数据处理基础的平台是云计算技术, 这就要求大多数算法都要适应云计算的框架, 算法要更具有兼容性, 最后当数据增加到一定规模时, 要运用统计学的方法, 从小量数据中挖掘出有用信息的算法同时适用于大数据, 邦弗朗尼原理就是一个典型的例子^[44]。

3) 数据解释。数据分析是大数据处理核心, 但怎样让经过数据分析后的结果呈现给用户, 并使之理解也同样重要, 如果分析的结果准确无误但不能很好的解释, 大数据也不能被很好的应用同时还可能产生误导等致命错误。大数据分析的结果往往极为复杂, 传统的解释方法已不能满足大数据时代的需求, 需要采用新的方法来提升大数据的解释能力: 引入可视化技术, 对分析的结果用可视化的方式形象地呈现给用户, 常用的可视化技术有标签云(tag cloud)、历史流(history flow)以及空间信息流(spatial information flow)等。让用户参与一定程度上的数据分析过程也能让其更好的了解和理解数据分析, 同时通过数据起源技术^[45], 也可以更好的让用户理解最终的分析结果。

关键技术分析

多种关键技术协同发挥作用, 产生化学反应才能完整的体现出大数据的巨大潜在价值。文件系统来支持充当提供最底层的存储能力, 在其上建立数据库系统来进行数据管理。通过构建索引提供例如高效数据查询等常用功能, 然后对数据库中的大数据运用数据分析技术提取出有益的知识。

1) 云计算: 运用大数据的关键技术。假如把大数据的各种运用比作一列列“动车”, 支撑起这些“动车”运行的“高速铁路”就是云计算。也正是由于云计算技术在存储数据、分析与管理等方面相对于传统算法的突破, 才使得大数据能最大化的发挥价值。

在所有的“高速铁路”中, 最具有代表性的是谷歌, 创新满足需求, 同时需求驱动创新。2006 年谷歌公司第一次提出了云计算的概念, 为了处理大量的 Web

数据，其自行研发了一系列云计算技术和工具。不可多得的是谷歌公司并没有将这些技术屏蔽起来，而是采用论文的方式部分公开出来，正是因为这些公开的论文，使得一系列大数据处理技术例如 GFS(Google file system) ,MapReduce ,Bigtable 得到了了解和应用，同时以 Hadoop^[46]为代表的云计算开源工具也得到了发展。GFS 是基于大量廉价服务器上构建的可扩展的分布式文件系统，主要应用于容量较大的文件，同时读大于写的情况，正是由于这样的读写应用不能满足需要，又促使谷歌公司重新设计 GFS，并改名为 Colossus。云计算包含很多的技术，通过谷歌公司介绍的云计算能够快速、准确的把握云计算的重点和精华。根据谷歌公司公开的论文以及查阅梳理的相关文献资料，联合大数据处理的需求，整理出谷歌公司的技术演化过程。如图 2.4 所示。

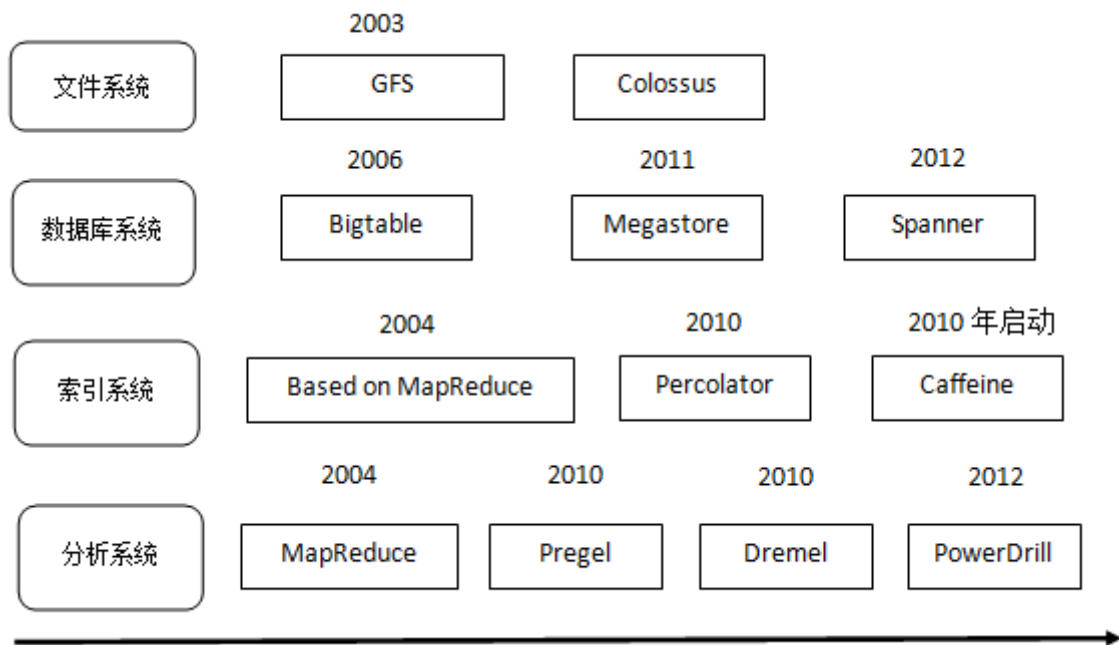


图 2.4 Google 技术演化图

2) 大数据处理工具：在相当长的一段时间里都会选择数据关系库进行数据管理，进入大数据时代以后，数据管理、分析等的多样化需求使得关系数据库已不再适用。Hadoop 是现在比较流行的大数据处理平台。最开始 Hadoop 的 Doug Cutting 模仿 GFS ,MapReduce 完成搭建一个云计算开源平台，然后奉献给 Apache。Hadoop 逐渐发展成为包含文件系统、数据库、数据处理等多功能的一系列生态系统。从一定意义上来说，Hadoop 是大数据处理工具的标准。如今的研究热点也发展为将 Hadoop 改进，并应用于不同场景的大数据处理中。研究成果主要体现在改进 Hadoop 平台的性能^[47]、查询处理的效率^[48]、建构使用索引^[49]、构建 Hadoop 上的数据仓库^[50]、连接 Hadoop 和数据库系统^[51]、数据挖掘^[52]以及推荐数据等^[53]。

除此之外，还有一些大数据处理的工具，它们有的是比较完整的处理平台，或者是专门针对特定的大数据处理应用，这部分处理平台和工具有的已经应用在商业领域，有的是开源软件。应用在商业领域的产品，很多是在 Hadoop 基础上的扩展功能，或者是提供与 Hadoop 的数据接口。

大数据时代面临的新挑战

大数据时代的数据和传统的数据管理比较起来有多源异构、分布广泛、动态增长、先有数据后有模式等特点，使得大数据时代面临着新的数据管理挑战。

1) 大数据集成

数据的广泛性体现在数据散乱的分布于不同的数据管理系统中，需要先对数据进行集合才能进行数据的分析。数据集成并不是什么新的技术，但是在大数据时代却面临着广泛的异构性、数据质量不高等新的挑战。

2) 大数据分析

在大数据时代，传统的数据分析技术越来越不能适应半结构化和非结构化数据的迅猛增长，包括数据处理的实时性、动态变化环境中索引的设计、先验知识的缺乏等。

3) 大数据隐私问题

隐私问题从计算机诞生的那一刻起就已经存在了，随着互联网的兴起，越来越多的数据以数字化的形式储存在计算机中，并能很方便的进行传播，但同时隐私问题也就变得越来越严重。主要体现在隐性数据的暴露、数据公开与隐私保护的矛盾、数据的动态性等方面。

4) 大数据能耗问题

高能耗在如今能源价格不断上涨、数据中心存储规模不断扩大的情况下，已经成为制约大数据发展的主要瓶颈。人们对小型集群和大规模数据中心如何降低能耗的问题都没有引起足够的重视，也没有太多的研究成果。从现有的一些研究成果来看，可以从采用新型低功耗硬件、引入可再生的新能源等两个方面来改善大数据的能耗问题。

5) 大数据处理与硬件的协同

大数据的发展离不开硬件的快速升级换代，但同时也对大量不同架构硬件共存产生了影响。大数据管理在日益复杂的硬件环境中面临的挑战包括硬件异构性带来的大数据处理难题、新硬件给大数据处理带来的变革两个方面。

6) 大数据管理异构性问题

从数据集成到数据分析，再到最后的数据解释，整个大数据流程中都有易用性的影子，易用性的挑战主要体现在大数据时代的数据量巨大，分析比较复杂，得出的结果形式呈多样化；其次是随着大数据的普及与广泛应用，各行各业都需要

分析应用大数据，但是往往从业人员又不是数据分析专家，所以繁杂的分析过程和形式多样的分析结果等这两个原因都导致易用性成为大数据时代所面临的巨大挑战。目前关于大数据易用性的研究还是处于初级水平，可以从关注可视化原则、匹配原则和反馈原则这 3 个基本原则来达到易用性。

2.1.3 房地产开发企业的大数据应用

大数据时代的来临能够为掌握大数据资源并充分挖掘利用其价值的企业带来巨大的收益和强劲的发展后劲，因此，应用大数据做好开发运营是我国房地产企业提高自身竞争力的关键性因素。大数据应用是房地产开发企业的重要资源，是房企商业能力的衡量标准，同时也是未来发展的方向。与传统的开发销售模式不同，大数据应用会贯穿从产品设计开发、生产、降低风险与节约成本、产品销售等的全过程，而不是单一的只注重拿地。大数据多样复杂的特点使得房地产开发企业、房地产中介或者物业管理企业的业务范围都向多元化和综合性转化，各企业之间的协同也越来越广泛，协作越来越紧密，本文主要介绍房地产开发和营销方面的应用。

大数据在房地产开发中的应用

大数据时代的到来为房地产企业的理性开发提供了强有力的数据保障；我国房地产业的兴盛与辉煌已有长达十五年的历史，掌握了大量开发投资方面的历史数据，其中包括城市地形地貌，经济发展与人口数量，城市定位与政策导向等，房地产企业可以有针对性的全面分析这些大数据，预测市场未来的供需情况，评估出项目的投资回报，从而理性地进行开发。利用大数据不仅可以进行住房供需分析、理性拿地，同时还可以使房地产开发企业业务范围多样化。万达集团和绿地集团等知名房企借助大数据来扩展其酒店和旅游等项目的多元化投资服务，探索出房地产市场以外的盈利空间。正如维克托所说，数据的重复使用不会降低数据的价值，同时还可以揭示其潜在的巨大价值，大数据经过专业的挖掘分析，可以带来额外的利润。在个人信息数据化和房地产业思维变革的热潮带动下，大数据条件下的创新性投资也能成为企业新的利润点；计算机和互联网的大力发展使得越来越多的购房者信息更容易被开发商捕捉和存取，经过专业的分析后，看似与房地产企业不相关的信息，如购房者的日常行为习惯、他们偏爱的出行路线等都可以成为房地产开发投资的新机遇，成为盈利的新突破点。表 2.2 整理出大数据在房地产开发中的应用。

表 2.2 大数据在房地产开发中的应用

相关企业	数据来源	数据应用	实践成果
万科	移动互联网客户信息	电信运营商等第三方提供大数据，系统分析和模型处理，估算地价	掌握土地资源市场情况，合理应对地价上涨
万达、绿地	历史开发信息	历史数据匹配商业用地需求信息	旅游、酒店等多元化投资
花样年	住户消费需求	手机 APP 连接客户与商户	社区电子商务、金融和酒店服务
万科	业主资料	配套消费、产业配套与 480 万业主数据资源衔接	构建“城市配套服务商”
世茂	业主健康状况	与东熙软件合作，以动态跟踪设备搜集数据，干预和管理个人健康	推出“健康云”业务向业主提供健康监督和咨询服务
金地、绿地	业主信息	借助互联网云计算形成 O2O 服务模式，构筑产业生态合作圈	提出“智慧城市”、“云服务”等概念
Windermere	GPS 导航仪	从近一亿匿名驾驶员行车导航信息中挖掘出最佳路线	为潜在购房者规划上下班行车路线和时间

大数据在房地产营销中的应用

房地产企业可以通过信息系统实现精确营销，房地产商可以通过其自身的数据优势，建立独立的客户信息系统，将客户进行系统的分类，再通过大数据挖掘，提炼出客户的信息，从而有针对性的实现精准营销。与此同时，一些知名房企主动转向了电子商务，对营销模式进行了深刻变革，比如新峰地产规划出五个大数据应用系统，房谱网可以为消费者筛选出他们中意的房产来满足其需要，自动评估系统通过大数据处理技术实现了对房产价格的自动评估功能，消费者只要输入房产的相关数据，系统就会自动估计出房价，同时为提供相应的贷款和税收等相关信息。与万达集团的电子商务运行相仿，新峰地产同样采取 O2O 的方式，线下的营销部会通过线上的客户信息与客户取得联系，这就要求企业既是大数据的拥有者同时又是数据处理的行家，从而提高了对房企的要求。

上述营销方式基本不用第三方数据平台，房地产商主动扩宽原来的业务范围即可以完成，而房地产商为了扩大其应用大数据的优势，可以与第三方平台寻求合作以共享数据资源。如 CNFS 房地产大数据中就包含了囊括中国 289 个城市从政府到房地产开发商再到二手房交易市场的房地产信息，这种第三方平台还有美国著名的众筹公司 Realty Mogul，Realty Mogul 以互联网众筹的方式成为房地产商和投资者之间纽带，为小资本投资提供了可能性，而它为小资本投资者所提供的房产信息和分析结果正是来源于其掌握的大量数据。轻资产运营的典型代表宜居中国最先提出中国房地产流通服务商的理念，借助成熟的计算机技术，宜居独具特色的推出了“克而瑞房价分析系统”，为广大的置业用户提供良好的服务。好屋中国则是大量储存个人购房信息，然后搭建起大数据资源库，通过特定的算法找

到购房者的需求，并以之为导向进行投资运营项目，从而大大提高了房地产交易的成交量，凤凰房产网的累计访问量已经超过 160 万，对海量信息进行有效的分析，可以准确的发现客户的真正需求，从而更好地为精准营销服务，以大数据思维促进房地产业良性的发展。表 2.3 呈现了相关企业应用大数据技术改变和发展营销的模式。

表 2.3 大数据在房地产营销中的应用

相关企业	数据来源	数据应用	营销成果
新峰	产品信息	房谱网等应用系统	房源自动推荐、房价预测等
中金	市场流通信息渠道	CNFS 房地产大数据系统	全面、实时、客观的房地产数据共享
Realty Mogul	投资者资质	网络众筹平台	助力房地产小规模投资
宜居中国	历史交易信息	克而瑞房价分析系统	房地产流通服务商理念
好屋中国	开发商、房产经纪人	房地产大数据库	匹配购房需求与房源信息
凤凰房产	用户访问日志	凤凰房产网	实现精准营销定位
腾讯&碧贵园	社交软件用户资料	大粤房产网	“十里银滩”创纪录开盘

2.2 能力成熟度模型

2.2.1 成熟度的概念

成熟度是一种比较新型的评价方式，在很多领域都有广泛的应用，John Schlichter 把成熟度定义为一个正在完全发展的过程阶段，这个过程可以解释为什么能够成功以及应该如何避免问题的理解和可见性^[54]。美国著名的质量大师菲利普·克劳士比（Philip B. Crosby）在他的质量著作《质量免费：确定质量的艺术》中首先提出了成熟度方格理论，他认为任何一个企业的质量管理水平都要经历从最初的不成熟逐渐走向成熟，同时不确定期，觉醒期，启蒙期，智慧期和确定期构成了衡量质量管理水平的五个不同阶段。这也为将来提出的成熟度模型和发展奠定了理论基础。

2.2.2 能力成熟度模型理论

最早的成熟度模型（Capability Maturity Model, CMM）研究是在软件开发项目管理上，在上个世纪 70 年代，美国国防部对一些软件开发项目进行了系统的研究，研究发现超过 70% 失败的项目都是由于管理问题，而不是因为开发技术的原因^[55]。1986 年，卡内基·梅隆大学软件工程研究所（CMU/SIE）和美国国防工业协

会在得到美国国防部的资助后开始研究能力成熟度模型，五年后推出了 CMM1.0 版本，用以帮助软件公司建立和实施过程改造计划。能力成熟度模型把软件公司开发软件看作一个过程，监控和研究整个软件的开发和维护过程，以实现软件开发过程的科学化、标准化。随后建立了一套在全世界范围内推广的软件评估标准，可以更好的帮助用户评价软件开发能力，随后不断对软件进行升级换代。在软件开发过程的管理和评估工程能力的应用中经常用到能力成熟度模型。能力成熟度模型从 1987 年实施认证，经过近三十年的发展，现在已经成为软件业最具权威的评估认证体系。其包含 5 个等级（如图 2.5），18 个过程域，52 个目标和 500 多个关键实践^[56]。表 2.4 列举了五个等级的基本特征与要求。

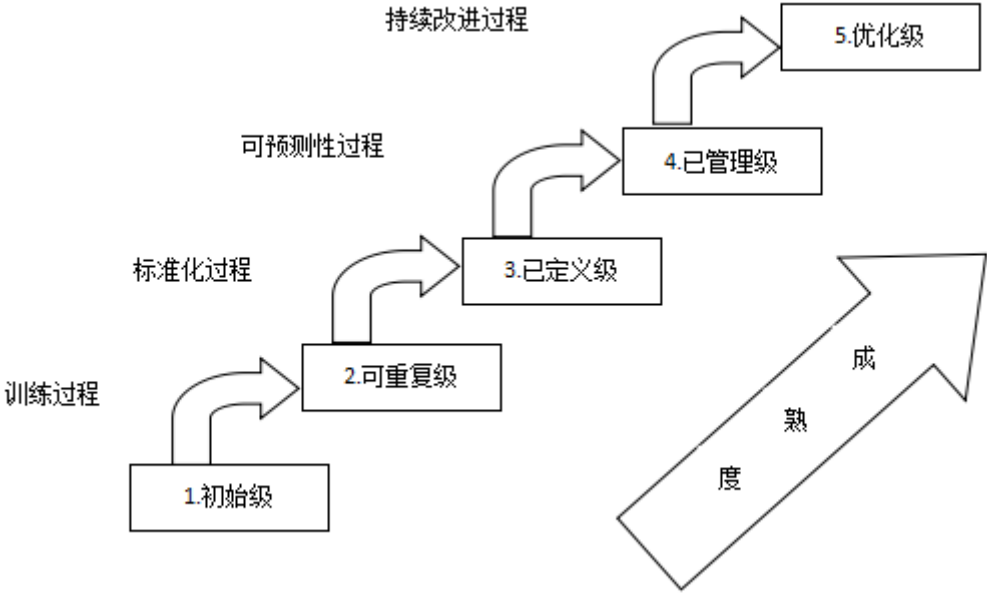


图 2.5 软件过程成熟度模型的五级图

表 2.4 五个等级的特征与要求

级别	名称	基本特征	主要特点与描述
1	初始级	杂乱无序	企业自身的软件开发环境并不为稳定，一旦出现问题后，就会放弃原定计划，从而集中精力编程和测试，此阶段杂乱无序，难以预测进度、预测、功能和质量。
2	可重复级	制度化、纪律性、重复性	在以前项目的经验上，建立管理软件项目的政策，同时按照政策制定相关措施，在这个阶段企业管理的项目过程逐渐制度化、有纪律且能重复。
3	已定义级	标准化	在已定义级中，企业需要制定关于软件工程和管理工程特定的软件开发与维护过程的文件，并集成这些过程，组成一个

			整体。该阶段的软件开发过程已经标准化。
4	已管理级	定量化	此级别中，企业要建立产品和过程定量的质量目标，还要明确制定度量标准，精细度量项目的活动生产率和质量。实现提高软件产品的预期质量目标。该级别的企业已达到过程定量化。
5	优化级	持续改进	在优化级中，企业需要不断优化过程，主动采取措施了解过程中存在的优缺点，预防可能出现的缺陷。同时，在优化级中，要求企业分析新技术产生的成本和效益，并验证各个过程资料的有效性，并提出修改性意见。优化级的企业可以自发的改进。

初始级 (Initial) :处于初级阶段中的组织，在实施项目时由于工程管理制度不健全，没有规范的标准和准则，每项工作只能按照特定的方法来进行。需要依靠有能力的管理者并组建起一只优秀的团队才能取得项目的成功。但是，项目实施的时间和费用在没有完善的管理体系和详细的计划下总会超过预期，从而导致很多行为都是在做紧急重要的事情，而不是按照事先规划好的先做重要但不紧急的事情。在该阶段的组织中，对当前人员的配置主要体现在项目管理的过程中，一旦人员发生了变化，也会引发过程的改变，从而加大了预测时间和费用的难度，因此，不可能在这个级别精准的预测产品的开发时间和费用。

可重复级 (Repeatable) :该级别中，根据类似项目和产品实施的经验可以得出该软件项目的管理行为、设计和管理技术，所以这一阶段被称为“可重复级”。在这个阶段，采取详细的跟踪费用和进度措施是一个完整过程中最为重要的第一步。管理人员能够在危机发生后及时发现，同时采取相应的措施来改正，防止出现真正的危机，这是较初始级提高的地方。如果没有采取一定措施详细的跟踪费用和进度，管理人员很难在真正的危机出现前发现。同时可以参考未来项目采取的措施，对拟定项目实现的期限和费用。

已定义级 (Defined) :在这个级别中，企业可以为软件生产过程编制较为完整的计划和管理体系文档。并对其管理和技术进行了定义，并对软件的开发过程进行持续的改进，最终通过评审来保证软件的质量，在定义级中，通过引用 CASE 环境来实现提高质量和生产率的目的。但在此阶段中，采用“高技术”会使这一危机的驱动过程变得更加混乱。

已管理级 (Managed) :处于这一级别中的企业，针对不同项目设定不同的质量与生产目标，项目的质量和生产目标也都在企业的实时监控内，可以随时采取有效措施修正与预期的偏差。通过管理部门运用统计质量来区分随机偏离与非随

机偏离，其中包括质量或生产目标偏离。

优化级（Optimizing）：该阶段的目标是一个持续改进软件的过程。通过运用统计质量和过程控制技术来指导优化级的组织，从项目中汲取各个方面的知识，并把正反馈循环融入在软件过程中以达到提升生产率和质量的目的。

表 2.4 描述了五个等级的特征和要求，为软件的过程能力提供一个阶梯式的改进框架，在总结以前软件工程的经验教训后得出了一个过程改进的流程图，说明一个软件开发项目在实施过程中要完成哪些工作，各项工作彼此之间的关联性，并根据逻辑关系安排工作的先后顺序，不断积累总结实践经验逐步使得软件组织趋于成熟。CMM 思想问世至今被不断修订，并广泛应用于软件业，对其产生了重要的意义，同时对建立项目管理成熟度模型也有深刻的影响。

2.2.3 几种成熟度模型的介绍

在广泛推广使用能力成熟度模型以后，一些组织和个人把项目管理知识体系也融入到能力成熟度模型之中。他们从项目的角度，研发出的成熟度模型能够满足不同需求的项目管理，包括以下几种：

① CMMI 模型

市场逐渐成熟后，继开发 CMM 模型之后，2000 年 8 月 SEI 推出了集成化能力成熟度 CMMI（Capability Maturity Model Integration）。CMMI 模型中融入的 6Sigma、TQM 和 ISO9000 等标准体系能够显著提高软件过程开发质量和效率^[57]。主体内容框架中 CMMI 模型包含了 SW-CMM（软件工程 CMM）、SE-CMM（系统工程 CMM）、SA-CMM（软件采购 CMM）、IPD-CMM（集成产品开发 CMM）、P-CMM（人力资源 CMM）以及 IPPD-CMM（集成产品和过程开发 CMM）和 CMM 比较起来，CMMI 更加适用于广泛的知识工程领域，而 CMM 主要是针对软件领域。CMMI 模型有 SW-CMMv2、EIA731 和 IPD-CMMv0.98 三种形式。CMMI 是建立在 CMM 的基础上，比 CMM 模型具有更好的延展性，有多种多样的组合形式，开发与维护实践活动囊括了项目管理、硬件工程、软件工程、流程管理、系统工程以及其他的支持过程。例如 CMMI-SW、CMMI-SE/SW、CMMI-SE/SW/IPPD 和 CMMI-SE/SW/IPPD/SS。

