

THE FORELAND OF
TRADING TECHNOLOGY

内部资料 免费交流
《准印证》编号沪 (K) 0671

交易技术前沿

2019年 第四期 总第37期

本期主题
金融科技

No.4



上海证券交易所
SHANGHAI STOCK EXCHANGE

ITRDC
证券信息技术研发中心 (上海)

内部资料 2019 年第四期（总第 37 期）
准印证号：沪（K）0671

NO.4

主管：上海证券交易所
主办：上交所技术有限责任公司
总编：黄红元
副总编：徐毅林
执行主编：王泊、陈治纲
责任编辑：黄俊杰、徐丹、郭望

上海市浦东南路 528 号
邮编：200120
电话：021-68813289 021-68800293
投稿邮箱：ftt.editor@sse.com.cn



篇首语

1990 年底，伴随着黄埔江边的开锣声，上交所开业完成了第一笔电子交易。当年，中国证券业先驱大胆放弃口头竞价方式直接跨入电子化领域，为技术赋能金融的时代敲响了余音不绝的锣声。今天，大数据、云计算、人工智能和区块链技术正在以前所未有的力度驱动着行业变革，不断创造出新的业务模式、应用、流程和产品。本期《交易技术前沿》以“金融科技”为主题，从科技评价、交易撮合、技术风险、数据安全、业务支持等方面洞察行业前沿，聆听历史回音。

其中，《智能科技评价体系研究之专利维度》分析了准确评价科创板上市企业“硬科技实力”的难点，提出以专利评价为基础、同业对比为核心的智能科技评价体系，通过聚类技术方向分布相似的企业，准确评价企业在细分行业的技术实力。《基于分布式架构打造兼顾稳态和敏态的证券交易新核心》介绍了分布式业务架构兼具稳态和敏态的优势，详述东方证券通过“小核心、大外延、分布式”的模式实现业务解耦，并成功切换分布式架构的开发实践。《证券行业金融科技风险若干问题思考：技术视角》指出当前金融科技风险防范工作集中在宏观层面的问题，详细剖析了技术层面的相关风险，同时探索了算法可解释性、数据隐私性、监管沙盒领域等场景中的风险防范实践，最后提出了相关规划建议。《安全多方计算与证券行业数据生态建设》结合当前行业数据生态建设面临的问题，以区块链技术为载体，结合安全多方计算、分布式身份标识 DID、可验证授权凭证等技术，构建了可信行业数据生态系统，促进数据安全流通，释放数据价值。《全球交易所行业区块链最新发展及思考》对全球交易所行业的区块链发展情况进行综述，并对我国证券行业区块链发展提出建议。

今天，随着资本市场改革进入深水区，行业要不忘初心牢记使命，努力学习先驱们攻坚克难的勇气，传承技术创新的精神，为资本市场稳定发展保驾护航。

潮平岸阔风正劲，扬帆起航正逢时！

《交易技术前沿》编辑部

2020 年 1 月 2 日

目录 Contents

本期热点 Hotspot

- | | |
|-----------------------------------|----|
| 1 智能科技