CMMI 模型常采用阶段式表示法和连续式表示法两种方法表示。为了便于企业组织间能力成熟度的比较，CMMI 模型保留了 CMM 中的阶段式表示法。同时，CMMI 中又增加了连续式的表示法用以促进企业组织更加切合实际的改进内部软件过程。

可用初始级、已管理级、已定义级、定量管理级和优化级来分别表示 CMMI

模型成熟度等级的进化过程；在此阶段中，软件过程呈现出杂乱无序的状态，企业组织没有提供维持流程的稳定环境，最后能够成功往往也是依靠个人的努力和团队的作用；在管理级中，有专门的项目管理过程对项目的成本、进度和功能进行监控，与此同时，这些项目管理过程可以对必要的过程进行规范化、分析以往近似应用项目的成功经验；文档化和标准化已定义级的软件过程，按照软件开发过程中所有项目的开发与维护标准来进行；定量管理级，制定出软件过程和产品质量的详细度量标准来加深对其的理解和控制；在优化级中，过程的改进利用过程的量化反馈和新技术、新方法来实现，促进最优的软件工程开发实践。五个等级中的每个等级都是下一个等级的基础，并且由不同的过程域组成，过程域通过实践体现，又通过实现实践活动来完成过程域的目标。

② OPM3 模型

OPM3 模型最早是由美国项目管理协会研究开发得出 ,它不仅可以提高组织有效管理各个项目的能力，同时也能起到提升组织战略能力的作用，具有重要的意义。

构成 OPM3 模型的三维立体结构分别是成熟度的 4 个等级(如图 2.6),9 个项目管理的知识领域和 5 个基本过程以及层次不同的项目管理对象，这样的三维结构看似复杂，但同时为用户提供了一个清晰的层次，使用户在理解项目管理成熟度模型^[58]时变得更加直观。

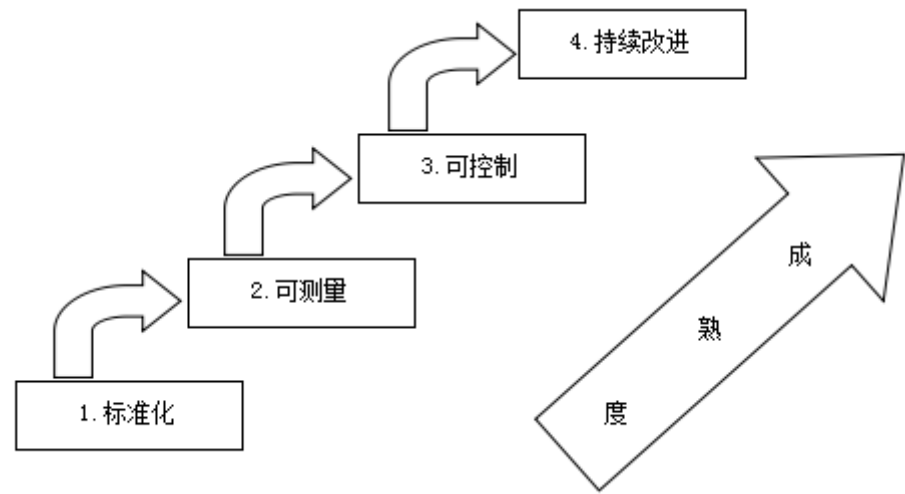


图 2.6 OPM3 的成熟度等级

图 2.6 是 OPM3 的成熟度等级。标准化、可测量、可控制和持续改进四个等级共同组成了 OPM3 模型的成熟度等级，这四个等级同时又处于不断改进不断提高的动态循环中。各个等级都按照其相应的路径和方法进行改进。如最佳实践、路径、能力、可见结果和主要绩效指标以及模型的范畴等构成 OPM3 模型的因素之

间存在着相互的关系，这些元素共同构成了 OPM3 组织项目管理成熟度模型的基础。

③ K-PMMM 模型

美国哈罗德·科兹纳博士在发表的“项目管理的战略规划：项目管理成熟度模型”中首次提出 K-PMMM 模型。项目管理成熟度模型（K-PMMM）有着相当广泛的应用，其应用领域包括软件行业 and 项目管理^[59]。北电网络公司对该模型的有效性和科学性进行了检验和证明^[60]。站在企业项目管理规划的角度，并结合项目管理战略任务和企业的战略发展，把 K-PMMM 模型划分为五个等级^[61]。如图 2.7 所示五个具体过程的阶梯模式。

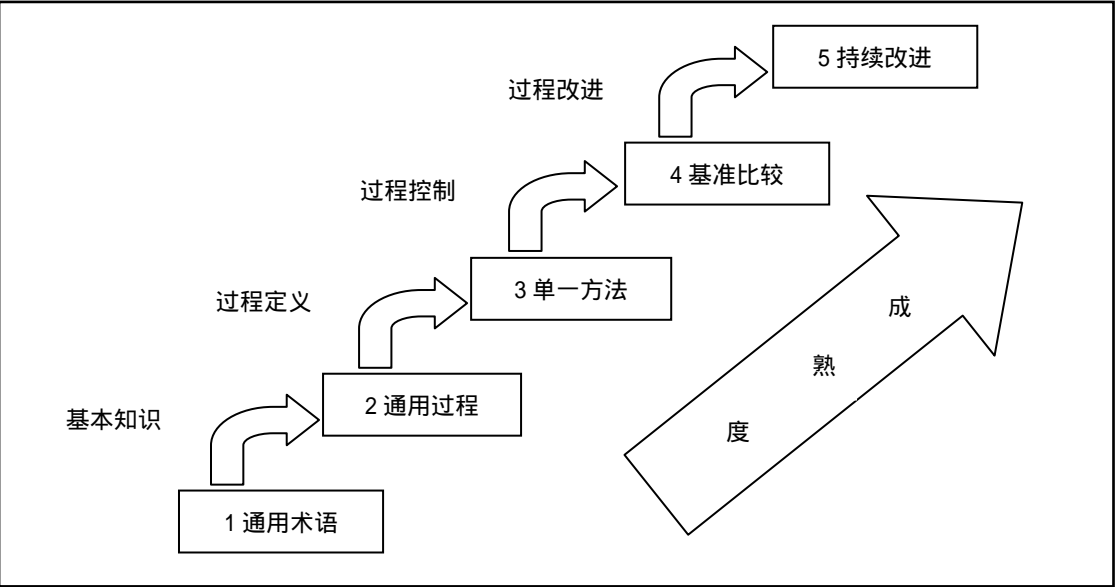


图 2.7 五个过程能力的阶段模式

通用术语、通用过程、单一方法、基准比较和持续改进共同构成了科兹纳博士项目管理成熟度模型的五个层次。第一个层次级别是通用术语，处于这一级别上的过程成熟度是最低的；第二层次的通用过程比第一层次的通用术语相对要完善一些，但总体来说这两个过程都还不太成熟，从第三个层次单一方法开始逐渐向成熟转化，较高级别的层次则是卓越管理的典范，企业在实施的过程中出现工作之间相互重叠也可能会使这五个过程能力层次之间相互重叠。但是从不成熟阶段向成熟阶段的关键跃进过程中则不能发生重叠。这五个层次是按照一定的步骤、顺序不断深入的，所以不能改变五个层次之间的顺序。表 2.5 中罗列了项目管理成熟度模型的五个层次成熟度水平特征：

表 2.5 五个层次成熟度水平特征

层次名称	特征
通用术语	对项目的认识还只是停留在概念上，同时项目管理的水平也较低；高层领导对其的支持力度不够，也没有充足的资金进行项目管理的培训和教育。
通用过程	对通用过程项目管理的重要性进行了解后，开始重视过程控制和成本控制，学习相关知识同时引进相关专业人士，也开始投入项目管理培训，并得到支持。
单一方法	管理领导层开始支持、培训和教育的经费也得到满足、支持综合过程、组织文化和非正式项目管理、卓越的行为。
基准比较	项目办公室和卓越中心重点关注基准比较、量化过程和方法的基准比较以及文化基准比较。
持续改进	总结基准比较的结果，进行分析后对形成的项目管理成熟度模型进行持续改进。

与 CMM 模型不同的是，K-PMMM 模型不仅仅是实现这一过程域就达到了该等级，而是达到每个等级的标准后循序渐进实现的，同时不能改变完成每个阶段的顺序。科兹纳博士在 K-PMMM 模型中也分析了成熟度等级提升过程中所面对的困难和障碍，并把主要关注点放在基本原理、过程定义、过程控制和持续改进上。

PMS-PMMM 模型

PMS-PMMM 模型最早是由 J.Kent Crawford 博士提出，他在 2002 年出版的《项目管理成熟度模型，提供一种达到卓越的项目管理途径》（Project Management Maturity Model, Providing a Proven Path to project Management Excellence）一书中提出了一个二维矩阵结构的项目管理成熟度模型。横向的五个等级与 CMM 成熟度相类似，纵向的 9 个项目管理知识域分别是整体管理、范围管理、人力资源管理、成本管理、沟通管理、时间管理、质量管理、采购管理和风险管理。这 9 个知识域被分成多个关键组件(Key Components)，每个关键组件中又细分设立了相应的关键绩效指标(Key Performance Index)。该模型适用于单项目的测评，不适用于项目群。管理者可以通过此模型判断自己企业的成熟度等级和改进的方法。如图 2.8 所示。

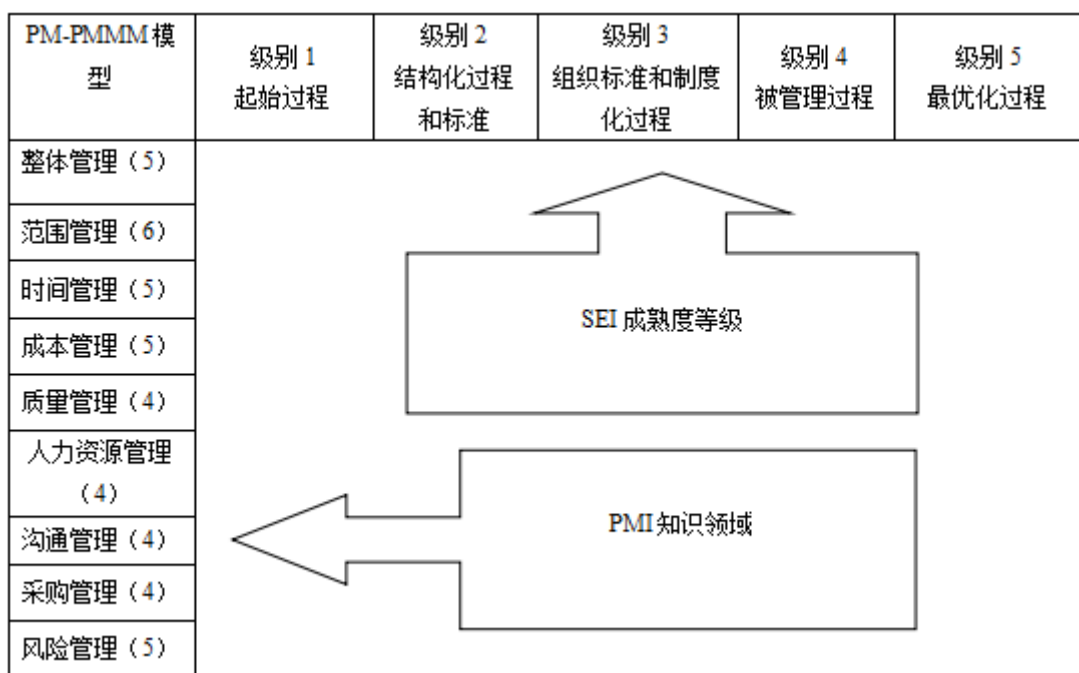


图 2.8 项目管理成熟度模型

按照结构框架和适用范围把项目管理成熟度模型分为两大类：一类是根据 CMM 模型建立的单一项目的项目管理成熟度模型。另一类则是将项目成熟度模型延伸到整个企业，对企业进行测评并提供改进方向的项目管理成熟度模型。尽管这些模型的侧重点和适用范围并不相同，相同的是他们都描述企业项目管理能力的模型。最低级别都表示管理的杂乱和无序，然后逐渐改进提升到下一个级别，而最高级别又都是处于不断改进和持续优化的动态过程之中。

2.2.4 能力成熟度模型的应用

不断发展和演化的能力成熟度的应用范围也变得越来越广泛，从最开始的软件工程领域，到如今能力成熟度模型已经可以用来评价各类工程领域的项目和组织能力。

由于技术性创新的过程往往比较复杂，具有较高的不确定性和高风险性，因此，技术创新项目的管理往往具有较高的难度。技术创新过程由很多活动或者流程衔接组成，每个过程包括多个次级过程，直至单一活动，通过衡量这些活动在实施过程中的有效性来保证其活动能按照企业的基准进行，最终使项目能够成功收尾。利用成熟度模型，从企业的现状出发，识别成熟度的等别，找准自己的定位，发现改进和提升其薄弱环节，从而实现上述要求。因此进一步确定和提高技术创新项目管理能力是提升技术创新项目成熟度等级的方法，我们应该重视和研

究成熟度模型这一前沿理论。

IT 行业、航空航天业、军工业等项目是最早用项目管理成熟度模型来测评的领域，但很少在建筑工程领域使用，同时项目质量管理的领域也是一片空白。许多年来，建筑工程项目的管理和质量管理领域一直没有一个成型的理论或者模型来指导和参考，即使最近几年我国的成熟度模型研究有所发展，呈现出上升的趋势，但是和国外的发展相比还是存在不小的差距。我国的建设项目有着自身的独特性，我们应该结合我国的政治、经济、管理体制以及文化差异等国情和建设项目自身特殊性的基础上来研究和应用国外成熟度模型理论的研究成果和实践经验。我国应用研究成熟度模型的必然要求是构建适用于我国实际情况的建筑工程项目成熟度模型，并提出相应的评价指标和评价方法。这同时也是我们分析研究成熟度模型并应用于我国工程实践所面临的巨大挑战和机遇。

2.3 房地产开发企业大数据应用能力成熟度模型

2.3.1 CMM 构建的必要性

房地产开发企业是以营利为目的，从事房地产开发与经营项目、具有法人资格、能够独立核算的企业的总称^[62]，是我国房地产企业的主体。房地产开发企业应用大数据能够从产品的设计开发、生产、降低风险与节约成本、产品销售等的全过程给企业带来新的盈利模式与空间，从而达到支持高层战略决策、同时缓解房地产行业的困境。大数据可以为房地产企业理性开发提供强大的数据支持，通过挖掘已有数据的潜在价值，房地产开发企业可以进行有针对性的多元化投资，同时大数据时代个人信息的数据化和整个房地产业的思维变革，都为房地产开发企业的创新性投资提供了新的利润增长点。

国外也有很多房地产企业运用自身的数据优势进行业务创新，美国的 Windermere 房地产就被作为大学教学的经典案例。该公司分析了上亿名驾驶员的行车导航信息，为潜在的购房者缜密的规划处在不同时段上下班的最优行车路线和时间。切实的满足了消费者的需求，同时也提高了服务的品质。

大数据的应用是房地产开发企业提高自身竞争力的关键因素，是房企的重要资源，是衡量房企商业竞争力的标准，同时也是未来发展的方向。因此，构建房地产开发企业大数据应用能力成熟度评价模型，对其进行指导从而发现房企自身的薄弱环节，并对其薄弱环节加以改进，从而提高自身竞争力是非常必要的。

2.3.2 CMM 构建的意义

房地产开发企业在发展过程中，持续改善和不断提高大数据应用能力的成熟度等级，而这一等级主要是用来表征房企从产品设计开发、生产、降低风险与节约成本、产品销售等的全过程应用大数据的水平，而等级的确定主要取决于大数据应用在房地产开发企业运营中各个不同阶段的管理与盈利情况。大数据应用的各个阶段都有一个成熟度等级与之相对应，而任意一个成熟度等级都是在房企应用大数据过程中的不断改进和演化得来的，每个等级对应一个具体的目标，当大数据应用能力达到了该目标的要求后，才表明房地产开发企业大数据应用能力上升到了这一等级层次。而当每次到达成熟度模型的一个等级时会构建起房企应用大数据过程中的一个新成分，从而持续提高房企的大数据应用能力。

房地产开发企业大数据应用能力包含从房地产开发运营过程中的某个阶段或某项工作应用大数据到全过程应用大数据的各种不同等级的能力，成熟度模型不仅是指最低级别的大数据应用能力，也不是指全过程应用大数据的能力，而是为如何提升房地产开发企业大数据应用能力提供了一套系统的框架，首先确定影响房地产开发企业大数据应用能力的关键因素，然后划分出能力成熟度的等级层次，进而运用一套评价方法来评价和分析关键因素，并提出相应的改善措施来提升大数据应用水平。因此，房地产开发企业大数据应用能力成熟度模型是一套完整的科学改进方法，能够体现出房地产开发企业大数据应用能力从较低等级逐渐改善到高等级的提升过程，并为房地产开发企业提高大数据应用能力提供了具体的指标标准，为房企的发展指明了方向具有重要意义。

2.3.3 CMM 构建的难度及工作的特点

影响房地产开发企业大数据应用能力的指标重多，通过大量阅读总结房地产开发企业应用大数据方面的书籍和文献，进行筛选，从而避免成熟度模型复杂和混淆，既要全面衡量大数据的应用能力，又要避免指标之间出现重复或包含从属关系，工作量较大，最后得到 8 个比较重要的一级标准分别是：企业战略影响指标、资源评价指标、市场评价指标、大数据分析/处理技术评价指标、应用环境评价指标、应用覆盖范围评价指标、组织评价指标和应用风险评价指标。其中资源评价指标是因为房地产开发企业在开发运营过程中要关注的焦点很多，比如土地资源，房地产开发企业要重视大数据背景下的土地市场，才能敏锐的洞察土地资源的市场走向，万科集团的土地资源数据就是来自第三方，国外的谷歌公司也通过分析海量的搜索词成功的预测了美国住房的市场供需和价格。因此，注重企

业内外资源的迅捷调配和相关经验的借鉴是房地产开发企业应用大数据的关键。大数据分析/处理技术则是因为大数据本身的价值密度低造成的，数据之间彼此又是以半结构化和非结构化形式存在的居多，这就要求数据要进行分析与处理后才能发现其所包含的巨大价值。房地产开发企业大数据应用能力成熟度模型的建立过程中影响评价指标的因素很多，如何保证评价指标选取的恰当和全面是房地产开发企业大数据应用能力成熟度模型的重点问题。

房地产开发企业大数据应用能力成熟度等级的确定也是构建该模型体系主要研究的内容之一，通常成熟度模型包含 5 个等级，成熟度等级的确定依据不同的行业和领域也有所不同。房地产开发企业在构建模型的层次划分时也应重点考虑。

2.4 本章小结

本章由三部分组成，第一部分主要介绍大数据内涵和大数据应用，包括大数据的概念、特征以及处理框架、运作流程和大数据在房地产业中的应用等；第二部分介绍了能力成熟度模型（Maturity Model）的相关理论，对成熟度模型的定义、发展和应用进行了阐述；第三部分则介绍了房地产开发企业大数据应用能力成熟度评价模型构建的必要性、意义、难度以及构建工作的特点等。本章的介绍为后文的研究建立数学评价模型奠定了基础。

第三章 房地产开发企业大数据应用能力成熟度模型的分析与构建

3.1 成熟度模型的构建分析

在上一章中分析了典型的能力成熟度模型的内容和结构，虽然这些模型关注的重点不同，形式也各异，但概括起来说，能力成熟度模型一般都包括了四个方面的内容，分别是模型研究的范畴、模型的维度、成熟度等级以及模型的内部结构，这几个方面的内容是一个成熟度模型的基础。房地产开发企业大数据应用能力成熟度模型的构建是在参考已有的几个经典成熟度模型的基础上建立的，它的构建仍然要从这几个方面入手。

模型的范畴

成熟度模型的范畴指的是模型研究的层面是什么，每个成熟度模型都有自己的研究层面，有的成熟度模型是立足于流程层面，有的则是立足于项目层面或过程层面，每个成熟度模型立足于什么层面的研究要具体看该模型的应用范围。

模型的维度

成熟度模型的维度指的是在空间上来讲，该成熟度模型是一维的、二维的还是三维的，并且指出每个维度的内容是什么。模型的维度是整个模型构建的空间基础。

成熟度等级

大数据应用成熟度等级是指房企的大数据应用水平在达到最终成熟的过程中所要经历的几个具有明显意义的成熟度水平中间平台，它是根据实际情况经过严格定义的，是实现最终成熟的发展阶梯。成熟度等级有两种划分方式，即连续式的划分和阶段式的划分，在划分的同时要给出各成熟度等级的数目、各等级的名称以及各等级的特征描述。成熟度等级的相关内容能力成熟度模型中的核心内容，一个房企的大数据应用能力水平就是通过成熟度等级来体现的。

模型的内部结构

在这里，模型的内部结构指的是成熟度模型的内部构成元素和它们之间的相互关系。可以说，各成熟度模型的实质差别就是通过其内部结构体现的。

3.2 能力成熟度模型(CMM) 对于房地产开发企业大数据应用能力的适用性分析

3.2.1 CMM 的特性

以 CMM 理念建立的房地产开发企业大数据应用能力评价模型具有以下几个特性：

最初的 CMM 模型是用于对软件开发能力评价和软件开发的过程，能够很好的帮助软件开发和能力的改进过程，而这与房地产开发企业大数据应用能力研究具有一致的目标性，这也为研究房地产开发企业大数据应用能力运用 CMM 模型扫清了障碍。

CMM 模型是一个拥有合适的抽象级别，且被广泛使用的过程改进模型。它并不限制一个企业怎样实现软件过程，这样的特点使得 CMM 具有很强的适用性，在很多行业都易于推广，不断演化发展的能力成熟度模型正好也说明了这点。

本文希望在房地产开发企业中建立一种能够使大数据应用能力逐渐提高的机制，也就是把提高作为房地产开发企业大数据应用能力整体框架的一部分，这正是 CMM 所具有的。

CMM 所具有的循序渐进的表达方式可以对每一等级建立共同的过程改进实践，继而更容易在多个重要方面聚焦房地产开发企业，从而实现其大数据应用能力能够更好更快地提高。

3.2.2 房地产开发企业大数据应用能力运用 CMM 的适用性

在实际的操作过程当中，能力成熟度模型是一个在过程中不断改进、提升的模型，其基于成熟度模型理念的大数据应用能力评价模型对房地产开发企业具有一定的现实意义。

能力成熟度是一个在过程中不断改进并提高的模型

CMM 模型是一个持续改进、不断提升的模型。通过不断的改进来实现房地产开发企业的各级目标，这种渐进性的改进和提升也在一定范围和程度上避免了革新所引发的在企业内部的冲突与动荡，从而保证了企业革新的顺利开展。

本文运用 CMM 主要是通过确定对大数据应用能力等级来发现房地产开发企业在运营和管理出现的若干问题，继而对不足和薄弱的地方提出有效的措施来保障房地产开发企业大数据应用能力的不断提高。

因此，运用能力成熟度模型的框架体系来建立房地产开发企业的大数据应用能力评价模型，将有利于房地产开发企业的决策者了解提高大数据应用能力的方

法，并为房地产开发企业在未来应用大数据时可能会遇到的问题提前作好应对措施，达到增大盈利空间的目的。

能力成熟度是一种不断提高的基准管理模型

目前许多行业对基准管理的关注度都非常高，其高效的管理方法使得很多中小企业自发的学习本行业的优秀企业，找出存在的差距，然后分析导致差距的原因，最后通过采取一定的措施来改进并不断提高自己的综合实力。能力成熟度模型中的 4 个不同等级可以作为基准管理的标准，用此来检验和校核企业的改进成果。

基于成熟度模型的理念，在房地产开发企业大数据应用能力等级评价时确定相应的等级标准作为整个大数据应用能力评价的标准，帮助房地产开发企业了解自身大数据应用能力所能达到的等级，然后正确的面对企业当前的综合现状，搭建起一座连接现状和目标之间的桥梁，在此过程中，也会激励着房地产开发企业向着既定目标不断发展。在较为详细的阐述了能力成熟度模型的适用性后，本文将引入到房地产开发企业大数据应用能力等级的研究中去。

3.3 房地产开发企业大数据应用能力成熟度模型的构建

3.3.1 房地产开发企业大数据应用能力成熟度模型的范畴

在房地产开发企业的全生命周期内，不同阶段的不同任务，往往由不同的组织部门参与完成，而各参与组织部门的工作任务和利益不同，从而形成不同目标类型的大数据应用。

大数据应用是房地产企业提高自身竞争力的关键因素，是房企的重要资源，是衡量房企商业能力的重要标准，同时，也是未来的发展方向。与传统的开发、销售模式不同，大数据应用贯穿从产品设计开发、建造生产、降低风险与节约成本到产品销售的全生命周期。房企战略的大数据应用在整个项目开发过程中始终处于主导地位，有着导向的作用。另外，房地产开发过程具有周期长，不确定因素多等特点，使得大数据应用的情况复杂多变，这就决定了房企必须在具体的开发建设客观实际情况上做好大数据应用的工作。

从上一章的分析可知，现有的典型能力成熟度模型主要分三个层面的研究，如下表 3.1 所示：

表 3.1 几个典型成熟度模型的研究层面

CMM (SEI)	PMS-PM3(PMS)	K-PM3(Kerzner)	OPM3 (PMI)
流程层面	项目层面	介于项目层面和组织层面	组织层面

以 CMM 为代表的流程层面的研究关注的是软件开发流程的研究，因此准确的说 CMM 还算不上是真正意义上的能力成熟度模型。虽然以企业层面的研究为背景的成熟度模型大都借鉴了 CMM 模型，但它们融合了应用能力的知识体系，有的模型还考虑了应用能力的生命周期，这些模型与应用能力的联系更为紧密。

本文研究的重点是基于房企开发过程的大数据应用，但是考虑到房地产开发企业内多组织部门的客观存在，以及大数据应用能力成熟度模型涉及到组织层面时所体现出的优势，所以会在模型中加入组织层面的研究。综上所述，本文拟够建的成熟度模型的主体是房地产开发企业，并且模型是介于项目层面和组织层面之间，模型的理论基础将借鉴 PMS-PM3 模型和 K-PM3 模型。

3.3.2 房地产开发企业大数据应用能力成熟度模型的维度

成熟度模型最直观的内部结构特征就是模型的维度。从上一节的叙述可知，目前常见的能力成熟度模型多数是一维或二维的，只有 OPM3 模型是三维的。最早的成熟度模型 CMM 是一个一维的模型，它的研究背景是基于软件开发流程。二维的成熟度模型基本上都是基于项目层面的研究，它们的模型融合了项目管理的知识体系和项目管理的生命周期，其中大多数是把成熟度等级作为一个维度，而把 PMBOK 的九大知识领域的全部或者部分作为第二个维度，从而构成一个矩阵式的结构模型，典型的代表是 PMS-PM3 模型。

虽然这些能力成熟度模型的维度不同，但他们也各具优势。一维和二维的模型结构简单，应用较为方便，而三维的模型结构则比较复杂，但三维的模型更能准确的反映企业的整体应用能力。但同时必须看到，并不是所有的成熟度模型都适用于三维的结构，目前所有模型的主流结构仍然是二维的。本文拟构建的房地产开发企业大数据应用能力成熟度模型的范畴是基于房地产开发企业和组织部门层次的，通过以上分析可知，该模型采用二维的结构形式最为适宜。

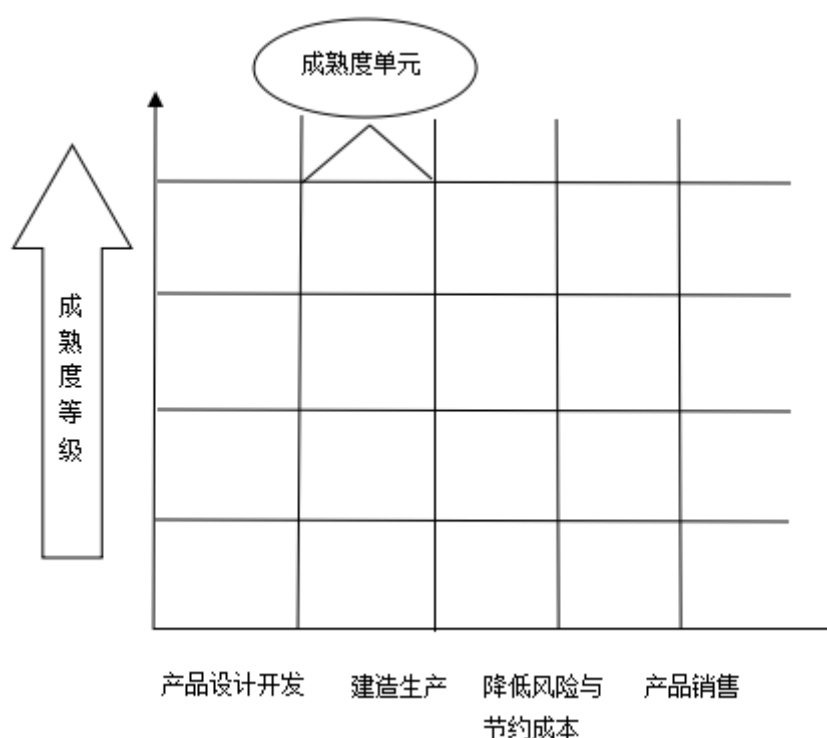


图3-1 房地产开发企业大数据应用能力成熟度模型结构

本文研究模型的第一维是房地产开发企业大数据应用能力成熟度的四个等级，第二维是大数据应用全过程的四个阶段，包括产品设计与开发阶段、建造生产阶段、降低风险与节约成本阶段以及产品销售阶段。其框架结构如图 3-1 所示，其中每个成熟度单元都是由关键过程域、关键目标和关键实践通过复杂的内部关系构成，具体内容会在后面的文章中进行分析。

3.3.3 房地产开发企业大数据应用能力成熟度模型的等级

现有成熟度模型的成熟度表示方式主要有阶段式的等级形式和连续式的百分比形式，在进行成熟度的测评过程中，阶段式和连续式的测评方法各有利弊，阶段式成熟度模型的测评方法简单易用，而连续式测评方法比较复杂，它的测评过程重在原因的分析和改进的途径。常见的 CMM 模型、K-PMMM 模型等都是阶段式的等级表示形式，而 OPM3 模型则是以百分比的形式来表示成熟度水平的高低。为了模型的测评方法简便易用，本文建立的成熟度模型将采用阶梯式的等级形式来表示。

房地产开发企业大数据应用能力成熟度模型的等级是指房企应用大数据水平从不成熟走向成熟的过程中所达到的几个具有重要意义的表征能力成熟度水平的中间台阶。这些成熟度等级具有明确的特征和定义，能在一定范围与程度上直接

表明房企应用大数据的成熟度水平，每一个成熟度等级都是房企达到下一个等级水平的基础。可以说，大数据应用能力成熟度等级为房企不断提升其大数据应用水平提供了一份引导图。

在《大数据的冲击》一书中专门设置了大数据运用模式一章，其中对大数据运用级别划分的叙述如下所示：

大数据运用级别一般分为对过去/现状的把握、发现模式、预测和优化。

对过去/现状的把握，大数据的运用是从数据收集开始的，对数据进行积累并努力找出事实，是大数据运用的第一步。

发现模式，在积累了大量数据之后，就需要使用数据挖掘、机器学习等技术，从海量数据中发现对业务有影响的模式（成功模式、失败模式等）。

预测，如果能够发现成功模式或者失败模式，接下来就可以将输入数据与这些模式相结合来进行预测了。

优化，从过去到现在积累的大量数据中发现某些模式，并对将来做出预测，到这一步，大数据的运用还没有结束。没有行动的预测，其价值就减少了一半。也就是说，基于预测的结果，需要执行优化这一行动。

基于上述大数据应用成熟度等级划分的思路，将应用能力等级划分如下，见图 3-2 所示。

第一层次：数据收集。像顾客所购买的房屋地段和购置费用，业主的个人信息、行为习惯、消费习惯、声音（通话）、血压和体重等健康数据、医疗病历等，其中包含有意进行记录的数据，但大多数都是无意中积累下来的生活日志数据。此外，这些数据也不仅仅是顾客产生的，还包括潜在业主、其他相关人员等产生的数据。

第二层次：总结关联。在积累了大量数据之后，通过关联分析，发现对业务有影响意义的模式，从而有针对性和导向性的对目标进行调整或者设置。

第三层次：预测未来。在这一层次中，房企往往能够对“具有某某属性的顾客”其个人属性与过去的关联数据进行对照分析，就可以更准确地预测出该潜在顾客会购买什么样的房源、需要什么样的服务。

第四层次：持续改进。优化包含了多种含义，例如，通过个人属性、购买记录，以及购买模式，预测出如果进行合适的推荐就很有可能会购买的房源，那么最后的画龙点睛之笔，就可以在售楼部将该房源的增值服务提供给顾客，这样的行动就是优化，其实也就是企业持续改进、不断提高服务的过程。

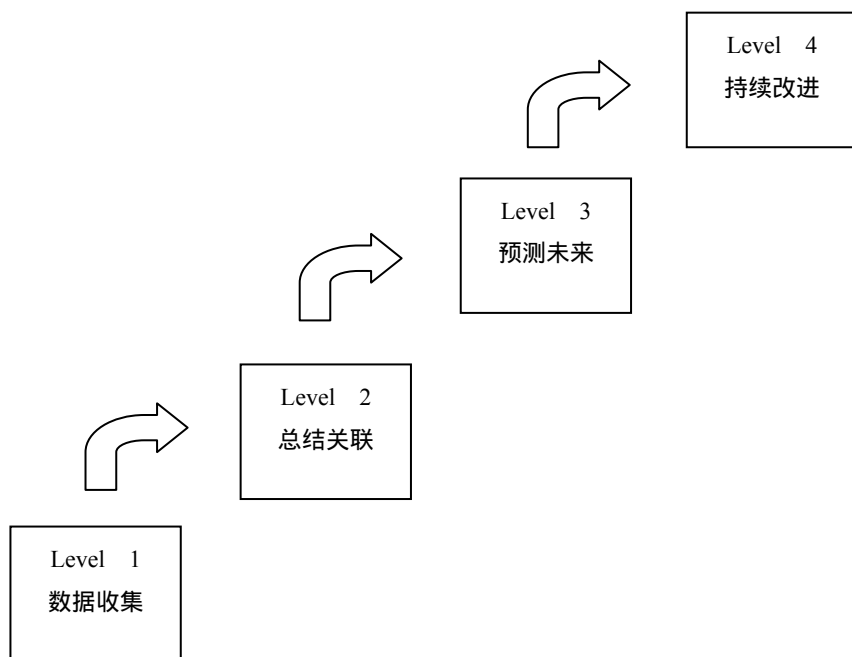


图 3-2 房地产开发企业大数据应用能力成熟度评价等级水平

对于上图中的 4 个成熟度等级,对过去/现状的把握级是最低级的别的成熟度,优化级是最高级别的,表 3.2 简要的列出了房地产开发企业大数据应用能力成熟度模型各等级的特征。房地产开发企业大数据应用能力成熟度级别的提高是一个循序渐进的过程,每个成熟度等级都是相互依赖的,上层包涵下层的目标和实践。同时,最高级并不是静止的,它与下面的等级始终存在着无休止的循环改进过程。需要说明的是,如果房企达到了较高级别的大数据应用能力成熟度水平,仍然可以使用处在较低等级时使用过的大数据应用的方法和工具,这并不意味着其大数据应用成熟度等级的降低。

表 3.2 房地产开发企业大数据应用能力成熟度等级及其特征描述

成熟度等级	描述	典型特征
1.对过去/现状的把握级	对数据进行积累并努力找出事实	数据收集。其中包含有意进行记录的数据,但大多数都是无意中积累下来的生活日志数据。
2.发现模式	使用数据挖掘、机器学习等技术,从海量数据中发现对业务有影响的模式	总结关联。通过关联分析,发现对业务有影响意义的模式,能够有针对性和导向性的对目标进行调整或者设置。

3. 预测	将输入数据与这些模式相结合来进行预测	预测未来。在这一层次中，房企往往能够对“具有某某属性的顾客”其个人属性与过去的关联数据进行对照分析，预测出该潜在顾客会购买什么样的房源、需要什么样的服务。
4. 优化	基于预测的结果，执行优化这一行动	持续改进。预测出如果进行合适的推荐就很有可能会购买的房源，就可以在售楼部将该房源的增值服务提供给顾客，这样的行动就是优化，其实也就是企业持续改进、不断提高服务的过程。

3.3.4 房地产开发企业大数据应用能力成熟度模型的内部结构

如前所述，房地产开发企业大数据应用能力成熟度模型在等级层次上分为 4 个等级，研究者往往将注意力集中于解决具有最高优先级的问题，从而容易忽略大数据应用过程改进的真正目的。因此需要一个结构框架，通过这个框架，可以使相关数据工程师检查和讨论大数据应用过程，最后根据对组织情况以及具体项目的详细了解，确定每个等级的关键问题，从而决定对什么进行改进，以及如何改进。因此，分析模型的内部结构，通过这个结构框架来更好地对房地产开发企业的大数据应用能力成熟度进行度量以及指明改进的方向。

房地产开发企业大数据应用能力成熟度模型定性的描述了成熟度的 4 个等级，该成熟度模型仍然需要采用类似传统模型的关键过程域(KPA)来衡量和判断是否达到了某个成熟度等级。同样，为了能更好的表明房地产开发企业大数据应用能力成熟度水平，本文研究在借鉴关键过程域的思想引入关键领域(Key Area，简称 KA)，关键领域定义的是房地产开发企业大数据应用能力成熟度在度量上需要关注的主要方面。传统的关键过程域关注的主要方面是“过程”，但是过程能力并不能代表大数据应用能力的全部，通常人们认为过程、技术和人员是能力的三个重要组成部分。没有任何过程能独立存在，组织方式、组织的文化等都会对项目的过程改进产生重要影响，因此本文的关键领域主要分为项目层面的关键领域和组织层面的关键领域。项目层面的领域更接近于传统意义上的关键过程域，组织层面的关键领域则拓展了关键领域的内涵，它主要包括组织的结构，组织的文化以及高层管理员的支持情况等方面。房地产开发企业大数据应用能力成熟度模型的内部结构如图 3-3 所示。

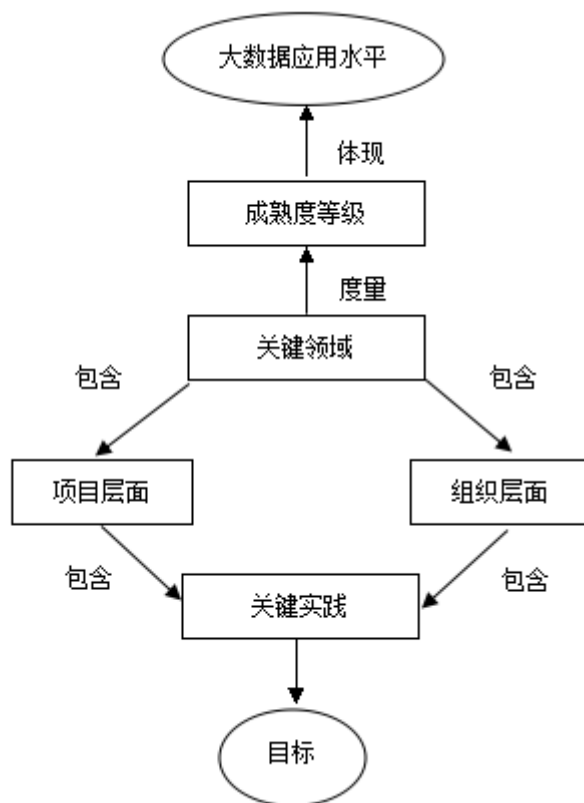


图3-3 房地产开发企业大数据应用能力成熟度模型的内部结构

“关键”领域表明这些领域是不能少的，起核心作用的。房地产开发企业大数据应用能力成熟度模型没有详细描述对大数据应用产生影响的所有领域或方面，只是描述了对大数据应用能力起着决定作用的关键领域或方面，这些关键领域对组织在改进大数据应用水平上最有成效，可以把它们看成是达到相应大数据应用能力成熟度等级的必要条件。每个成熟度等级都包含若干个关键领域，每个关键领域都指明了一组相关联的关键实践（KP），KP 描述了对有效执行和制度化该 KPA 起主要作用的从属活动和基础构造，实施这些 KP 就能实现这个 KA 的目标，当所有项目层面的 KA 和组织层面的 KA 都到达了预先设定的目标时，房企的大数据应用水平就达到了相对应的成熟度级别。可见，每个成熟度等级都可以逐级向下分解到它们在关键实践方面的可操作的定义，每个成熟度等级都有图3.3所示的内部结构。

项目层面的关键领域主要包括项目的市场预期、风险管理、盈利能力这几个方面，组织层面的关键领域主要包括了组织的结构、人员、文化及计划这四个方面。表3.3、3.4分别给出了项目层面的关键领域和组织层面的关键领域以及它们对应的目标。

表3.3CMM 项目层面的关键领域及对应目标

项目层面的关键领域	目标
市场预期	项目符合基本的市场预期供需情况，大数据应用规范化、制度化
风险管理	大致满足项目的投资回报，规避具有重大风险的项目投资
盈利能力	逐步提升项目开发的盈利能力，以便提高房企的整体盈利率

表3.4CMM 组织层面的关键领域及对应目标

组织层面的关键领域	目标
组织	确定大数据应用的管理权属，提高大数据语言使用的频率及频数，对大数据应用构架进行监督、审查和改进。
人员	从业人员要进行定期培训，满足大数据应用的基本要求，对从业人员有培养和考核的机制
文化	逐步提升项目开发的盈利能力，以便提高房企的整体盈利率，在公司内部树立大数据应用意识文化
计划	制定大数据应用计划、建立与第三方数据公司的共享机制

3.4 本章小结

本章的主要内容是房地产开发企业大数据应用能力成熟度模型的构建。分析了成熟度模型构建的一般规律，即通过模型的范畴、维度、等级及内部结构四个方面来构建成熟度模型，并对能力成熟度模型对于房地产开发企业大数据应用能力的适用性进行了分析，最后，构建了房地产开发企业大数据应用能力成熟度模型。

第四章 房地产开发企业大数据应用能力成熟度评价指标体系构建

4.1 CMM 指标体系设置要求

房地产开发企业大数据应用能力评价是在众多应用领域中，通过将定量与定性分析二者结合起来的评价方法来确定企业的应用能力。而在确定企业大数据应用能力这一过程中，首先要创建评价指标体系，指标体系创建是否合理将直接影响到评价结果的客观、公正和准确。因此，在构建房地产开发企业大数据应用能力评价指标体系时应满足以下要求。

实用性要求

指标体系里的每个一级指标和其对应的二级指标都要有明确的含义和目标导向，都能够直接的或者间接的反映房地产开发企业大数据的应用能力。一级指标的选择和层次的划分要符合逻辑思维，二级指标的划分标准应该统一，不应该出现各子标准相容的情况。

可操作性要求

指标体系的构建应该尽量简单，排除各个指标之间的模糊性，尽量用最少的指标反映出最多的信息，设定指标时还应考虑收集数据的难易程度，以便简化评价时指标的计算过程，简单且容易操作。

定量与定性分析相结合的要求

房地产开发企业大数据应用能力评价本身包含了定量和定性的因素，二者是相辅相成的关系，所以指标体系的设置也应该既有定性指标也有定量指标，这样才能确保根据该指标体系作出的评价是客观公正的，只有将二者有机的结合起来，才能对其作出准确、客观的评价。

4.2 CMM 指标体系的识别

为了准确且较为深刻的了解大数据的应用能力，合理的评价房地产开发企业的大数据应用能力，在确定了房地产开发企业大数据应用能力四个应用等级，遵循以上的构建指标体系的要求后，本小节将讨论怎样确定在房地产开发企业中大数据应用能力的关键因素以不断提高实现该模型的表达方法。结合评价指标体系的层次性和复杂性，本章构建的评价指标体系分为目标层、一级指标层、二级指标层和备选方案层。

指标体系设立的最终目的是目标层，即是房地产开发企业大数据应用能力成熟度评价，目标层下设置一级指标层，达到全面准确反映总目标的目的；为了使

指标体系更加详细，需要进一步细化指标体系，二级指标层就是具体指标的设置，是影响大数据应用能力的基本因素，同时也是构成指标体系中最重要的一部分^[63]。本文重点选择了 20 篇相关度较高的国内外有关大数据应用的文献资料，通过分析这些文献资料，结合专家的访谈调研，提炼筛选出有关大数据应用能力的多方面评价指标，总共有企业战略影响指标、资源评价指标、市场评价指标、大数据分析/处理技术评价指标、应用环境评价指标、应用覆盖范围评价指标、组织评价指标和应用风险评价指标等 8 个方面（见表 4.1）47 项指标（见表 4.2-4.10）。再经过问卷调查，邀请有关专业人士对各个指标的关联度进行打分，最后通过系统工程理论计算出每个指标的关联度评价价值，最后确定出房地产开发企业大数据应用能力成熟度评价指标。

表 4.1 大数据应用能力成熟度评价指标

代码		B01	B02	B03	B04	B05	B06	B07	B08	总计
大数据应用能力成熟度标准		企业战略影响指标	资源评价指标	市场评价指标	大数据分析处理技术评价指标	应用环境评价指标	应用覆盖范围评价指标	组织评价指标	应用风险评价指标	
参考文献	严娟 (2013)	√	√			√	√	√	√	6
	杜丹阳 (2014)	√		√			√		√	4
	Battle LM (2013)	√	√			√	√			4
	Lohr S (2012)			√	√			√		3
	Wu L (2013)	√				√			√	3
	李国杰 (2012)	√	√		√		√	√		5
	Zhihan Lv (2016)		√							1
	Wooldridge, J. M. (2010)			√	√			√	√	4
	Sheppard, S. (1999)	√				√				2
	王珊 (2011)		√	√			√		√	4
	肖剑 (2004)	√	√	√		√	√		√	6
	Wonshik Na (2016)		√		√		√		√	4
	李芬 (2013)				√	√	√			3
	Maier, G. (2010)		√							1
	Wang (2015)			√	√	√	√			4
	程明书 (2013)		√	√	√	√	√		√	6
	孟小峰 (2013)							√	√	2
	Seokhoon Kim (2016)		√		√		√	√	√	5
	Chen (2015)			√		√	√			3
	彭晓愈 (2013)		√	√			√			3

4.3 房地产开发企业大数据应用能力成熟度评价指标体系的提取

通过问卷调查对根据文献和专家访谈调研初步梳理的房地产开发企业大数据应用能力成熟度评价指标进行选择，以理论结合实践，确定大数据应用能力成熟度评价指标体系；这次初步研究听取了四位专业人士对问卷的可读性、全面性以及准确度的意见，在这四位专业人士当中，其中两位来自房地产开发企业，一位来自于大数据应用公司，一位来自高校，都具有十年以上的工作或研究经验，因为他们相信，这 8 种标准能够全面反映出一个成熟房地产开发企业大数据应用的特征，因此没有再增加新的标准。此外，根据四位专业人士的意见，新增了标准指标下的最佳实践、并对最佳实践的说明进行了修订以增强可读性和准确性。本文通过对大数据应用指标进行调查问卷直接设计、间接设计和整体设计来提取应用能力指标^[64]。具体见以下步骤。

第一，问卷直接设计。查阅以前学者有关大数据应用的文献资料，归纳整理出企业战略影响指标、资源评价指标、市场评价指标、大数据分析/处理评价指标、应用环境评价指标、应用覆盖范围评价指标、组织评价指标和应用风险评价等指标，形成房地产开发企业大数据应用能力成熟度评价指标直接调查问卷。问卷直接设计主要来源于理论研究。

第二，问卷间接设计。首先，确定问卷调查的目标企业和目标人群，考虑到大数据的普及程度和房地产开发企业的大数据应用现状，本次问卷调查的对象是重庆市部分知名房地产开发企业的中层及以上管理人员和专家。其次，向目标人群阐述本次调研的目的和步骤，同时向目标人群提供问卷直接设计得出的指标体系。然后，受调查人群依据自身对大数据应用的感知，找出大数据应用影响因素，并通过网络技术进行多次循环调查，引导受调查人群找出更多的影响因素，这些影响因素可以作为潜在的大数据应用评价指标^[65]。最后，参照行业术语，重新命名专家访谈调研得出的影响因素，根据其被提出的频率大小序列进行房地产开发企业大数据应用能力评价指标的提取，分析和筛选后得到间接调查问卷。问卷间接设计主要来源于实践调研。

第三，整体设计。总结归纳以上问卷直接设计得出的指标与问卷间接设计得出的影响因素后，识别出比较科学全面的房地产开发企业大数据应用能力评价指标如下所示：

企业战略影响指标

企业战略关乎到企业的最终目标，是一个企业自身定位与发展的指向标。它

关系到企业未来的发展，大数据应用的普及已经让众多行业受益，房地产业能否转变思想，把大数据的应用上升到企业的战略高度，已成为房地产业抢占先机，转变危机的重中之重。Mayer-SchonbergerV(2011)、Wu L(2013)、Xishang W(2010)、孟小峰(2013)、周刚华(2013)等人都在这方面做了大量研究，总结他们的研究成果，把企业战略影响指标总结为以下六个方面，具体见表 4.2。

表 4.2 企业战略影响指标

M01 企业战略影响指标		参考文献
B01.1	房地产企业高级管理层的支持	Mayer-SchonbergerV(2011)、Wu L(2013)、Xishang W(2010)、孟小峰(2013)、周刚华(2013)
B01.2	定期组织员工进行培训	
B01.3	大数据语言沟通能力的提升，大数据语言术语及尝试的普及	
B01.4	建立房地产大数据共享平台	
B01.5	在公司内部树立大数据应用意识文化	
B01.6	与第三方数据公司的数据共享机制，资源共享模式，抓住痛点，各取所需	

资源评价指标

这里的资源配置应用能力主要是针对企业内部的。众所周知，资源对于一个企业的重要性，它是企业能否正常运转的前提条件和重要保障。同样，如何利用配置好一个企业所拥有的资源，为企业用最少的成本创造出最大的价值同样是一个非常值得思考的研究问题。一个企业发展的平台与规模都是建立在其自身拥有的资源上，大数据的应用更是需要企业满足各方面的资源配置，才能展开日常的运行和服务。通过文献阅览可以发现，几乎所有研究大数据应用的学者都会把资源指标作为影响大数据应用的重要因素。这里，本文在资源指标下设置了以下 5 个二级指标，具体见表 4.3。第一个指标是优化企业的人力资源，优化企业的人力资源主要是提高数据科学家的数量，提高其在公司总员工中的比重，另一方面是提高数据科学家的素质，包括数据科学家所需要具备的各项技能。财务资源的利用指标主要是增强资金的投入，并使其得到合理利用。除了这些研究成果以外，还有很多学者如 Cukier K(2013)、Zhihan Lv(2016)^[66]、Brynjolfsson E(2013)、慈祥(2013)、钱放(2013)、王卉(2013)等人逐步开始关注房地产开发企业大数据资源评价指标的影响，越来越多的房企把资源评价作为大数据应用能力的重要考虑指标。

表 4.3 资源评价指标

M02 资源评价指标		参考文献
B02.1	优化人力资源，提高数据科学家的数量，增强数据科学家的综合素质	Cukier K(2013)、Zhihan Lv(2016)、Brynjolfsson E(2013)、慈祥(2013)、钱放(2013)、王卉(2013)
B02.2	财务资源的利用，加大大数据应用的资金投入	
B02.3	内外资源的迅捷调配	
B02.4	相关经验的借鉴，对以往同样或类似案例的处理方法加以改进或直接套用	
B02.5	土地资源数据的挖掘能力，合理购买土地	

市场评价指标

通过查阅前人的资料发现，很多研究房地产开发的学者都会把市场指标作为影响企业发展的重要因素。在此，本文在市场评价指标下设置了以下 9 个大数据应用实践，具体见表 4.4。这些市场评价指标对于评价房地产开发企业大数据应用能力都是扮演着很重要的角色，应该给予足够的重视。

表 4.4 市场评价指标

M03 市场评价指标		参考文献
B03.1	大数据应用的市场效益	Albert P.C.Chen (2002) Eddie W.L Chen(2005) Hakyeon Lee (2010) Wang (2015) ^[67] 、
B03.2	对客户消费水平的掌握以及消费习惯了解情况	
B03.3	大数据应用市场的潜力	
B03.4	竞争对手情况，调查了解同质化企业的应用水平及应用模式	
B03.5	消费者未来的市场偏好，未来市场的走势与风口导向	
B03.6	大数据应用市场的规模，探索发展空间	
B03.7	大数据应用的市场定位及布局导向	
B03.8	大数据应用市场的激烈竞争程度	
B03.9	大数据应用商业化成功的可能性，推进产品商业化	

大数据分析、处理技术评价指标

通过查阅相关学者 Sonka S(2014)、Suning T(2013)、Tene O(2012)、Polonetsky J(2012)、Wonshik Na (2016)^[68]、田溯宁(2013)、陈大川(2012)、张宝山(2012)等多人的研究，本文在大数据分析、处理技术指标下设置了 8 个二级指标，具体见表 4.5。

大型的房地产企业掌握着大量的信息，他们搜集的数据不仅仅局限于户主姓名、家庭结构、收入状况以及购房的意向等方面，互联网的普及和信息的快速广泛传播使购房者的个人信息变得更容易捕捉和存取。通过专业分析这些大数据后，可以挖掘出一些表面上与房地产开发企业毫无关系的信息，比如购房者的日常行为偏好和消费习惯等。通常情况下这些数据呈现出非结构性或半结构性的特征，但它们具有巨大的潜在价值，是房地产企业开发投资的新机遇，是盈利的新突破点。所以，大数据应用的基石是大数据分析、处理等关键技术，房地产开发企业能否真正实现大数据应用，很大程度上取决于其掌握的关键技术能否达到应用的标准。

表 4.5 大数据分析、处理技术评价指标

M04 大数据分析、处理技术评价指标		参考文献
B04.1	评估项目投资价值能力	Sonka S(2014)、Suning T(2013)、Tene O(2012)、Polonetsky J(2012) 、
B04.2	客户数据搜集与存储水平	
B04.3	数据挖掘能力，在大量低密度的数据中提炼出有价值的数据	

B04.4	非结构化数据提炼技术，不同形式或格式的数据相互间运用时兼容可视化	Wonshik Na(2016)、田溯宁(2013)、陈大川(2012)、张宝山(2012)
B04.5	机器学习能力，计算机模拟或实现人类的学习行为，不断提高的性能	
B04.6	住房供需分析能力，预测未来市场的需求	
B04.7	自然语言处理，实现人与计算机之间用自然语言进行有效通信	
B04.8	沉默数据激活能力，对以往忽略认为没有价值的数据重新开始使用	

应用环境评价指标

对房地产开发企业大数据应用能力产生深远影响的还有一个国家和地区的经济、政治、文化等等一系列软指标，在进行房地产开发企业大数据应用时应重点考虑这些影响。因此，应用环境评价指标不仅包括了房地产开发企业建设环境情况，同时还包括了其他周围与之相关的自然环境等各方面要求。通过调查问卷整体设计后，把应用环境评价指标划分为以下四个方面，具体见表 4.6。

表 4.6 应用环境评价指标

M05 应用环境评价指标		参考文献
B05.1	经济环境，区域经济发展状况，房产及信贷泡沫对实体经济的影响	David、K.-S.Chin、 R.P.Mohanty
B05.2	社会环境，当地社会人文对应用大数据的接受	
B05.3	政策环境，当地政府对大数据应用的支持力度，导向性政策的制定	
B05.4	地缘环境，周边各省市对大数据应用的状况	

应用覆盖范围评价指标

随着大数据的普及，其应用领域也是越来越广泛，覆盖范围也越来越宽阔，房地产开发企业也不例外。在理性投资、多元化开发、创新性投资、精准营销等众多方面都已开始应用大数据，同时也取得了不错的效果。Yan W，Wanli Y，Juan Y，白山，程明书，彭晓愈，肖剑等人都在此进行过研究，总结他们的成果，本文将在应用覆盖范围指标下设置以下 4 个大数据应用实践，具体见表 4.7。在这些指标中，理性投资应用是因为我国的房价受地理位置、经济发展状况、城市规划和政府政策导向等多方面的影响，投资的热度也不尽相同。多元化开发指标是指利用大数据不仅可以进行住房供需分析、理性拿地，同时还可以使房地产开发企业业务范围多样化。在应用大数据进行创新性投资方面万科集团和花样年集团的经验值得分析。花样年集团拥有上千万的购房者数据信息，基于消费者的需求，通过手机 APP 的形式将商户和消费者联系起来，建立一个能够精准营销、便捷高效的“社区电子商务”平台。此外，花样年集团控股有限公司还构建了基于移动互联网的大数据业务布局，以金融服务、酒店服务、以及文化旅游等为主的领域。也极大的扩展了传统意义上的房企业务范围。同样，万科集团提高大数据的处理技术也为其带来了巨大的商机。万科集团经过处理 480 万业主的信息，结合社区商业、社区物流、社区医疗和养老以及业主的大数据信息，提出了构建“城

市配套服务商”的理念。营销应用能力是指根据庞大的数据来源准确的对客户进行定位和有效的营销，首先，房地产企业在建立客户基本系统的基础上对客户进行分类，通过大数据挖掘，然后提炼出有用的信息，实现有针对性的精确营销。表 4.7 呈现了相关企业大数据应用覆盖范围评价指标。

表 4.7 大数据应用覆盖范围评价指标

M06 大数据应用覆盖范围评价指标		参考文献
B06.1	理性投资应用能力，各地受不同因素影响导致投资的热度不尽相同	Yan W、Wanli Y、Juan Y、 白山、程明书、彭晓愈、 肖剑
B06.2	多元化开发应用能力，房地产企业业务范围多元化	
B06.3	创新性投资应用能力，扩展传统意义上的房地产企业业务范围	
B06.4	精准营销应用能力，服务模式代替传统的开发—建设—销售模式	

组织评价指标

人们为了实现预先设定的目标，相互协作结合而形成的集团或团体叫做组织，因此，从某种意义上说，组织能力的好坏直接关系到大数据应用能否实施，本文在总结了前人的研究成果和结合实际情况后，归纳整理，将组织评价指标主要分为以下 4 个二级指标，分别是企业大数据应用管理的权属、大数据应用语言的普及、企业大数据应用构架的监督、审查和改进以及大数据应用融入业务流程。（见表 4.8）

表 4.8 组织评价指标

M07 组织评价指标		参考文献
B07.1	企业大数据应用管理的权属，应该有专门的部门来统一管理大数据应用	David Di Its、K.-F.Pun、 K.-S.Chin 、 Hakyeon Lee、Chulhyun Kim
B07.2	大数据应用语言的普及，在组织内使用大数据语言的频率及频数	
B07.3	企业大数据应用构架的监督、审查和改进	
B07.4	大数据应用融入业务流程，具体的应用模块植入具体的业务	

应用风险评价指标

房地产业具有两面性，它既是实体经济的支柱，但同时又具有一些虚拟经济的特点，比如复杂性、亚稳态和高风险等。这些特点提高了大数据应用的不确定性。由于虚拟经济体系对预期高度敏感，大数据的公开性和共享性很容易影响人们对房地产的预期，投资需求会更加不稳定。房地产虚拟经济的亚稳态一旦受到冲击也会破坏房地产业的稳定性，影响国民经济发展。基于对 Siwei C (2009)，Zuoming L(2009)，Lingling Zhang(2014)，Huijia Li(2014)，Seokhoon Kim(2016)^[68]，杜丹阳(2014)，李爱华(2014)等的学术研究成果，加上实践调研总结。本文设置了以下 7 个大数据应用实践来诠释风险管理与控制应用能力，具体见表 4.9。

表 4.9 应用风险评价指标

M08 应用风险评价指标		参考文献
B08.1	政策性风险，政府宏观调控力度	Siwei C(2009)、 Zuoming L(2009) 、 Lingling Zhang(2014)、 Huijia Li (2014)、 Seokhoon Kim(2016)、 杜 丹 阳 (2014)、 李爱华(2014)
B08.2	财务风险，财务投入比例以及收益状况	
B08.3	信息识别风险，大数据在识别过程中克服的障碍	
B08.4	个人信息数据模糊化处理水平	
B08.5	数据所有权的保留度	
B08.6	数据潜在价值保留度	
B08.7	利润风险，达到预计利润的风险大小	

4.4 房地产开发企业大数据应用能力成熟度评价指标体系的确定

本文的问卷调查涉及两个主要研究内容，一是通过调查对根据文献初步和专家调查访谈梳理的指标进行选择以确定大数据应用能力成熟度评价指标体系，具体见附录 A。二是根据问卷调查确定案例企业的大数据应用能力成熟度的指标权重和成熟度评价指标情况。

本节的主要任务是设计本次指标体系构建的调查问卷，在现有相关文献资料中的指标和通过专家访谈调研得出的指标基础上设计的。本次调查问卷是为了确定大数据应用能力评价指标体系中的各项指标。在房地产开发企业评价活动中不仅具有理论意义，同时还具有现实意义，这些指标是房地产开发企业大数据应用中需要考虑的因素。

调查问卷设计

根据本次问卷调查的目的和上节所识别的评价指标体系，尽可能的简化问卷调查中的题目，使被调查者更容易理解指标问题，从而提高问卷调查的有效性。另外保证本次调查仅用于学术交流，绝不涉及学术以外的商业用途^[69]。

第一部分，调查对象基本信息，这部分题目主要由被调查者的性别、职务/职称、工作年限、被调查企业的规模等题目组成。这部分题目均为单项选择题，被调查者根据各个公司和个人情况如实填写即可。

第二部分，调查每个指标与上一级指标的关联度，指标即为上节识别出的所有指标。题目包括大数据应用一级指标与总目标的关联度，二级指标与其对应的一级指标的关联度。这部分的评价指标关联度调查均采用 Likert 学者的五点式评量制度，具体含义为，“1”表示极其相关，“2”表示较强相关，“3”表示相关，“4”表示较弱相关，“5”表示弱相关。

设计问卷的修改

为了对设计的调查问卷进行进一步的完善与修改，增强问卷的可读性、全面

性以及准确度，同时使被调查者能够便于正确理解调查的目的与意义。

经过问卷调查分析，发现问卷中有如下问题：

1) 在第二题中对调查者的年龄进行了区分，但经过分析可以发现，对调查者的年龄进行区分没有明显的意义，因为年龄不能代表调查者的实际工作经验，也不能代表其在房地产开发企业对大数据应用的掌握情况，所以决定将第二题改为“您从事与房地产开发相关工作的年限”。

2) 经过分析，发现问卷中的提问方式并不规范。在最后的问卷中，考虑到被调查者均是从事与房地产开发相关工作有一定年限的专业人员，并为了表达对他们们的配合调查问卷工作的感谢与尊敬，于是将所有问卷中的提问方式均改为“请问您”。

正式调查

此次成熟度评价的目的是要从读者和从事与房地产开发相关工作的员工角度了解房地产开发企业大数据应用能力水平，为了确保本次问卷调查的有效性，本次问卷由重庆、成都部分知名房地产开发企业的中层及以上管理人员和专家填写，这些被调查者均具有一定的大数据应用知识和房地产开发经验，其中还有一部分直接从事大数据应用的数据挖掘分析师，因而他们有能力对问卷调查设计的题目做出合理的判断，并根据行业和房企实际情况完成调查问卷。本次调查问卷主要采用纸质问卷和电子邮件相结合的方式发放。本次调查的重要目的就是获取有效数据，获取数据的质量将会影响研究结果的有效性和可靠性^[70]。本次调查总共向从业人员发放纸质问卷和随机发放电子邮件200份，有173份完成的问卷被返回，有效问卷160份。回复率为86.5%，有效率为80%，包括34份来自专业学者，126份来自行业从业人员，绝大部分调查对象拥有丰富的工作或研究经验，因此可以保证回复的质量。

通过对回收的有效问卷进行分析，笔者将有效问卷的性别比例、从事与房地产开发相关工作的年限比例以及被调查者的职务/职称比例统计如表4.10。

表 4.10 样本基本情况统计

性别比例			工作年限			职务/职称		
	频次	百分比		频次	百分比		频次	百分比
男	84	52.50%	5年以下	42	26.25%	教授/研究员级工程师	6	3.75%
女	76	47.50%	5至10年	73	45.63%	副高级工程师	30	18.75%
			11至15年	38	23.75%	中级工程师	59	36.88%
			15年以上	7	4.38%	初级/助理工程师	42	26.25%
						其他	23	14.38%

如表 4.10 所示，一半以上的调查对象拥有超过 10 年的工作或研究经验，因此可以保证回复的质量。此外，160 位行业调查对象中，有 6 位教授/研究员级工程师，30 位副高级工程师，59 位中级工程师，42 位初级/助理工程师和 23 位其他人员。他们所占全部问卷调查的比例依次为 3.75%、18.75%、36.88%、26.25%和 14.38%。该表的人员比例与目前我国房地产开发企业大数据应用和普及程度大致相吻合，因此，表中的数据能够反映房地产开发企业大数据应用能力的评价。

图 2 说明了研究框架。具有重大意义的房地产企业大数据应用成熟度标准以及具有显著相关度的房企大数据应用最佳实践已包含在建议模型中，而那些不重要的标准和实践已经排除。

将每个成熟度标准与房地产开发企业大数据应用成熟度评价总目标的相关程度与其对应成熟度标准的关联程度分为“非常相关”、“相关”、“较相关”、“一般”和“弱相关”五个等级。在向上述的调查专家人员发放问卷时，建议他们尽量区分同一级的指标相对于上一级指标的关联程度，做到准确分明。任何混淆不清、模棱两可的选择都会对评价结果造成影响^[71]。最后，通过回收的有效调查问卷，归纳整理出房地产开发企业大数据应用能力成熟度评价指标体系关联密切度数据如表 4.11 和 4.12 所示。表中对应的数值为选择该关联度的人数。

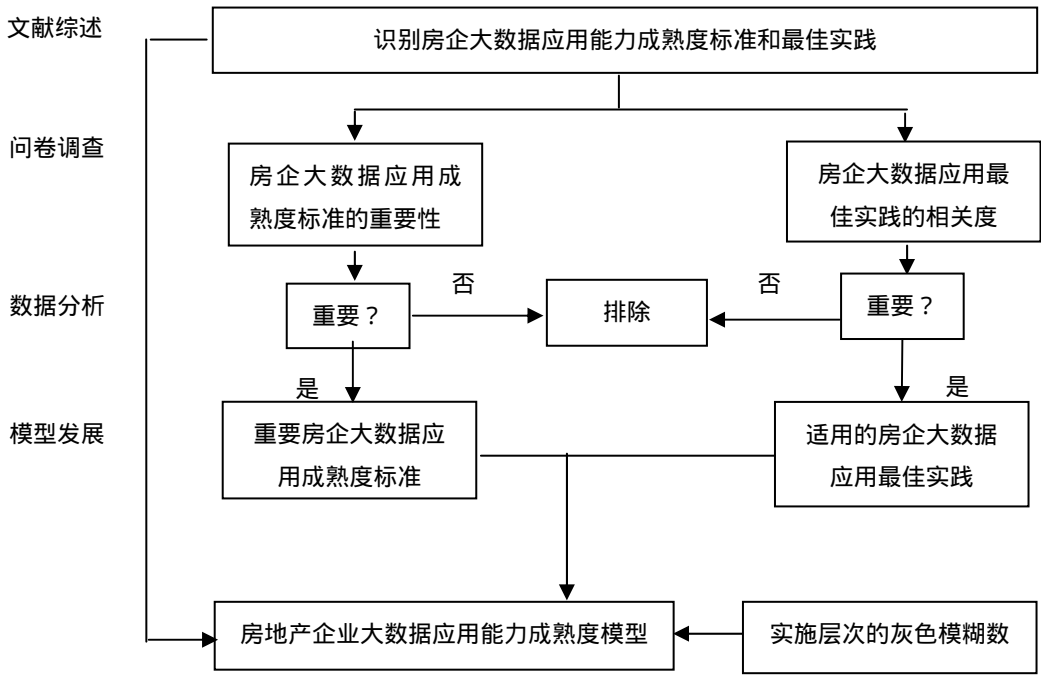


图 4-2 研究框架图

分析处理数据、关联度评价

根据表 4.11 和 4.12 中的数据引进系统工程原理，按照下面介绍的步骤依次对大数据应用能力指标进行关联度评价。然后，根据评价结果选择出相对有效的评价指标体系。

1) 确定评价集 A

将每个一级指标与房地产开发企业大数据应用能力成熟度评价总目标和二级指标与其对应的一级指标的关联度分为五个等级，分别是“非常相关”、“相关”、“较相关”、“一般”和“弱相关”。将“极其相关”评价值设为 1，“弱相关”评价值设为 0，其余关联度等级介于两者之间。那么根据系统工程原理将相关度评价集设为 $A = (a_1, a_2, a_3, a_4, a_5) = (1, 0.75, 0.5, 0.25, 0)$ 。

2) 确定一、二级指标评价矩阵 D 和 D_i

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} \\ d_{41} & d_{42} & d_{43} & d_{44} & d_{45} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} \\ d_{71} & d_{72} & d_{73} & d_{74} & d_{75} \\ d_{81} & d_{82} & d_{83} & d_{84} & d_{85} \end{bmatrix} = \begin{bmatrix} 104/160 & 46/160 & 5/160 & 2/160 & 3/160 \\ 134/160 & 12/160 & 7/160 & 5/160 & 2/160 \\ 1/160 & 11/160 & 31/160 & 114/160 & 3/160 \\ 109/160 & 37/160 & 8/160 & 5/160 & 1/160 \\ 4/160 & 29/160 & 81/160 & 39/160 & 7/160 \\ 75/160 & 67/160 & 11/160 & 3/160 & 4/160 \\ 5/160 & 38/160 & 57/160 & 51/160 & 9/160 \\ 3/160 & 23/160 & 12/160 & 116/160 & 6/160 \end{bmatrix} = \begin{bmatrix} 0.650 & 0.288 & 0.031 & 0.013 & 0.019 \\ 0.838 & 0.075 & 0.044 & 0.031 & 0.013 \\ 0.006 & 0.069 & 0.194 & 0.713 & 0.019 \\ 0.681 & 0.231 & 0.050 & 0.031 & 0.006 \\ 0.025 & 0.181 & 0.506 & 0.244 & 0.044 \\ 0.469 & 0.419 & 0.069 & 0.019 & 0.025 \\ 0.031 & 0.238 & 0.356 & 0.319 & 0.056 \\ 0.019 & 0.144 & 0.075 & 0.725 & 0.038 \end{bmatrix}$$

$$D_1 = \begin{bmatrix} d_{111} & d_{112} & d_{113} & d_{114} & d_{115} \\ d_{121} & d_{122} & d_{123} & d_{124} & d_{125} \\ d_{131} & d_{132} & d_{133} & d_{134} & d_{135} \\ d_{141} & d_{142} & d_{143} & d_{144} & d_{145} \\ d_{151} & d_{152} & d_{153} & d_{154} & d_{155} \\ d_{161} & d_{162} & d_{163} & d_{164} & d_{165} \end{bmatrix} = \begin{bmatrix} 0.075 & 0.881 & 0.025 & 0.013 & 0.006 \\ 0.288 & 0.581 & 0.094 & 0.019 & 0.019 \\ 0.019 & 0.169 & 0.506 & 0.256 & 0.050 \\ 0.281 & 0.600 & 0.038 & 0.056 & 0.025 \\ 0.044 & 0.213 & 0.356 & 0.275 & 0.113 \\ 0 & 0.113 & 0.081 & 0.750 & 0.056 \end{bmatrix}$$

以上矩阵中，我们将其中 $d_{ik} = \frac{r_{ik}}{r_n}$ ， r_{ik} 表示选择第 i 个一级指标的第 k 个评价值的
人 数， r_n 为参与此次评价的总人数。 $d_{ijk} = \frac{r_{ijk}}{r_n}$ ， r_{ijk} 表示选择第 i 个一级指标下 第
 j 个二级指标的第 k 个评价值的人数。同理可以求得其他大数据应用实践的评价矩
阵 $D_2, D_3, D_4, D_5, D_6, D_7, D_8$ 。每个指标的专家统计见表 4.11 和 4.12。

3) 确定关联度评价 N 和 N_i

$$N = D \circ A^T = \begin{bmatrix} 0.650 & 0.288 & 0.031 & 0.013 & 0.019 \\ 0.838 & 0.075 & 0.044 & 0.031 & 0.013 \\ 0.006 & 0.069 & 0.194 & 0.713 & 0.019 \\ 0.681 & 0.231 & 0.050 & 0.031 & 0.006 \\ 0.025 & 0.181 & 0.506 & 0.244 & 0.044 \\ 0.469 & 0.419 & 0.069 & 0.019 & 0.025 \\ 0.031 & 0.238 & 0.356 & 0.319 & 0.056 \\ 0.019 & 0.144 & 0.075 & 0.725 & 0.038 \end{bmatrix} \circ \begin{bmatrix} 1 \\ 0.75 \\ 0.5 \\ 0.25 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.884 \\ 0.923 \\ 0.333 \\ 0.888 \\ 0.475 \\ 0.822 \\ 0.467 \\ 0.345 \end{bmatrix}^T$$

$$N_1 = D_1 \circ A^T = \begin{bmatrix} 0.075 & 0.881 & 0.025 & 0.013 & 0.006 \\ 0.288 & 0.581 & 0.094 & 0.019 & 0.019 \\ 0.019 & 0.169 & 0.506 & 0.256 & 0.050 \\ 0.281 & 0.600 & 0.038 & 0.056 & 0.025 \\ 0.044 & 0.213 & 0.356 & 0.275 & 0.113 \\ 0 & 0.113 & 0.081 & 0.750 & 0.056 \end{bmatrix} \circ \begin{bmatrix} 1 \\ 0.75 \\ 0.5 \\ 0.25 \\ 0 \end{bmatrix} \\ = [0.752 \quad 0.775 \quad 0.463 \quad 0.764 \quad 0.450 \quad 0.313]$$

从关联度 N 和 N_1 可知，成熟度标准中的市场评价指标、应用环境评价指标、组织评价指标和应用风险评价指标的关联度都小于 0.5，低于较相关标准，说明关联度不高，故舍去。那么，一级指标简化为企业战略影响指标、资源评价指标、大数据分析/处理技术评价指标和应用覆盖范围评价指标。其中，企业战略影响指标中的树立大数据应用意识文化、大数据信息沟通能力的提升和与第三方数据共享机制的关联度均小于 0.5，舍去。同理可以求得其三个一级指标的大数据应用实践相关度评价 N_2 、 N_3 、 N_4 （见表 4.11 和 4.12）。

表 4.11 房地产开发企业大数据应用能力评价指标体系识别统计表

总目标	一级指标	成熟度标准与总目标的关系					评价值	取舍
		非常相关	相关	较相关	一般	弱相关		
房地产开发企业大数据应用能力评价	企业战略影响指标	104	46	5	2	3	0.884	留
	资源评价指标	134	12	7	5	2	0.923	留
	市场评价指标	1	11	31	114	3	0.333	舍
	大数据分析、处理技术评价指标	109	37	8	5	1	0.888	留
	应用环境评价指标	4	29	81	39	7	0.475	舍
	应用覆盖范围评价指标	75	67	11	3	4	0.822	留
	组织评价指标	5	38	57	51	9	0.467	舍
	应用风险评价指标	3	23	12	116	6	0.345	舍

表 4.12 房地产开发企业大数据应用能力成熟度评价指标体系识别统计表

房地产开发企业大数据应用能力评价指标体系		二级指标与一级指标的关系					评价值	取舍
一级指标	二级指标	非常相关	相关	较相关	一般	弱相关		
企业战略影响指标	管理层的支持	12	141	4	2	1	0.752	留
	员工培训计划	46	93	15	3	3	0.775	留
	大数据信息沟通能力的提升	3	27	81	41	8	0.463	舍
	建立房地产大数据共享平台	45	96	6	9	4	0.764	留
	树立大数据应用意识文化	7	34	57	44	18	0.450	舍
	与第三方数据共享机制	0	18	13	120	9	0.313	舍
资源评价指标	人力资源的优化	61	88	9	0	2	0.822	留
	财务资源的利用	4	12	114	25	5	0.477	舍
	内外资源的迅捷调配	8	15	113	18	6	0.502	舍
	相关经验的借鉴	9	11	24	109	7	0.353	舍
	土地资源数据的挖掘	103	42	8	7	0	0.877	留
大数据分析、处理技术评价指标	评估项目投资价值	36	119	4	0	1	0.795	留
	客户数据搜集与存储水平	89	65	4	2	0	0.877	留
	数据挖掘能力	2	14	25	116	3	0.338	舍
	非结构化数据提炼技术	0	32	84	44	0	0.481	舍
	机器学习能力	2	20	16	118	4	0.341	舍
	住房供需分析能力	67	84	3	1	5	0.823	留
	自然语言处理	0	12	32	116	0	0.338	舍
	沉没数据激活能力	3	32	69	43	13	0.452	舍
应用覆盖范围评价指标	理性投资应用能力	46	80	29	5	0	0.761	留
	多元化开发应用能力	17	127	12	3	1	0.744	留
	创新性投资应用能力	5	11	113	24	7	0.473	舍
	精准营销应用能力	27	85	40	4	4	0.698	留

4) 确定评价指标体系

根据前面确定的评价集及表 3.11 和 3.12 中各一级指标、二级指标的评价值，把关联度评价值低于 0.5（关联度为较相关）的所有指标舍掉，即删除关联度评价值小于 0.5 的市场评价指标、应用环境评价指标、组织评价指标和应用风险评价指标 4 个一级指标和大数据信息沟通能力的提升、树立大数据应用意识文化、与第三方数据共享机制、财务资源的优化、内外资源的迅捷调配、相关经验的借鉴、数据挖掘能力、非结构化数据提炼技术、机器学习能力、自然语言处理等二级指标。其余指标的关联度评价值均大于 0.5，说明关联度为较相关及以上，指标关联度较高，均保留有效。最后，得出整个房地产开发企业大数据应用能力成熟度指标体系包括 4 个一级指标和 11 个二级，具体如图 4-3 所示。

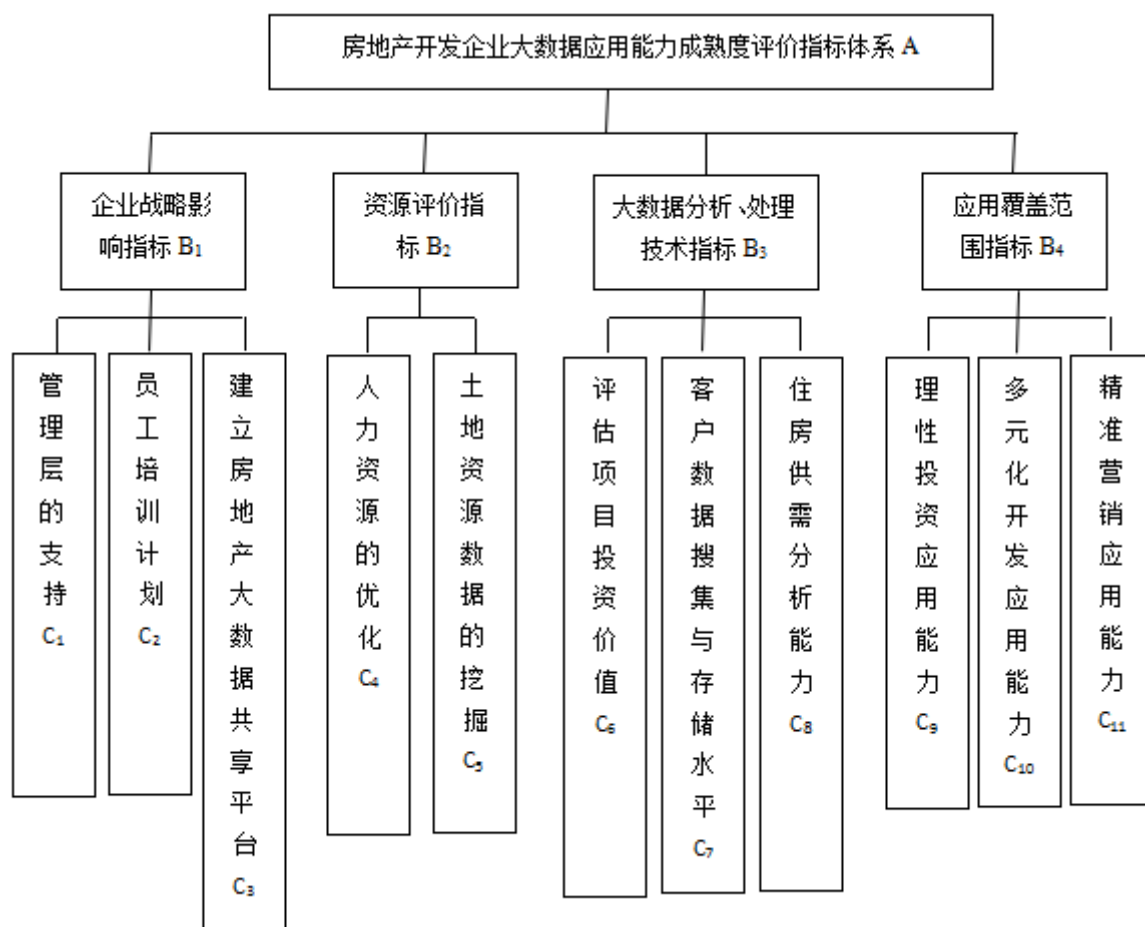


图 4-3 房地产项开发企业大数据应用能力成熟度评价指标体系

对于为何删除掉指标中的组织评价和应用风险评价等指标作出以下尝试性解释：

相对而言，组织评价指标在大型、知名房地产公司应用大数据时会更加重要，因为他们掌握的数据量要远远大于其他房地产开发企业，他们捕捉市场动态的能力也在很大程度上强于业界同行，怎样有针对性地组织利用好掌握的资源就显得尤为重要；而在普通的房地产开发企业中，组织评价能力固然重要，但是可以在其他指标中综合考虑。

在保留下来的企业战略影响指标、资源评价指标、大数据分析/处理技术指标和应用覆盖范围评价指标中已经包含了应用风险评价指标中的重点指标。

4.5 本章小节

本章首先进行了成熟度模型在房地产开发企业大数据应用能力评价研究中的适用性分析，得出了 CMM 适用于房地产开发企业大数据应用的能力等级；其次，

通过总结相关文献和问卷调查构建了房地产开发企业大数据应用能力成熟度评价的初步指标体系；最后，应用系统工程原理对初步建立的指标进行了优化，最后得到了 4 个成熟度标准，11 个大数据应用实践，据此建立起房地产开发企业大数据应用能力成熟度评价体系。

第五章 房地产开发企业大数据应用能力的评价方法

在第四章中，我们构建了房地产开发企业大数据应用能力成熟度评价指标体系，应该运用相应的方法确定评价指标的权重，以便于后面的评价。构建出的房地产开发企业大数据应用能力成熟度评价指标大部分是定性指标，所以应该依据每个指标的不同特点，选择简单、适用且可靠的方法对指标进行量化分析。在确定了各个指标的权重后，还应该运用相应的评价方法作为房地产开发企业大数据应用能力等级的量化工具，选择不同的评价方法，对产生的结果也有不同的意义。因此，在现实情况中应该针对不同的评价目标和评价过程选择适当的评价方法。

5.1 几种常用评价方法介绍

在第四章构建的指标体系中包含多个评价指标，因此对房地产开发企业大数据应用能力等级的评价是关于多目标的评价。阅读国内外大量研究文献我们发现关于多目标的评价方法不胜枚举，不同的评价方法也具有各自的优缺点，本节将针对几种常用评价方法的特点做出分析和总结，从而选择出最适合房地产开发企业大数据应用能力评价的方法。

层析分析法（AHP）

该方法的基本思路是按照具有递阶结构的目标和子目标准则及约束条件来建立评价的方案，通过运用两两元素间相比较的方法来确定判断矩阵，接着把判断矩阵中与最大特征相对应的特征分量作为相应系数，然后综合计算出各个方案的权重优先程度。该方法优点是清晰的量化和系统化房地产开发企业的大数据应用能力，并且容易计算。而劣势是当房企大数据应用指标很多并且框架庞大时，会出现很多意想不到的问题。

多目标数学规划法

该方法论述过程缜密，具有较强的说服力，容易让人信服，但运用该方法建立房地产开发企业大数据应用能力成熟度模型的步骤太繁琐太复杂，大数据应用的多指标因素很难用该方法解决。

模糊综合评价法

模糊综合评价法是一个综合考虑各种模糊因素有关的事物，应用模糊变换原理和最大隶属度原则，量化一些边界不清、不易定量的因素再对其进行综合评价的一种方法。但它对房企大数据应用评价指标的量化较少、定性成分多，不易令人信服，当指标过多数据统计量大时，难以确定权重。

④ 数据包络分析法（DEA）

建立在相对概念为基础上的 DEA 法，是一种以凸分析和线性规划为工具的评价方法，应用数学规划模型计算出比较决策单元之间的相对效率，对评价对象做出评价能够充分考虑对于决策单元本身最优的投入产出方案，不仅能够更理想地反映评价自身的信息和特点，同时对于评价复杂系统的多投入多产出分析也具有独到之处。数据包络分析法衡量的生产函数边界是确定的，所有随机干扰项则是效率因素。同时，极值也会影响到该方法的评价。

BP 神经网络法

该方法自身带有联想—记忆及非线性映射的特点，在实际运用时仅仅对神经的评价模型进行的模拟训练，可以在各个神经元的连接权重中反映出此类型的问题特点，然后输入评价对象的相关数据客观的得到评价结果。但在相对比较复杂且大数据应用评价因素众多的情况下，运用该方法不易操作成功。

作业成本法

该方法仅仅是在分析了房地产开发企业参与大数据应用而发生的直接成本与间接成本之间进行比较来衡量大数据的应用能力，运用此类方法考虑的因素太过于单一，不能真正评价出房地产开发企业的大数据应用能力。

灰色综合评价法

灰色综合评估法是一种以灰色关联分析理论为指导，基于专家评判的综合性评估方法。其过程是首先建立灰色综合评估模型，再对各种评价因素进行权重选择，最后进行综合评估。其中，灰色综合评估法中的权重选择可以结合层次分析法，以提高评估的准确性。

将上述 7 种常用评价方法经过仔细对比分析以后，我们发现其中每种方法都有自身的一些局限性，同时侧重的使用领域也不尽相同。综合比较上述几种方法，

相对而言多目标数学规划法建立模型的步骤太繁琐太复杂，涉及房地产开发企业大数据应用能力成熟度评价等多指标因素的问题时很难用该方法解决，模糊综合评价法可以量化房地产开发企业大数据应用能力成熟度评价体系中一些边界不清、不易定量的指标进行综合评价。但它定量数据较少、定性成分多，不易令人信服，同时当涉及指标过多数据统计量大时，难以确定权重。DEA 对异常值相当敏感，在实际生活中，由于统计数据质量、测量误差等问题，构成数据包络曲线的那些点是非常敏感的，或者说，其它效率不是最优的点都是和数据包络曲线上最好的点相比，而这些点其实是不稳定的，在此基础上得出的处理结果也是不稳定的。BP 神经网络法比较客观，只要确保数据来源的客观、可靠以及全面，通过该方法得出的评价结果就具有很强的说服力，但针对本文评价指标因素较多，评价体系也较为复杂的情况，此种方法达不到理想的效果。作业成本法仅仅是在分析了房地产开发企业参与大数据应用而发生的直接成本与间接成本之间进行比

较来衡量大数据的应用能力，运用此类方法考虑的因素太过于单一，不能真正评价出房地产开发企业的大数据应用能力。目前应用的最多的评价方法是层次分析法，但是当因素较多单独使用层次分析法时给出的结果往往是不全面和不准确的，且容易出错，工作量也比较大。而灰色理论认为现实世界并不是清清楚楚的白色系统，也不是一无所知的黑色系统，而是略知一二的灰色系统。以往用白色的思想处理问题，要找到因素间明确的映射关系，而且要求样本有大容量。但事实上，实践工作中大部分的问题的各个因素都不能建立非常明确映射关系。因此灰色系统理论则主要研究“小样本、贫信息、信息不确定问题”，这些在白色系统无法解决的。因此综合考虑以上各种因素后，结合本文的研究内容选择适用于问题较复杂且指标因素较多的评价方法—灰色综合评价法，同时灰色综合评估法中的权重选择可以结合层次分析法，以提高评估的准确性。

5.2 应用灰色综合评价法进行评价的适用性分析

灰色系统理论认为，尽管系统的信息不像白色系统那样充分，但作为系统必然是有特定功能和有序的，存在某些外露或内在的规律。对于随机量、无规则的干扰成分、杂乱无章的数据序列等，从灰色系统理论的观点看，并不认为是不可捉摸的，或者说只有依据过去大量的数据才能整理出一个经验性的统计规律。相反的，灰色系统理论将随机量看作是在一定范围内变化的灰色量，通过数据的处理，将灰色数变换为生成数，从生成数得到规律性较强的生成函数。这个过程不必依据先验数据的统计规律，或者说灰色系统理论的量化基础不是原始数据，而是生成数，不是经验性的统计规律，而是现实性的生成规律。这样就可使灰色系统理论突破了概率统计的局限性，生成数成了灰色系统理论的奠基石。灰色系统理论提出了动态关联度的概念以及量化的关系式，避免了数理统计回归分析的缺点，揭示了事物动态关联的特征与程度。灰数是灰色系统的基本“单元”或“细胞”，实际上指在某一个区间或某个一般的数集内取值的不确定数，灰色理论中把只知道大概范围而不知其确切值的数称为灰数。“信息不完全”就是灰关系的表述，一般包括以下几个方面：一是信息系统不完全明确；二是因素关系不完全清楚；三是系统结构不完全知道；四是系统的作用原理不完全明了。灰色综合评判是在一般的综合评判模型与灰色模型相结合的情况下进行的评判，其主要思想是根据评判对象的特性，求出评判矩阵，根据评价方程式计算评价值。灰色理论从提出至今已经在国内外产生了极大的影响，得到了广泛的认可和研究。世界上许多国家都有学者对其进行研究和拓展，其应用范围也遍及农业、工业、能源、交通、石油、地质、气象、水文、生态、环境、医学、军事、经济、社会等众多科学领

域。在灰色系统理论的基础上，还衍生出了一系列如：运用灰色系统模型进行水文预报、灰色聚类法在地质学中的应用、灰色系统理论与作物育种学相结合、灰色理论在医学中的应用、灰理论与机电控制的结合研究、灰色理论与交通系统结合和灰色混沌神经网络的灰色混沌理论、“区域 经济灰色系统分析 ”等一批新兴的交叉学科。

之所以可以运用灰色理论解决成熟度评价问题，原因在于参与评价的人员组成是有限的，同时虽然评价模型已经详细给出评价指标，但指标并不是绝对完全的，造成能够获得的信息量有限。再者由于评判人员的个性差异及知识结构的不同，认识问题的侧重方面不同，以及判断问题时存在受到其他因素干扰而出现偏差的可能性等原因，所以对于定性指标的量化判断存在不完全肯定性。正是由于成熟度评价也同样面临着信息不完备、不确定，数据较少的灰色系统的难题，因此将灰色评判理论引入到成熟度评价中，能同时解决评判过程的信息量有限，以及判定不完全肯定的现象。

在房地产开发企业大数据应用能力评价研究中，涉及到种类繁多层次繁杂的指标因素，必须要依据多个指标才能对目标做出评价，而不能单独依据某些指标的好坏简单的进行评判。且房地产开发企业大数据应用能力的影响因素并不相同，有些指标可以运用统计的方法求出它的概率分布，但是绝大多数的指标概念自身都不是很明确，这样一来用一般的数学方法就难以进行定量描述，此时就只有根据专家的专业知识和技能经验通过运用灰色系统理论来描述出它们所具有的性质以及可能产生的结果。

灰色综合评价法正好是一种关系到部分信息明确、部分信息不明确的且比较系统化的综合评价方法，可以按照以下步骤进行： 建立灰色综合评价模型；对各种评价因素进行权重选择； 最后进行综合评价。

该方法的主要优势在于： 容易构建较为简单的数学模型并且掌握起来比较方便； 对那些指标因素较多且层次结果较为繁杂的问题综合评价的效果较好；易于定量较为主观性的经验。

5.3 基于灰色综合评价法的评价步骤

灰色综合评价法是以灰色关联分析理论为指导，在专家评判的基础上进行的综合性评估方法。该方法可以把定性评价定量化，最后对受到多个指标因素影响的对象给出一个综合性的评价。该方法系统性强并且结果清晰明了，可以容易的将那些不明确且难以量化的问题一一解决，同时也适用于解决各种不确定性的问题。

5.3.1 权重计算及权重集的确定

计算所有评价指标的权重值

通过向 5 位相关的房地产开发企业大数据应用专家进行问卷打分，然后回收有效问卷并对其进行处理后获得类别指标层的权重统计表，通过计算平均值得出所有指标因素的权重值。

关键过程域指标层的权重计算

AHP 可以针对每一个一级指标下的二级指标进行审核，并且对每个评价指标依次进行两两间的对比，得出的结果即为最终的判断矩阵。表 5.1 和 5.2 分别为判断尺度表和判断矩阵。

表 5.1 判断尺度表

判断尺度	定义
1	A 和 B 相同重要
3	A 比 B 较为重要
5	A 比 B 重要
7	A 比 B 非常重要
9	A 比 B 绝对重要
2、4、6、8	介于上述两个相邻判断尺度中间
倒数	A 比 B 的重要性比为 a，则 B 比 A 的重要性比为 1/a

判断矩阵：

表 5.2 判断矩阵

D	X ₁	X ₂	...	X _n
X ₁	a ₁₁	a ₁₂	...	a _{1n}
X ₂	a ₂₁	a ₂₂	...	a _{2n}
...
X _n	X _{n1}	X _{n2}	...	X _{nn}

其他依次类推。

计算得出判断矩阵后采用方根法计算各个指标的权重值；

1) 将判断矩阵 A 中的每一行指标依次相乘后，计算它的 n 次方根，即：

$$\omega_i^{\bullet} = \left(\prod_{j=1}^n a_{ij} \right)^{\frac{1}{n}}, i = 1, 2, 3, \dots, n \quad (5.1)$$

2) 对矩阵中所有元素进行归一化处理,即：

$$\omega_i = \frac{\omega_i}{\sum_{j=1}^n \omega_j} \quad (5.2)$$

对计算结果进行一致性检验：

当且仅当判断矩阵完全一致时，判断矩阵 A 会有 $\lambda_{\max}=n$ ，若判断矩阵不完全一致时有 $\lambda_{\max}>n$ ，所以可以用 $\lambda_{\max}-n$ 的差值来验证判断矩阵 A 的一致性程度，一般而言，一致性的指标是用 C.I.值来计算的，而这个值越小，表明一致性程度就越高。

$$C.I. = \frac{\lambda_{\max} - n}{n - 1} \quad (5.3)$$

与此同时，在一些情况下 C.I.的偏差是由于随机因素产生的。因此，在验证判断矩阵 A 的一致性程度时，需要将 C.I.与 R.I.的比值进行比较，得出检验数 C.R.即：

$$C.R. = \frac{C.I.}{R.I.} \quad (5.4)$$

接下来开始进行矩阵 A 的一致性判断，若 $C.R.<0.1$ ，可以认为判断矩阵 A 中各个评价指标基本一致，各个评价指标的权数即为上述公式的计算结果；但若 $C.R.>0.1$ ，则说明矩阵 A 中各个指标不满足一致性要求，应该重新评估。表 5.3 为平均随机一致性指标。

表 5.3 平均随机一致性指标

n	1	2	3	4	5	6	7	8	9	10
R.I.	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49

5.3.2 灰色评价的过程

制定评分标准

“对过去/现状的把握”、“发现模式”、“预测”、“优化”四个等级的赋分区间分别是[1, 2], [2, 3], [3, 4], [4, 5]，如果分数不在赋分区间里则说明成熟度不满足任何一个等级。

组织评审

请以上专家对问卷进行选择，然后根据评分说明汇总分数，最后确定评价样本矩阵。表 5.1 中 C_{nij} 表示已经根据评分准则将分数汇总的二级指标评估值，n 表示评审人员的个数， $n \in [1, n]$ ，n 表示一级指标个数，j 表示二级指标个数。

5.1 评分统计表

一级指标	权重	二级指标	权重	评分					
				评审 1	评审 2	...	评审 n	...	评审 N
U ₁	W ₁	U ₁₁	W ₁₁	C ₁₁₁	C ₂₁₁	...	C _{n11}	...	C _{N11}
		U ₁₂	W ₁₂	C ₁₁₂	C ₂₁₂	...	C _{n12}	...	C _{N12}
		U ₁₃	W ₁₃	C ₁₁₃	C ₂₁₃	...	C _{n13}	...	C _{N13}
U ₂	W ₂	U ₂₁	W ₂₁	C ₁₂₁	C ₂₂₁	...	C _{n21}	...	C _{N21}
		U ₂₂	W ₂₂	C ₁₂₂	C ₂₂₂	...	C _{n22}	...	C _{N22}
U ₃	W ₃	U ₃₁	W ₃₁	C ₁₃₁	C ₂₃₁	...	C _{n31}	...	C _{N31}
		U ₃₂	W ₃₂	C ₁₃₂	C ₂₃₂	...	C _{n32}	...	C _{N32}
		U ₃₃	W ₃₃	C ₁₃₃	C ₂₃₃	...	C _{n33}	...	C _{N33}
U ₄	W ₄	U ₄₁	W ₄₁	C ₁₄₁	C ₂₄₁	...	C _{n41}	...	C _{N41}
		U ₄₂	W ₄₂	C ₁₄₂	C ₂₄₂	...	C _{n42}	...	C _{N42}
		U ₄₃	W ₄₃	C ₁₄₃	C ₂₄₃	...	C _{n43}	...	C _{N43}

确定评价灰类

在灰色评价中要真实的反应出属于某类的程度，必须确定评价的灰类，灰数以及白化权函数。

1) 灰类，即按照评估要求所需要归纳的类别，本文将其分为“对过去/现状的把握”、“发现模式”、“预测”、“优化”四个灰类，用 e_k 表示 $k \in [1, s]$ 。

2) 灰类的灰数，即可判定隶属于某一灰类的数值区间以及阈值。根据灰类和相应的评分标准确定出灰数以及端点值。从“对过去/现状的把握”向“优化”共四个灰类。每一个等级的取值范围分别为 $[a_1, a_2], [a_2, a_3] \dots [a_{k-1}, a_k] \dots [a_s, a_{s+1}]$ ，在这里确定为 $[1, 2], [2, 3], [3, 4], [4, 5]$ ，由于使用端点三角函数，灰数的中间值是隶属度最高的值，因此每一等级阈值为 1, 2, 3, 4。

3) 白化权函数是指在直角坐标中的一条三折线，它可以定量地描述某一评估对象隶属于某个灰类的程度（称权函数）。白化权值就是评价对象属于某个灰类的隶属度值，完全属于则为 1，完全不属于则为 0，取值介于 0, 1 之间。

本文选择构造端点三角函数。

设 $(a_k + a_{k+1})/2$ 属于第 k 个灰类的隶属度为 1，在直角坐标系中将点 $((a_k + a_{k+1})/2, 1)$ 和第 $k-1$ 个灰类的起点 a_{k-1} 以及第 $k+1$ 个灰类的终点 a_{k+2} 连接起来，得到指标关于第 k 个灰类的三角隶属函数 $f_k(c_{nij})$ ，在三角白化权函数中，将取数域分别向左右两边延拓至 a_0 与 a_{s+2} 后得到函数 $f(\bullet)$ 和 $f_s(\bullet)$ 。在这里 a_0 到 a_{s+2} 依次为 0, 1, 2, ..., 5, 6。

灰类 k 的白化权函数为 $f_k(c_{nij})$ ，函数定义域范围以外值的白化权函数为 0，在下述公式中将定义域超过的范围省略了。 c_{nij} 是第 n 个评价者对指标 u_{ij} 的评价值。

下面是灰度 1-4 的白化权函数 f_k ：

$$f_1(c_{nij}) = \begin{cases} \frac{c_{nij}}{1.5} & c_{nij} \in [0, 1.5] \\ \frac{3-c_{nij}}{3-1.5} & c_{nij} \in [1.5, 3] \end{cases} \quad f_2(c_{nij}) = \begin{cases} \frac{c_{nij}-1}{1.5} & c_{nij} \in [1, 2.5] \\ \frac{4-c_{nij}}{4-2.5} & c_{nij} \in [2.5, 4] \end{cases}$$

$$f_3(c_{nij}) = \begin{cases} \frac{c_{nij}-2}{3.5-2} & c_{nij} \in [2, 3.5] \\ \frac{5-c_{nij}}{5-3.5} & c_{nij} \in [3.5, 5] \end{cases} \quad f_4(c_{nij}) = \begin{cases} \frac{c_{nij}-3}{4.5-3} & c_{nij} \in [3, 4.5] \\ \frac{6-c_{nij}}{6-4.5} & c_{nij} \in [4.5, 6] \end{cases}$$

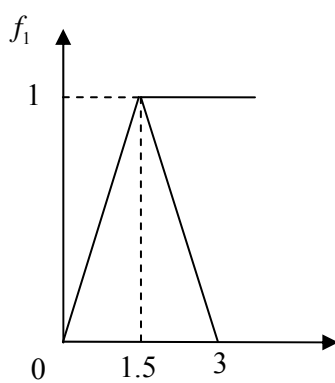


图 5-1

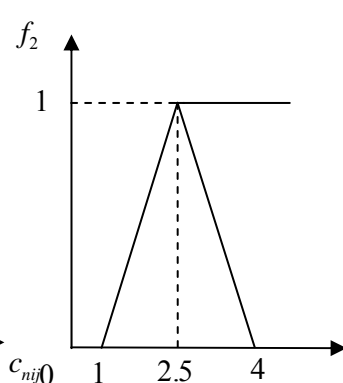


图 5-2

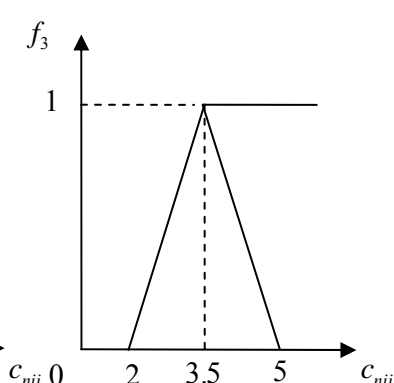


图 5-3

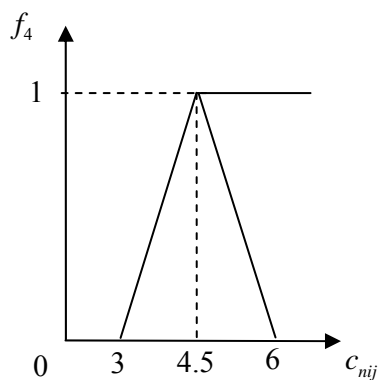


图 5-4

图 5-1 ~ 5-4 是 4 个灰类的白化权函数 f_k 在坐标轴上的表示。

4) 计算每个二级指标的灰色评价权值及判断等级

指标的灰色评价权值是综合各个评审所认为某个二级指标都有归属于某一灰度（等级）的隶属度。

$$\left. \begin{aligned} e_1 \quad x_{ij}^1 &= \sum_{n=1}^N f_1 \\ e_2 \quad x_{ij}^2 &= \sum_{n=1}^N f_2 \\ e_3 \quad x_{ij}^3 &= \sum_{n=1}^N f_3 \\ e_4 \quad x_{ij}^4 &= \sum_{n=1}^N f_4 \end{aligned} \right\} \quad (5.5)$$

每个二级指标的每个灰度都有其对应的白化权值。

u_{ij} 隶属于某个灰度的白化权值算法如式 (5.5) 所示。

每个二级指标 u_{ij} 的总白化权根据式 (5.6) 计算：

$$x_{ij} = \sum_{k=1}^4 x_{ij}^k \quad (5.6)$$

按式 (5.7) 计算 u_{ij} 评价权 (归一化)。

$$\left. \begin{aligned} e_1 \quad x_{ij}^{*1} &= \frac{x_{ij}^1}{x_{ij}} \\ e_2 \quad x_{ij}^{*2} &= \frac{x_{ij}^2}{x_{ij}} \\ e_3 \quad x_{ij}^{*3} &= \frac{x_{ij}^3}{x_{ij}} \\ e_4 \quad x_{ij}^{*4} &= \frac{x_{ij}^4}{x_{ij}} \end{aligned} \right\} \quad (5.7)$$

$$\text{且 } \sum_{e=1}^s x_{ij}^{*k} = 1$$

则 u_{ij} 对于各个灰类的灰色评价权向量 X_{ij}^* ：

$$X_{ij}^* = [x_{ij}^{*1} \quad x_{ij}^{*2} \quad x_{ij}^{*3} \quad x_{ij}^{*4}]$$

U_{ij} 等级判断。根据最大隶属度原则，认为 $\max x_{ij}^{*k}$ 所在的灰度 k 为该二级指标所属的等级。

5) 一级指标综合评价。

计算一级指标评价权，一级指标评价权向量为 X_{i0}^* 。

$$X_i^* = \begin{bmatrix} X_{i1}^* \\ X_{i2}^* \\ X_{i3}^* \\ \vdots \\ X_{ij}^* \end{bmatrix} = \begin{bmatrix} x_{i1}^{*1} & x_{i1}^{*2} & x_{i1}^{*3} & x_{i1}^{*4} & x_{i1}^{*5} \\ x_{i2}^{*1} & x_{i2}^{*2} & x_{i2}^{*3} & x_{i2}^{*4} & x_{i2}^{*5} \\ x_{i3}^{*1} & x_{i3}^{*2} & x_{i3}^{*3} & x_{i3}^{*4} & x_{i3}^{*5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{ij}^{*1} & x_{ij}^{*2} & x_{ij}^{*3} & x_{ij}^{*4} & x_{ij}^{*5} \end{bmatrix}$$

二级指标权向量为 W_{ij} :

$$W_1 = [w_{11}, w_{12}, w_{13}]^T$$

$$W_2 = [w_{21}, w_{22}]^T$$

$$W_3 = [w_{31}, w_{32}, w_{33}]^T$$

$$W_4 = [w_{41}, w_{42}, w_{43}]^T$$

U_1, U_2, U_3, U_4 一级指标综合评价权矩阵为 G_i

$$G_i = W_i^T \times X_i^* = [g_i^1 \quad g_i^2 \quad g_i^3 \quad g_i^4]$$

$$G = \begin{bmatrix} G1 \\ G2 \\ G3 \\ G4 \end{bmatrix} = \begin{bmatrix} g_1^1 & g_1^2 & g_1^3 & g_1^4 \\ g_2^1 & g_2^2 & g_2^3 & g_2^4 \\ g_3^1 & g_3^2 & g_3^3 & g_3^4 \\ g_4^1 & g_4^2 & g_4^3 & g_4^4 \end{bmatrix}$$

U_i 等级判断。根据最大隶属度原则, 认为 $\max g_i^k$ 所在的灰度 k 为该一级指标所属的等级。

6) 计算灰色综合评价值判断总目标等级

$$\text{总评价目标 } U \text{ 的权向量为 } W^T = [W_1 \quad W_2 \quad W_3 \quad W_4]$$

总评价目标 U 的评价权向量为 G^*

$$G^* = W^T \times G = [g^{*1} \quad g^{*2} \quad g^{*3} \quad g^{*4}]$$

灰类 $e_1 \sim e_5$ 的阈值依次为 1, 2, 3, 4。

$$\text{灰色综合评价值 } y = [1, 2, 3, 4] \times G^{*T}$$

再根据计算得到的灰色综合评价值, 带入公式求出相应的 $f_k(y)$ 值, 得到最终的灰色综合评价权向量 F 。

$$F = [f_1(y) f_2(y) f_3(y) f_4(y)]$$

根据最大隶属度原则, 判定 $\max f_k(y)$ 所在的灰度为最终的成熟度等级。

5.4 本章小节

本章首先介绍了几种常见的多目标评价方法, 通过对比分析各种方法之间的优缺点最终确定了灰色综合评价法; 其次, 鉴于房地产开发企业大数据应用能力

评价指标体系数量大且种类多的特点，本章详细叙述了层次分析法和灰色综合评价法具体运用的过程。

第六章 案例分析

6.1 某房地产开发企业的现状

国内某知名房地产开发企业是一家大型的上市公司，经过长期的发展，在全国许多城市都成立了分公司。所以，该企业的客户数量十分庞大。进入大数据时代后，越来越多的人对从大量的数据中找到规律，并对其进行运用感兴趣。该房地产开发企业也引入了大数据应用理念，从企业战略的层面制定大数据应用计划，并成立相关的部门对员工进行大数据培训，对公司人员结构进行优化，提高公司创新、研发的投入使用比例，就是希望改变传统企业的开发—建设—销售思维模式，建立一套收集、存储客户数据信息，并系统的对庞大数据进行分析处理的完整大数据思维体系，继而能够发现有价值的潜在客户，深入认识客户的行为习惯、消费偏好，探索理性投资、多元化开发、创新性开发等商业机会，以及交叉销售和向上销售的机会，从而实现利润最大化。

该房地产开发企业凭借自身的数据优势，建立客户信息系统，将客户进行分类，然后通过大数据挖掘，提炼出客户的信息，有针对性的实现精准营销。该房地产开发企业精准营销的流程如下：根据以往的历史数据以及第三方数据，确定出产品的市场定位；把不同的客户按照需求分析及其聚类群体进行划分；通过分析客户个人资料和需求信息后，精准的找到其购房的需求，然后从不同方面实现房源情况与客户需求的匹配，也可以通过楼市调研以及竞争环境数据，建立精确营销体系，推送适用于不同口径和渠道的产品信息，并对数据进行实时更新；再通过财务数据和潜在价值等非结构化数据进行评估和反馈绩效，对已有的营销体系进行闭环修正与优化（具体见图 6-1）。

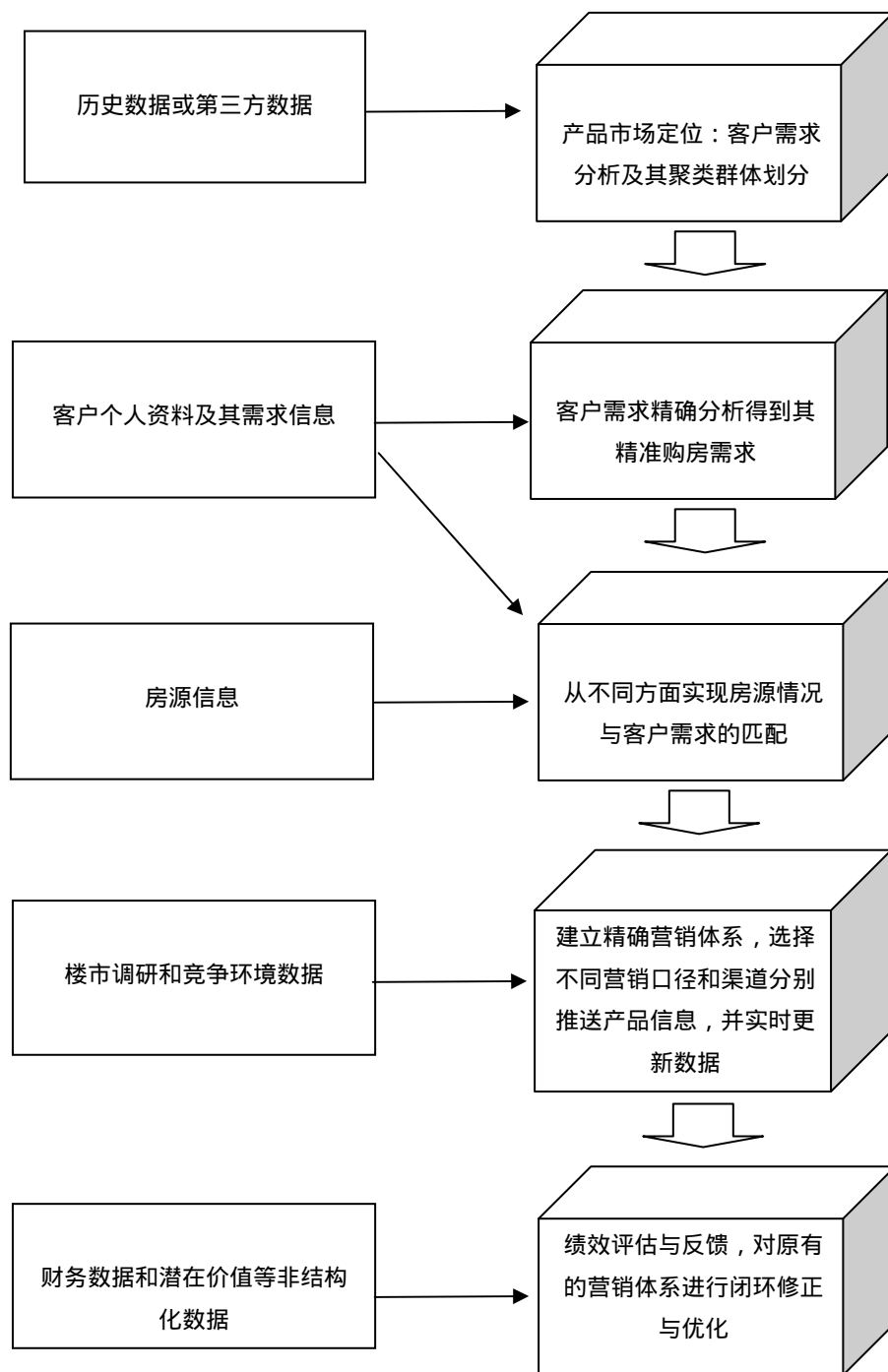


图 6-1 某房地产企业精准营销流程

6.2 房地产开发企业大数据应用能力成熟度模型的应用

通过对该房地产开发企业现状的介绍，得出的大数据应用能力应对能力成熟度模型主要应用于房地产开发企业的产品开发、服务开发、销售创新、交易及服

务的整个体系中，用本文所建立的大数据应用能力来探究房地产开发企业的大数据能力并对其进行评价，最后根据评价结果针对具体薄弱的地方对房地产开发企业提出相应的改善建议和措施。

下面将结合某房地产开发企业应用大数据的案例，通过前文所构建的成熟度评价指标体系和测评方法，对其大数据应用能力的成熟度水平进行测评。

6.2.1 评价指标集

首先确定所有指标隶属的层次和指标集，如图 6-2 所示。

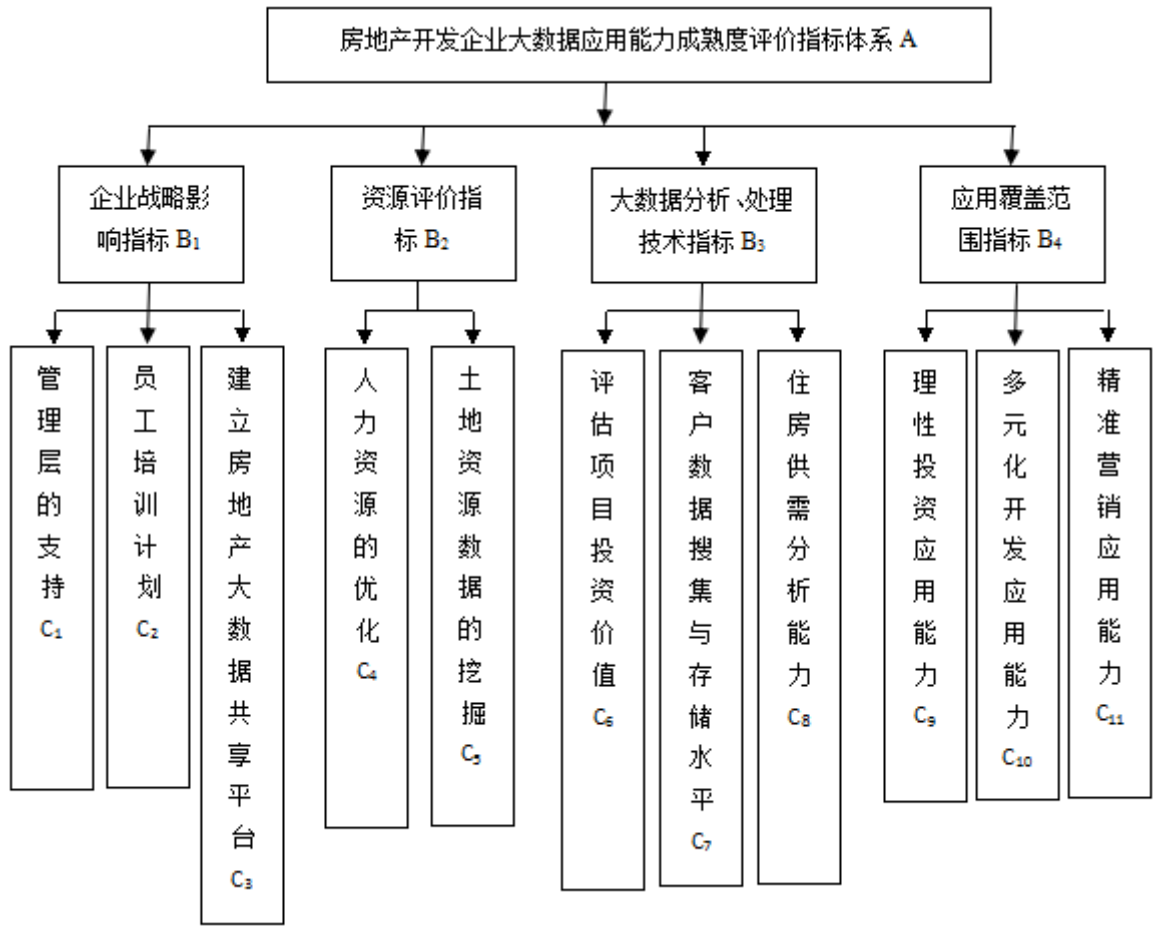


图 6-2 指标体系

6.2.2 权重计算

一级指标层权重计算

邀请5位案例企业有关专业人士对该企业的大数据应用的成熟度评价指标权重和成熟度评价指标进行打分，从而确定出案例企业大数据应用能力的成熟度情

况。在该企业的五位专业人士当中，分别来自于工程建设部、技术规划部、投资发展部、经营管理部和市场销售部等不同部门的领导干部或技术负责人，基本涵盖了房地产开企业从产品设计开发、建造生产、降低风险与节约成本、产品销售等全过程，且具有丰富的工作经验，因此，他们对本企业的大数据应用情况具有深刻的了解，可以保证回复的质量。

专家对一级指标层按照相对重要性的程度来进行两两比较打分，然后建立起判断矩阵。每个专家最后得出一个判断矩阵，接着对所给的矩阵中与之相对应的元素值求出其算术平均值，并作为最终判断矩阵中的对应元素，继而求出最终的综合判断矩阵（见附录 B）。通常把最终的综合判断矩阵做为计算 4 个类别指标层权重的凭据，最后得出最终的综合判断矩阵是：

$$\begin{bmatrix} 1 & 4 & 3 & 1/3 \\ 1/4 & 1 & 1/2 & 1/6 \\ 1/3 & 2 & 1 & 1/5 \\ 3 & 6 & 5 & 1 \end{bmatrix}$$

通过方根法能够计算出权数，具体的计算方法见式（5.1）式（5.2）。因此能够得到如表 6.1 所列出的成熟度标准的权数。

表 6.1 一级指标层权重

X	X ₁	X ₂	X ₃	X ₄	按行相乘	开四次方根	归一化得权重
X ₁	1	4	3	1/3	4	1.414213562	0.265423807
X ₂	1/4	1	1/2	1/6	1/48	0.229574884	0.043087296
X ₃	1/3	2	1	1/5	2/15	0.604275079	0.113412143
X ₄	3	6	5	1	90	3.080070288	0.578076751

由表 6.1 计算表格得到一级指标权重集 $A = (W_1, W_2, W_3, W_4) = (0.2654, 0.0431, 0.1134, 0.5781)$ 。本例中的指标权重保留到小数点后 4 位。

二级指标层权重计算

通过判断尺度表的判断对比之后，分别得到企业战略影响指标、资源评价指标、大数据分析、处理技术评价指标、应用覆盖范围评价指标的四个比较判断矩阵如下：

$$D_1 = \begin{bmatrix} 1 & 1/2 & 1/6 \\ 2 & 1 & 1/4 \\ 6 & 4 & 1 \end{bmatrix}$$

$$D_2 = \begin{bmatrix} 1 & 3 \\ 1/3 & 1 \end{bmatrix}$$

$$D_3 = \begin{bmatrix} 1 & 5 & 4 \\ 1/5 & 1 & 1/3 \\ 1/4 & 3 & 1 \end{bmatrix} \quad D_4 = \begin{bmatrix} 1 & 2 & 1/4 \\ 1/2 & 1 & 1/5 \\ 4 & 5 & 1 \end{bmatrix}$$

通过上述方法求得权重，计算方法见式（5.1）式（5.2），从而可以分别得出如下面所示的各层指标的权数的表。

表 6.2 企业战略影响指标权重

X	X ₁	X ₂	X ₃	按行相乘	开三次方根	归一化得权重
X ₁	1	1/2	1/6	1/12	0.436790232	0.106146124
X ₂	2	1	1/4	1/2	0.793700525	0.192880309
X ₃	6	4	1	24	2.884499141	0.700973566

由表 6.2 得到企业战略影响指标层权重集 $A_1 = (W_{11}, W_{12}, W_{13}) = (0.1061, 0.1929, 0.7010)$

表 6.3 资源评价影响指标权重

X	X ₁	X ₂	按行相乘	开平方根	归一化得权重
X ₁	1	3	3	1.732050807	0.7500
X ₂	1/3	1	1/3	0.577350269	0.2500

由表 6.3 得到资源评价指标层权重集 $A_2 = (W_{21}, W_{22}) = (0.7500, 0.2500)$

表 6.4 大数据分析、处理技术评价指标权重

X	X ₁	X ₂	X ₃	按行相乘	开三次方根	归一化得权重
X ₁	1	5	4	20	2.714417617	0.673810571
X ₂	1/5	1	1/3	1/15	0.405480133	0.100653930
X ₃	1/4	3	1	3/4	0.908560296	0.225535667

由表 6.4 得到大数据分析、处理技术评价指标层权重集 $A_3 = (W_{31}, W_{32}, W_{33}) = (0.6738, 0.1007, 0.2255)$ 。

表 6.5 应用覆盖范围评价指标权重

X	X ₁	X ₂	X ₃	按行相乘	开三次方根	归一化得权重
X ₁	1	2	1/4	1/2	0.793700526	0.235337743
X ₂	1/2	1	1/5	1/10	0.464158883	0.137626347
X ₃	4	5	1	20	2.114742527	0.627035911

由表 6.5 得到应用覆盖范围评价指标层权重集 $A_4 = (W_{41}, W_{42}, W_{43}) = (0.2354, 0.1376, 0.6270)$ 。

下面对各权重进行一致性检验，根据式（5.3）式（5.4），得到 C.I.和 C.R.的值。

表 6.6 一致性判断

指标类别	指标名称	λ_{\max}	C.I.	R.I.	C.R.
成熟度标准层	四个成熟度标准	4.1045	0.0348	0.90	0.0387
大数据应用实践层	企业战略影响指标	3.0092	0.0046	0.58	0.0079
	资源评价指标	2.0000	0.0000	0.00	
	大数据分析、处理技术评价指标	3.0869	0.0435	0.58	0.0750
	应用覆盖范围评价指标	3.0247	0.0124	0.58	0.0214

从表 6.6 可知:各级指标的 C.R.值都小于 0.1 ,这个结果符合逻辑循环且判断矩阵满足了一致性的要求，因此计算得到的权重是合理的。

6.2.3 测评结果计算

通过向以上专家进行问卷打分，n [1,N]，N=5，该一级指标下各专家得出的二级指标分值和用层次分析法计算出专家对权重的设置，如表 6.7 所示。

表 6.7 企业战略影响一级指标下 3 个二级指标的评审打分结果

一级指标	权重	二级指标	权重	评分				
				专家 1	专家 2	专家 3	专家 4	专家 5
U ₁	0.2654	U ₁₁	0.1061	3.6	2.7	3.0	3.8	3.2
		U ₁₂	0.1927	2.4	3.0	2.6	3.3	3.5
		U ₁₃	0.7001	2.3	2.5	2.0	1.9	2.8

计算每个专家对于二级指标所有灰类的白化权值

将 C₁₁₁=3.6，C₂₁₁=2.7，C₃₁₁=3.0，C₄₁₁=3.8，C₅₁₁=3.2，分别带入 f_k(c_{nij})，得到

每个专家的 $f_1, f_2 \cdots f_k \cdots f_s, k \in [1, s], s = 5, x_{11}^k = \sum_{n=1}^N f_k, x_{11}^{*k} = \frac{x_{11}^k}{\sum_{k=1}^s x_{11}^k}$ 。U₁₁、U₁₂、U₁₃ 的

白化权值分别如表 6.8~6.10 所示。

表 6.8 各专家对指标 U₁₁ 各灰度的白化权值

x_{11}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{11}^{*k}
0.2	f_1	0	0.2	0	0	0	0.027
2.467	f_2	0.267	0.867	0.667	0.133	0.533	0.333
3.667	f_3	0.933	0.467	0.667	0.8	0.8	0.496
1.066	f_4	0.4	0	0	0.533	0.133	0.144

表 6.9 各专家对指标 U_{12} 各灰度的白化权值

x_{12}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{12}^{*k}
0.667	f_1	0.4	0	0.267	0	0	0.086
3.333	f_2	0.933	0.667	0.933	0.467	0.333	0.431
3.201	f_3	0.267	0.667	0.4	0.867	1	0.414
0.533	f_4	0	0	0	0.2	0.333	0.069

表 6.10 各专家对指标 U_{13} 各灰度的白化权值

x_{13}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{13}^{*k}
2.333	f_1	0.467	0.333	0.667	0.733	0.133	0.318
3.934	f_2	0.867	1	0.667	0.6	0.8	0.537
1.066	f_3	0.2	0.333	0	0	0.533	0.145
0	f_4	0	0	0	0	0	0

等级判断。计算 U_{11} - U_{13} 属于各个灰度的隶属度 x_{ij}^* 。根据最大隶属度原则对各指标的等级进行判断。

$$x_{11}^* = [0.027, 0.333, 0.496, 0.144]$$

$$x_{12}^* = [0.086, 0.431, 0.414, 0.069]$$

$$x_{13}^* = [0.318, 0.537, 0.145, 0]$$

U_{11} 属于预测级, U_{12} 属于发现模式级, U_{13} 属于发现模式级。

计算得出该一级指标评价权。企业战略影响一级指标的白化权向量为 X_1^* 。

$$X_1^* = \begin{bmatrix} x_{11}^* \\ x_{12}^* \\ x_{13}^* \end{bmatrix} = \begin{bmatrix} 0.027 & 0.333 & 0.496 & 0.144 \\ 0.086 & 0.431 & 0.414 & 0.069 \\ 0.318 & 0.537 & 0.145 & 0 \end{bmatrix}$$

权向量 $W_1 = [w_1, w_2, w_3]^T = [0.1061, 0.1929, 0.7010]^T$ 。

$$G_1 = W_1^T \times X_1^* = [0.1061, 0.1929, 0.7010] \times \begin{bmatrix} 0.027 & 0.333 & 0.496 & 0.144 \\ 0.086 & 0.431 & 0.414 & 0.069 \\ 0.318 & 0.537 & 0.145 & 0 \end{bmatrix}$$

该案例的企业战略影响一级指标的评价权向量为：

$$G_1 = [0.242, 0.495, 0.234, 0.029]$$

可以判断该案例的企业战略影响指标是属于发现模式级的。

资源评价专家打分结果

表 6.11 资源评价一级指标下 2 个二级指标的评审打分结果

一级指标	权重	二级指标	权重	评分				
				专家 1	专家 2	专家 3	专家 4	专家 5
U ₂	0.0431	U ₂₁	0.7500	2.7	2.0	2.3	2.5	3.0
		U ₂₂	0.2500	3.0	3.1	3.3	2.9	3.0

表 6.12 各评审对指标 U₂₁ 各灰度的白化权值

x_{21}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{21}^{*k}
1.667	f_1	0.2	0.667	0.467	0.333	0	0.225
4.068	f_2	0.867	0.667	0.867	1	0.667	0.550
1.667	f_3	0.467	0	0.2	0.333	0.667	0.225
0	f_4	0	0	0	0	0	0

表 6.13 各评审对指标 U₂₂ 各灰度的白化权值

x_{22}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{22}^{*k}
0.067	f_1	0	0	0	0.067	0	0.009
3.134	f_2	0.667	0.6	0.467	0.733	0.667	0.448
3.534	f_3	0.667	0.733	0.867	0.6	0.667	0.505
0.267	f_4	0	0.067	0.2	0	0	0.038

由表 6.12~6.13 可得

$$x_{21}^* = [0.225, 0.550, 0.225, 0]$$

$$x_{22}^* = [0.009, 0.448, 0.505, 0.038]$$

U₂₁ 属于发现模式级，U₂₂ 属于预测级。

资源评价一级指标的白化权向量为 X_2^* 。

$$X_2^* = \begin{bmatrix} x_{21}^* \\ x_{22}^* \end{bmatrix} = \begin{bmatrix} 0.225 & 0.550 & 0.225 & 0 \\ 0.009 & 0.448 & 0.505 & 0.038 \end{bmatrix}$$

权向量 $W_2 = [w_1, w_2]^T = [0.7500, 0.2500]^T$ 。

$$G_2 = W_2^T \times X_2^* = [0.7500, 0.2500] \times \begin{bmatrix} 0.225 & 0.550 & 0.225 & 0 \\ 0.009 & 0.448 & 0.505 & 0.038 \end{bmatrix}$$

该案例的资源评价一级指标的评价权向量为：

$$G_2 = [0.171, 0.525, 0.295, 0.009]$$

可以判断该案例的资源评价指标是属于发现模式级的。

大数据分析、处理技术评价指标专家打分结果

表 6.14 大数据分析、处理技术评价指标下 3 个二级指标的评审打分结果

一级指标	权重	二级指标	权重	评分				
				专家 1	专家 2	专家 3	专家 4	专家 5
U ₃	0.1134	U ₃₁	0.6738	3.0	2.8	2.0	2.5	2.8
		U ₃₂	0.1007	2.5	3.6	3.8	3.6	3.2
		U ₃₃	0.2255	3.3	3.0	3.0	2.7	2.4

表 6.15 各评审对指标 U₃₁ 各灰度的白化权值

x_{21}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{21}^{*k}
1.266	f_1	0	0.133	0.667	0.333	0.133	0.174
3.934	f_2	0.667	0.800	0.667	1	0.800	0.542
2.066	f_3	0.667	0.533	0	0.333	0.533	0.284
0	f_4	0	0	0	0	0	0

表 6.16 各评审对指标 U₃₂ 各灰度的白化权值

x_{21}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{21}^{*k}
0.333	f_1	0.333	0	0	0	0	0.043
2.200	f_2	1	0.267	0.133	0.267	0.533	0.282
3.799	f_3	0.333	0.933	0.800	0.933	0.8	0.487
1.466	f_4	0	0.4	0.533	0.4	0.133	0.188

表 6.17 各评审对指标 U₃₃ 各灰度的白化权值

x_{21}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{21}^{*k}
0.600	f_1	0	0	0	0.2	0.4	0.082
3.601	f_2	0.467	0.667	0.667	0.867	0.933	0.491
2.935	f_3	0.867	0.667	0.667	0.467	0.267	0.400
0.2	f_4	0.2	0	0	0	0	0.027

由表 6.15~6.17 可得

$$x_{31}^* = [0.174, 0.542, 0.284, 0]$$

$$x_{32}^* = [0.043, 0.282, 0.487, 0.188]$$

$$x_{33}^* = [0.082, 0.491, 0.400, 0.027]$$

U₃₁ 属于发现模式级，U₃₂ 属于预测级，U₃₃ 属于发现模式级。

大数据分析、处理技术一级指标的白化权向量为 X_3^* 。

$$X_3^* = \begin{bmatrix} x_{31}^* \\ x_{32}^* \\ x_{33}^* \end{bmatrix} = \begin{bmatrix} 0.174 & 0.542 & 0.284 & 0 \\ 0.043 & 0.282 & 0.487 & 0.188 \\ 0.082 & 0.491 & 0.400 & 0.027 \end{bmatrix}$$

$$\text{权向量 } W_3 = [w_1, w_2, w_3]^T = [0.6738, 0.1007, 0.2255]^T$$

$$G_3 = W_3^T \times X_3^* = [0.6738, 0.1007, 0.2255] \times \begin{bmatrix} 0.174 & 0.542 & 0.284 & 0 \\ 0.043 & 0.282 & 0.487 & 0.188 \\ 0.082 & 0.491 & 0.400 & 0.027 \end{bmatrix}$$

该案例的大数据分析、处理技术一级指标的评价权向量为：

$$G_3 = [0.140, 0.504, 0.331, 0.025]$$

可以判断该案例的大数据分析、处理技术指标是属于发现模式级的。

应用覆盖范围指标专家打分结果

表 6.18 应用覆盖范围一级指标下 3 个二级指标的评审打分结果

成熟度标准	权重	大数据应用实践	权重	评分				
				专家 1	专家 2	专家 3	专家 4	专家 5
U ₄	0.5781	U ₄₁	0.2354	2.7	2.4	2.8	2.5	2.4
		U ₄₂	0.1376	3.2	3.0	3.3	3.2	3.8
		U ₄₃	0.6270	3.0	3.2	2.7	3.6	3.2

表 6.19 各评审对指标 U₄₁ 各灰度的白化权值

x_{21}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{21}^{*k}
1.466	f_1	0.2	0.4	0.133	0.333	0.4	0.187
4.533	f_2	0.867	0.933	0.800	1	0.933	0.576
1.867	f_3	0.467	0.267	0.533	0.333	0.267	0.237
0	f_4	0	0	0	0	0	0

表 6.20 各评审对指标 U₄₂ 各灰度的白化权值

x_{21}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{21}^{*k}
0	f_1	0	0	0	0	0	0
2.333	f_2	0.533	0.667	0.467	0.533	0.133	0.321
3.934	f_3	0.8	0.667	0.867	0.8	0.8	0.541
1	f_4	0.133	0	0.2	0.133	0.533	0.138

表 6.21 各评审对指标 U₄₃ 各灰度的白化权值

x_{21}^k	f_k	专家 1	专家 2	专家 3	专家 4	专家 5	x_{21}^{*k}
0.2	f_1	0	0	0.2	0	0	0.027
2.867	f_2	0.667	0.533	0.867	0.267	0.533	0.387
3.667	f_3	0.667	0.8	0.467	0.933	0.8	0.496
0.666	f_4	0	0.133	0	0.4	0.133	0.090

由表 6.19~6.21 可得

$$x_{41}^* = [0.187, 0.576, 0.237, 0]$$

$$x_{42}^* = [0, 0.321, 0.541, 0.138]$$

$$x_{43}^* = [0.027, 0.387, 0.496, 0.090]$$

U_{41} 属于发现模式级， U_{42} 属于预测级， U_{43} 属于预测级。

应用覆盖范围一级指标的白化权向量为 X_4^* 。

$$X_4^* = \begin{bmatrix} x_{41}^* \\ x_{42}^* \\ x_{43}^* \end{bmatrix} = \begin{bmatrix} 0.187 & 0.576 & 0.237 & 0 \\ 0 & 0.321 & 0.541 & 0.138 \\ 0.027 & 0.387 & 0.496 & 0.090 \end{bmatrix}$$

$$\text{权向量 } W_4 = [w_1, w_2, w_3]^T = [0.2354, 0.1376, 0.6270]^T$$

$$G_4 = W_4^T \times X_4^* = [0.2354, 0.1376, 0.6270] \times \begin{bmatrix} 0.187 & 0.576 & 0.237 & 0 \\ 0 & 0.321 & 0.541 & 0.138 \\ 0.027 & 0.387 & 0.496 & 0.090 \end{bmatrix}$$

该案例的应用覆盖范围一级指标的评价权向量为：

$$G_4 = [0.061, 0.422, 0.442, 0.075]$$

可以判断该案例的应用覆盖范围指标是属于预测级的。

计算综合评价权。

$$\text{总评价目标 } U \text{ 的权向量 } W = [W_1, W_2, W_3, W_4]^T = [0.2654, 0.0431, 0.1134, 0.5781]^T$$

所有一级指标的评价权向量为 G 。

$$G^* = W^T \times G = [0.2654, 0.0431, 0.1134, 0.5781] \times \begin{bmatrix} 0.242 & 0.495 & 0.234 & 0.029 \\ 0.146 & 0.513 & 0.327 & 0.014 \\ 0.140 & 0.504 & 0.331 & 0.025 \\ 0.061 & 0.422 & 0.442 & 0.075 \end{bmatrix}$$

$$= [0.122, 0.455, 0.369, 0.054]$$

计算综合评价权值和综合评价值的隶属度值。综合评价权值 y 等于：

$$y = [1, 2, 3, 4] \times G^{*T}$$

$$= [1, 2, 3, 4] \times [0.122, 0.455, 0.369, 0.054] = 2.545$$

将 y 值 2.545 带入 f_k ，归一化后综合评价值隶属于每个灰度的隶属度为：

$$F = [0.185, 0.593, 0.222, 0]。$$

判断综合等级。判断该案例综合指标分析，其隶属度为 0.593，属于发现模式级。也可以认为有 0.593 比例的专家认为该企业是属于发现模式级的。

案例中该房地产开发企业大数据应用能力成熟度水平处于发现模式级，说明该房地产开发企业在运行过程中，已经积累了大量的数据，并且能够从海量的数据中发现对业务有影响的模式，包括成功模式或者失败模式，从而可以有针对性

和导向性的对目标进行调整或者设置。但同时也说明大数据应用能力成熟度水平还不够高，房企需要进行内部分析总结，同时结合评估指标和评估因素分析，对“具有某某属性的顾客”其个人属性与过去的关联数据进行对照分析，从而更加精准地预测出该潜在顾客会购买什么样的房源、需要什么样的服务。形成新的企业目标从而提高成熟度水平，最终提高大数据的应用能力。

6.3 改善房地产开发企业大数据应用能力的相应措施

由上一小节中算法的演算，可以分别得出每个大数据二级指标的灰色评价权向量，一级指标的灰色评价权向量和综合灰色评价权向量，由评价结果可知，该房地产开发企业的大数据应用能力成熟度水平还不够高，因此根据上一小节的计算结果针对其薄弱环节的地方提出相应的改善措施，其中包括从应用大数据分析、大数据处理能力、数据科学家的数量和质量、关键技术等方面。

提高数据存储技术

房地产开发企业大数据应用的影响因素有很多，随着结构化数据和非结构化数据量的持续增长，未来我们将生活在信息轰炸的时代，房地产开发企业也将会变成数据驱动型公司，越来越多样化的数据来源，使得此前设计的存储系统已经无法满足大数据应用的需要。必须通过软件的需求推动硬件的发展、修改基于块和文件的存储系统的架构设计以适应新的要求。

1) 通过提高相应海量数据存储系统的扩展能力。在解决容量的问题上，房地产开发企业可以针对本企业的特点来解决数据储存的容量问题，同时可以通过增加模块或磁盘柜来增加容量。

2) 解决延迟、强化实时性的能力。房地产开发企业在涉及资金流、特别是与网上交易或者金融类相关的应用往往是其关注的重点，一般而言“大数据”应用环境需要较高的 IOPS 性能，例如 HPC 高性能计算，房地产开发企业应该着重研发各种模式的固态存储设备。

3) 提高安全标准和保密性要求。房地产开发企业在应用大数据时往往要相互参考多种类型的数据，但在以往则并不会出现这种数据混合访问的情况，大数据应用也使得安全性问题变得越来越重要，增强企业级闪存的性能和可靠性，从而实现简单、透明的应用加速，达到安全、方便的目的。

加强大数据处理能力

伴随着计算机技术与互联网技术的发展，在产生大数据的同时，也为人们带来了全新的云计算技术。云计算技术带来的大数据处理能力，使得分析和掌握大数据中蕴藏的无尽信息、知识和智慧成为可能。

1) 提升计算节点的处理能力和计算架构的优化与革新。数据的计算能力是房地产开发企业大数据处理过程中的核心能力。大数据处理不仅带来了数据存储管理方面的挑战,更是对处理平台的计算能力提出了极高的要求,在大数据环境下,随着需要处理的与房企相关的数据量由 TB 级迈向 PB 甚至 ZB 级,即使一个简单的数据排序或者分类问题也会因为数据量的异常庞大而变得十分的困难,房企产开发企业可以采用多处理技术和并行计算两种方式提升数据的计算能力。

2) 改善大数据处理系统的容错性。房地产开发企业大数据处理系统的容错性又包括数据存储容错与计算任务容错,数据存储容错是指当系统中的部件或节点由于硬件或软件出现故障,导致数据、文件损坏或丢失时,系统能够自动将这些损坏或丢失的文件和数据恢复到故障发生前的状态,使系统能够维持正常运行的技术,企业可以采用磁盘镜像和磁盘双工、基于 RAID 的磁盘容错以及基于集群的数据容错技术来处理。要完成一个完整的计算任务容错,必须解决失效节点检测、计算任务迁移及数据定位与获取三个重要问题。

提高数据科学家的数量与质量。

1) 建立房企数据科学家的培养机制。数据科学家通过运用统计、分析工具,从汗牛充栋的数据中找到有价值的,并以通俗易懂的形式将数据的价值传达给房企决策者,达到辅助决策的目的。但是拥有这种能力的数据科学家却凤毛麟角,房地产开发企业应该根据自身的特点,建立科学的培养机制,定向培养自己的人才。

2) 提高数据科学家的综合素质。数据科学家应该具备多种技能,大多要求数据科学家具备编程、计算机科学相关的专业背景。通俗的讲,就是掌握对处理大数据所必需的 Hadoop、Mahout 等大规模并行技术和机器学习相关的技能技术。除了需要掌握数学、统计方面的知识外,还需要具备使用 SPSS、SAS 等主流统计分析软件的技能。分析数字罗列所组成的数据中所包含的意义,开发网页原型,使用外部 API 将图表、地图、Dashboard 等其他服务统一起来,使分析结果变得可视化,这也是数据科学家必须掌握的技能之一。

关键技术的掌握和使用。

对于目前的房企应用大数据的浪潮,在技术层面上提供支持的,是开源分布式处理框架 Hadoop ,Hadoop 是以开源形式发布的一种对大规模数据进行分布式处理的技术。特别是处理起大数据时代所必需的非结构化数据时,Hadoop 在性能和成本方面都具有优势,而且通过横向扩展进行扩容也相对容易。考虑到 Hadoop 还处于发展阶段且开发投资成本过高,一般房企难以承受,因此,一般房企不必非要推翻和替换现有的技术,在销售数据和客户数据等结构化数据的存储和处理上,只要使用传统的关系型数据库和数据仓库就可以了。

6.4 本章小结

在前文构建的房地产开发企业大数据应用能力评价模型的基础上，通过案例的分析进一步对成熟度模型的应用进行了详细的阐述，案例中所有的数据都进行了处理能够保证评价结果的合理性，并且在上述所有理论知识的基础上详细说明了房地产开发企业大数据应用能力的具体运用过程，计算结果表明成熟度模型能够较容易地计算出企业在建设过程中对大数据应用的成熟度等级值，对其能力等级给出了正确的判断，最后并依据评价结果及详细的分析对企业的薄弱点提出改进的方法。

第七章 结论与展望

7.1 结论

本文在研究了房地产开发企业大数据应用能力的现状与实践成果的基础上，对房地产开发企业大数据应用能力评价指标体系和能力等级测评两方面分别进行了深入的研究，具体成果如下：

本文站在房地产开发企业的角度，通过对国内外大量研究文献资料的整理分析，总结出了房地产开发企业应用大数据时的 47 条评价因素，并依附这些因素建立了房地产开发企业大数据应用能力成熟度指标体系，该指标体系是由目标层、一级指标、二级指标这三个层次组成，其中目标层是用于设立指标体系的终极目标，就是进行房地产开发企业大数据应用能力的等级评价，而一级指标是为总目标层而展开设置的，可以全面并且准确的反映出总目标，本文将一级指标设置为企业战略影响指标、资源评价指标、大数据分析/处理技术评价指标和应用覆盖范围评价指标四个方面。而二级指标又是对一级指标层中各指标的进一步细化表达，是更为详细具体的指标设置，同时也是指标体系中最重要的一部分，本文整个体系包括 4 个一级指标，11 个二级指标，该指标体系的建立为引入成熟度理念以及评价房地产开发企业大数据应用能力等级奠定了良好的基础。

本文在对比分析了层次分析法、模糊综合评价法、数据包络分析法、多目标数学规划法、BP 神经网络法、作业成本法等 7 种多目标评价方法的优缺点后最终确定了灰色理论的评价方法，并且介绍了应用灰色理论进行评价的适用性分析。最后依据改进了的能力等级、用系统工程原理对指标进行处理后的最优指标和新的评价原则这三个方面融合在一起所构建出新的能力成熟度模型。该模型非常适用于这些指标因素较多，同时过程较为复杂的情况。并且该指标体系还详细叙述了层次分析法和灰色综合评价法的具体应用过程，进一步表明这个模型可以较为准确的量化研究房地产开发企业大数据应用能力等级。

7.2 不足与展望

7.2.1 本文研究的不足

由于本人的认识和能力限制，本文的研究在以下方面略有不足：

理论研究不够深入，通过查阅有关的书籍和资料，查看前人的理论深度，自知自己理论知识的短缺与不足，还需要向各位前辈好好学习；

指标体系不够完善，由于实际应用中需要考虑的因素非常多，本文考虑的

可能不够完善，定性指标较多，还需要进一步的细化和量化。

在房地产开发企业大数据应用能力成熟度等级评价方面，所用的测评方法比较简单。

7.2.2 未来研究的展望

由于我国建筑行业引入成熟度模型的时间较短，对该模型的研究及运用还处于初级阶段，并且涉及到房地产开发企业大数据应用能力方面可以用来参考和学习的文献也不多，因此本文的研究可能还是停留在初级阶段，尤其是构建的房地产开发企业大数据应用能力的成熟度评价模型方面还有很多缺陷，但这是个非常有价值的研究方向，有待于以后进一步研究与探讨。

致 谢

时光任然，物是人非，三年弹指一挥间，研究生生活即将告一段落，一想起这段时段时光让我感慨万千，这期间不仅扩展了我的知识储备同时也改变了我的知识结构，树立了坚韧的意志品质和独立生活的能力。

本论文是在我的导师孟卫军老师的悉心指导下完成的，在本文的构思和写作过程中，他给了我一个清晰的研究方向，并对我在研究过程中遇到的困难给予耐心的讲解。对此，我表示衷心的感谢，感谢他的默默付出和悉心教导。

其次，要感谢所有重庆交通大学经济与管理学院的老师们，他们的授课给了我很大的启发，让我认识到了自己的不足，同时给我的论文写作也提供了良好的思路，带来了很大的帮助。

然后，感谢我的寝室室友和同门师兄弟，是他们陪我一起过度了研究生三年的生活，在生活相互关照，在学习上相互探讨，在我松懈时给予我动力，在我迷茫时给我指明了方向，感谢他们在我生活和学习中的帮助，有了他们，我的研究生生活才更加精彩。

最后，感谢我亲爱的父母们，感谢他们多年的培育和养育之恩，感谢他们对我的无私奉献，感谢他们在我人生路上关爱、支持与包容。

参考文献

- [1] 中国共产党新闻网.发展大数据确实有道理[EB/OL].<http://cpc.people.com.cn8/c38547527176083.html>
- [2] 刘江娜,大数据时代:为什么数据分析能让你的企业脱颖而出[J].现代企业育,2013(3) :32
- [3] Science. Special online collection: Dealing with data [EB/OL].[2012-10-02].<http://www.sciencemag.org/site/special/data/>.2011
- [4] Manyika J, Chui M, Brown B.et al. Big data: The next frontier for innovation, competition, and productivity [R/OL]. [2012-10-02].http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation
- [5] Nature.Big Data [EB/OL].[2012-10-02].<http://www.nature.com/news/specials/bigdata/index.html>
- [6] Agrawal D, Bernstein E, et, al. Challenges and opportunities with big data—A community white paper developed by leading researchers across the United States[R/OL]. [2012-10-02]. <http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf>
- [7] World Economic Forum. Big data, big impact: New possibilities for international development [R/OL].[2012-10-02].http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf
- [8] UN Global Pulse. Big Data for Development: Challenges&Opportunities[R/OL].[2012-10-02].<http://www.unglobalpulse.org/projects/BigDataforDevelopment>
- [9] Big Data Across the Federal Government [EB/OL]. [2012-10-02].http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_fact_sheet_final_1.pdf
- [10] Bryant R E, Katz R H, Lazowska E D. Big-Data computing: Creating revolutionary breakthroughs in commerce, science, and society[R]. [2012-10-02]. http://www.cra.org/ccc/docs/init/Big_Data.pdf
- [11] 李国杰,程学旗.大数据研究:未来科技及经济社会发展的重大战略领域—大数据的研究现状与科学思考.中国科学院院刊. 2012.3
- [12] 孟小峰,慈祥.大数据管理:概念、技术与挑战[J].2013,50(1): 146-169
- [13] 杜丹阳,李爱华.大数据在我国房地产企业中的应用研究,2014, (8): 68-74
- [14] 刘翌,张玉娟.大数据时代:房地产业的机遇与挑战[J].河南商业高等专科学校校报, 2013, (12): 20-22
- [15] 严娟.基于大数据的房地产企业精确营销研究[J].市场周刊,2013,(9):66-67
- [16] 李鹏,吕欣.以大数据辅助房地产宏观决策[J].宏观经济管理,2014,(8):34-36
- [17] 陈大川,张宝山.大数据技术在住房信息系统中的应用[J].业务与运营,2013

- [18] 王珊,王会举,覃维派.架构大数据：挑战、现状与展望[J].计算机学报,2011,(10):1741-1752
- [19] 胡雄伟,张宝林,李抵飞.大数据研究与应用综述[J].标准科学,2013 (10) : 18-21
- [20] 李芬,朱志祥,刘盛辉.大数据发展现状及面临的问题[J].西安邮电大学学报,2013,18 (5) : 100-103.
- [21] 包达尔罕,罗敏,刘智伟,赵静.大数据及其应用技术探讨[J].内蒙古民族大学学报(自然科学版). 2014.(3):156-158
- [22] 陈志.战略性新兴产业发展中的商业模式创新研究[J].经济体制改革,2012,(5):112-116
- [23] 程明书,彭晓愈.未来房地产网络营销在线交易模式[J].房地产市场,2013,(8):31-36
- [24] 肖剑.房地产营销新热点—体验式营销[J].房地产营销,2004,(4):51-52
- [25] 国家统计局浙江调查总队课题组[J].统计科学与实践,2014,(8):12-15
- [26] Bill Curtis. William E Hefley, Sally Miller. People Capability Maturity Model Version 2.0 C-MU/SEI.July 2001
- [27] William E Hefley, Bill Curtis People CMM-Based Assessment Method Description Version 1.0[M]. CMU/SEI, August 2008
- [28] 吕潇潇.房地产项目管理能力成熟度评价研究[J].现代物业,2014,(08):122-124
- [29] 冉秀阶.房地产企业项目管理成熟度模型构建及评价研究[D].:[硕士学位论文].重庆:重庆大学,2009
- [30] Big data, http://en.wikipedia.org/wiki/Big_data.
- [31] Benjamin Woo World wide Big Data Technology and Service 2012-2015 Forecast.2012.5[16]
Big data <http://www.gartner.com/it-glossary/big-data>.
- [32] 马建广,姜巍.大数据的概念、特征及其应用[J].国防科技,2013,(4):10-16
- [33] obama Administration Unveils “Big Data” Initiative: Announces \$200 Million in New R&D Investment. Office of Science and Technology Policy Executive Office of the President.2012.5
- [34] The Economist. Data, data, everywhere—A special report on managing information [EB/OL]. [2012-10-02]. <http://www.economist.com/node/15557443>
- [35] Kumar R. Two computational paradigm for big data [EB/OL]. [2012-10-02]. KDD summ school, 2012. <http://kdd2012.sigkdd.org/sites/images/summerschool/Ravi-Kumar.pdf>
- [36] InformationWeek Report. The big data management challenge[R/OL]. [2012-10-02]. <http://reports.informationweek.com/abstract/81/8766/business-intelligence-and-information-management/research-the-big-data-management-challenge.Html>
- [37] Storm [EB/OL]. [2012-10-02]. <http://github.com/nathanmarz/storm>
- [38] Neumeyer L, Robbins B, Nair A, et al. S4: Distributed Stream Computing Platform[C]//Proc of ICDM Workshops 2010. Piscataway, NJ: IEEE, 2010:170-177
- [39] Goodhope K, Koshy J, Kreps J, et al. Building LinkedIn’s Real-time Activity Data Pipeline[J]

- Data Engineering,2012,35(2):33-45
- [40] Dean J, Ghemawat S. MapReduce: Simplified processing on large clusters [C] //Proc of OSDI2004. Berkeley, CA;USENIX association,2004:137-150
- [41] Das S. Data Infrastructure at LinkedIn[C/OL]//Proc of the 5th Extremely Large Databases Conf.http://www-conf.slac.stanford.edu/xldb2011/talks/xldb2011_tue_1005_Linkedln.pdf.2011
- [42] ScholarSpace[EB/OL]. [2012-10-02]. <http://www.cdblp.cn/>
- [43] Hass L. Integrating Extremely Large Data is Extremely Challenging[C/OL]//proc of XLDB Asia 2012.<http://idke.ruc.edu.cn/xldb/www.xldb-asia.org/program.html>.
- [44] Rajaraman A, Jeff Ullman. Mining of Massive Datasets[M/OL]. [2012-10-02]. [http://i. Stanford.edu/ullman/mmds.html](http://i.Stanford.edu/ullman/mmds.html)
- [45] Chapman A, Allen M D, Blaustein B. It's About the Data: Provenance as a Tool for Assessing Data Fitness[C]//Proc of the 4th USENIX Workshop on the Theory and Practice of Provenance. Berkeley, CA; USENIX Association, 2012
- [46] Hadoop[EB/OL]. [2012-10-02]. <http://hadoop, apache, org/index. html>
- [47] Seo S, Jang I, Woo K, et al. HPMR: Prefetching and preshuffling in shared MapReduce computation environment[C]//Proc of Cluster 2009. Piscataway, NJ: IEEE, 2009:1-8
- [48] Lu Wei, Shen Yanyan, Chen Su, et al: Efficient processing of k nearest joins using MapReduce[J]. PVLDB,2012, 5(10): 1016-1027
- [49] Gudmundsson G p, Amsaleg L, Jonsson B p, Distributed high-dimensional index creation using Hadoop, HDFS and C++[C]//Proc of CBMI2012. Piscataway, NJ:IEEE, 2012:1-6
- [50] Thusoo A, Sarma J S, Jain N, et al. Hive-A petabyte scale data warehouse using HadoopP[C]//Proc of ICDE 2010. Piscataway, NJ: IEEE, 2010:996-1005
- [51] Su Xueyuan, Swart G. Oracle in-database hadoop: when mapreduce meets RDBMS[C]//Proc of SIGMOD 2012:779-790
- [52] Yang Yan, Ni Xianhua, Wang Hongjun, et al. Parallel implementation of ant-based clustering algorithm based on Hadoop[C]//Proc of ICSI 2012. Berlin: Springer, 2012:190-197
- [53] Chiky R, Ghislotti R, Kazi-Aoul Z. Development of a distributed recommender system using the Hadoop framework[C]//Proc of EGC. 2012: 495-500
- [54] (美) J.肯特. 克劳福德著, 肖艳颖译. 项目管理成熟度模型. 北京: 机械工业出版社, 2008.5
- [55] International Standards Organization. ISO 9000-3 Guidelines for the application of 9001-Model for Quality Assurance in Design/Development, Production, Installation and Service. Geneva, 1987: 3-7
- [56] 何新贵, 王玮, 软件能力成熟度模型[M], 北京: 清华大学出版社, 2000

- [57] Mary Beth Chrissis Mike Konrad and Sandy Shrum.CMMI: Guidelines for Process Integration and Product Improvement. Addison-Wesley 2003
- [58] Mary Beth Chrissis, Mike Konrad, and Sandy Shrum. CMMI: Guidelines for process Integration and Product Improvement. Addison-Wesley, 2003
- [59] 关芬娜, 基于能力成熟度模型 CMM 的低碳经济评价体系的研究[D].华南理工大学,2013
- [60] Harold Kerzner, Strategie planning for project management using a project management Maturity model New York NY. 2001
- [61] 冯芳, 能力成熟度模型在通信软件企业应用研究. [J].中国新通信.2013,(17): 67-68
- [62] 房地产开发企业.http://wiki.mbalib.com/wiki/[EB/OL].MBA.智库百科.2011.9.30
- [63] Robert G. Cooper, Scott J. Edgett, Portfolio Management for New Products: Picking the Winners. Working Paper, 2001(11):56-78
- [64] 刘书庆,韩亚辉.项目施工战略合作伙伴评价与选择方案研究[C].中国科学技术大学,2008
- [65] Keith Goffin, Fred Lemke, Marek Szwajczewski. An exploratory study of “close”suppliermanufacturer relationships. Journal of Operations Management,2006(4):189-209
- [66] Zhihan Lv, Javier Chirivella, Pabto Gagliardo(2016). Bigdata Oriented Multimedia Mobile Health Applications. J Med Syst, 40(3), 120.
- [67] Qu, L., Wang, Y., Orgun, M. A., Liu, H., & Bouguettaya, A.(2015). CCClould: Context-aware and credible cloud service selection based on subjective assessment and objective assessment IEEE Transactions on Services Computing, 8(3), 369-383.
- [68] Seokhoon Kim, Wonshik Na(2016). Safe Data Transmission Architecture Based on Cloud for Internet of Things. Wireless Pers Commun, 86(10), 287-300.
- [69] Fowler,F. J. Survey research methods. Sage, CA: Newbury Park, 1988(5):23-45
- [70] 高忠仕.知识转移、知识搜索及纺织学习绩效关系研究[D].杭州:浙江大学,2008
- [71] 刘书庆,董雅文.基于 ISO9000 族标准的质量管理体系有效性评价方案设计[J].工业工程, 2006,9(3):88-93

附录 A

房地产开发企业大数据应用能力成熟度评价指标调查问卷

先生/女士：

问卷编码：

您好！我是重庆交通大学 2013 级管理科学与工程专业的研究生，正在对房地产开发企业大数据应用能力评价指标进行调查研究，旨在了解房地产开发企业应用大数据时的重点分析考虑的相关指标体系，以理论结合实践，制定出一套房地产开发企业大数据应用能力成熟度评价的指标体系，以期对房企应用大数据有所帮助。希望你能积极参与，你的答案无对错之分，只需保证客观真实即可，我们将对您提供的答案和个人信息完全保密。非常感谢您的支持与配合。

请将您的答案填写在题后的划线处，或在相关选项上打钩。如有疑问，请咨询有关人员。

第一部分

1. 请问您的性别是？

A 男 B 女

2. 请问您从事与房地产开发的时间为多久？

A 5 年以下 B 5—10 年 C 11—15 年 D 15 年以上

3. 请问您的职务/职称是什么？

A 教授/研究院级工程师 B 副高级工程师 C 中级工程师 D 初级工程师 E 其他

4. 请问您公司的员工规模？

A 0—50 人 B 51—100 人 C 101—500 人 D 500 人以上

第二部分 影响指标调查

在这一部分中，请您根据理论知识和实际工作经验，对下面的一级指标隶属于总目标，二级指标隶属于一级指标的关联度进行评价。其中，1 表示非常相关，2 表示相关，3 表示比较相关，4 表示一般，5 表示弱相关。

请您认为在正确的数字上面打

总指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	企业战略影响指标	1	2	3	4	5
2	资源评价指标	1	2	3	4	5
3	市场评价指标	1	2	3	4	5
4	大数据分析/处理技术评价指标	1	2	3	4	5

5	应用环境评价指标	1	2	3	4	5
6	应用覆盖范围评价指标	1	2	3	4	5
7	组织评价指标	1	2	3	4	5
8	应用风险评价指标	1	2	3	4	5

企业战略影响指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	管理层的支持	1	2	3	4	5
2	员工培训计划	1	2	3	4	5
3	大数据信息沟通能力的提升	1	2	3	4	5
4	制定企业大数据应用计划	1	2	3	4	5
5	树立大数据应用意识文化	1	2	3	4	5
6	与第三方数据共享机制	1	2	3	4	5

资源评价指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	人力资源的优化	1	2	3	4	5
2	财务资源的利用	1	2	3	4	5
3	内外资源的迅捷调配	1	2	3	4	5
4	相关经验的借鉴	1	2	3	4	5
5	创新、研发资源的使用	1	2	3	4	5

市场评价指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	应用的市场效益	1	2	3	4	5
2	住户消费需求掌握情况	1	2	3	4	5
3	大数据应用市场的潜力	1	2	3	4	5
4	竞争对手情况	1	2	3	4	5
5	消费者未来的市场偏好	1	2	3	4	5
6	大数据应用市场的规模	1	2	3	4	5
7	大数据应用的市场定位	1	2	3	4	5
8	大数据应用市场的竞争激烈程度	1	2	3	4	5
9	大数据应用商业化成功的可能性	1	2	3	4	5

大数据分析/处理技术评价指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	云计算水平	1	2	3	4	5
2	数据搜索与存储能力	1	2	3	4	5
3	数据挖掘能力	1	2	3	4	5
4	非结构化数据提炼技术	1	2	3	4	5
5	机器学习能力	1	2	3	4	5
6	统计分析技术	1	2	3	4	5
7	自然语言处理	1	2	3	4	5
8	沉默数据激活能力	1	2	3	4	5

应用环境评价指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	经济环境	1	2	3	4	5
2	社会环境	1	2	3	4	5
3	政策环境	1	2	3	4	5
4	地缘环境	1	2	3	4	5

大数据应用覆盖范围评价指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	理性投资应用能力	1	2	3	4	5
2	多元化开发应用能力	1	2	3	4	5
3	创业性投资应用能力	1	2	3	4	5
4	营销应用能力	1	2	3	4	5

组织评价指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	企业大数据应用管理的权属	1	2	3	4	5
2	大数据应用语言的普及	1	2	3	4	5
3	企业大数据应用构架的监督、审查和改进	1	2	3	4	5
4	大数据应用融入业务流程	1	2	3	4	5

应用风险评价指标：（1=非常相关，2=相关，3=比较相关，4=一般，5=弱相关）

序号	指标	相关程度				
1	政策性风险	1	2	3	4	5
2	财务风险	1	2	3	4	5
3	信息识别能力	1	2	3	4	5
4	个人信息数据模拟化处理水平	1	2	3	4	5
5	数据所有权的保留度	1	2	3	4	5
6	数据潜在价值保留度	1	2	3	4	5
7	利润风险	1	2	3	4	5

附录 B

房地产开发企业大数据应用能力的成熟度评价指标权重打分调查表

先生/女士：

问卷编码：

您好！我是重庆交通大学 2013 级管理科学与工程专业的研究生，正在对房地产开发企业大数据应用能力评价权重进行调查研究，旨在了解房地产开发企业应用大数据时的重点分析考虑的相关指标体系的指标权重，以期对房地产开发企业应用大数据有所帮助。希望你能积极参与，你的答案无对错之分，只需保证客观真实即可，我们将对您提供的答案和个人信息完全保密。非常感谢您的支持与配合。

评分原则如下：

判断尺度	定义
1	A 和 B 相同重要
3	A 比 B 较为重要
5	A 比 B 重要
7	A 比 B 非常重要
9	A 比 B 绝对重要
2、4、6、8	介于上述两个相邻判断尺度中间
倒数	A 比 B 的重要性比为 a ，则 B 比 A 的重要性比为 $1/a$

表 B-1 房地产开发企业大数据应用能力成熟度评价专家评断矩阵

	企业战略影响 X_1	资源评价 X_2	大数据分析、处理技术指标 X_3	应用覆盖范围指标 X_4
企业战略影响 X_1	1	×	×	×
资源评价 X_2		1	×	×
大数据分析、处理技术指标 X_3			1	×
应用覆盖范围指标 X_4				1

表 B-2 企业战略影响指标专家评判矩阵

	管理层的支持 X_{11}	员工培训计划 X_{12}	制定大数据应用计划 X_{13}
管理层的支持 X_{11}	1	×	×
员工培训计划 X_{12}		1	×
制定大数据应用计划 X_{13}			1

表 B-3 资源评价指标专家评判矩阵

	人力资源的优化 X_{11}	创新、研发资源的使用 X_{12}
人力资源的优化 X_{11}	1	×
创新、研发资源的使用 X_{12}		1

表 B-4 大数据分析、处理技术指标专家评判矩阵

	云计算水平 X_{11}	数据搜集与存储水平 X_{12}	统计分析技术 X_{13}
云计算水平 X_{11}	1	×	×
数据搜集与存储水平 X_{12}		1	×
统计分析技术 X_{13}			1

表 B-5 应用覆盖范围指标专家评判矩阵

	理性投资应用能力 X_{11}	多元化开发应用能力 X_{12}	营销应用能力 X_{13}
理性投资应用能力 X_{11}	1	×	×
多元化开发应用能力 X_{12}		1	×
营销应用能力 X_{13}			1

附录 C

房地产开发企业大数据应用能力的成熟度评价指标打分调查表

先生/女士：

问卷编码：

您好！我是 2013 级管理科学与工程专业的研究生，正在对房地产开发企业大数据应用能力的成熟度评价指标进行打分，以期对房企在未来的大数据应用时有所帮助。您的回答没有对错之分，只要能反映真实情况即可，希望您能积极参与。我们将对您的回答及其他信息完全保密。非常感谢您的支持和配合。

一级指标	权重	二级指标	权重	评分					
				评审 1	评审 2	...	评审 n	...	评审 N
U ₁	W ₁	U ₁₁	w ₁₁	C ₁₁₁	C ₂₁₁	...	C _{n11}	...	C _{N11}
		U ₁₂	w ₁₂	C ₁₁₂	C ₂₁₂	...	C _{n12}	...	C _{N12}
		U ₁₃	w ₁₃	C ₁₁₃	C ₂₁₃	...	C _{n13}	...	C _{N13}
U ₂	W ₂	U ₂₁	w ₂₁	C ₁₂₁	C ₂₂₁	...	C _{n21}	...	C _{N21}
		U ₂₂	w ₂₂	C ₁₂₂	C ₂₂₂	...	C _{n22}	...	C _{N22}
U ₃	W ₃	U ₃₁	w ₃₁	C ₁₃₁	C ₂₃₁	...	C _{n31}	...	C _{N31}
		U ₃₂	w ₃₂	C ₁₃₂	C ₂₃₂	...	C _{n32}	...	C _{N32}
		U ₃₃	w ₃₃	C ₁₃₃	C ₂₃₃	...	C _{n33}	...	C _{N33}
U ₄	W ₄	U ₄₁	w ₄₁	C ₁₄₁	C ₂₄₁	...	C _{n41}	...	C _{N41}
		U ₄₂	w ₄₂	C ₁₄₂	C ₂₄₂	...	C _{n42}	...	C _{N42}
		U ₄₃	w ₄₃	C ₁₄₃	C ₂₄₃	...	C _{n43}	...	C _{N43}

攻读学位期间取得的研究成果

- [1] 蒋鸿鹏. 建筑企业成熟度模型的创建与应用研究.[J]. 建筑工程技术与设计.2014,(12):359-360.
- [2] 蒋鸿鹏. 基于模糊综合评价的项目管理成熟度应用研究[J].工程技术.2015,(1):143-144.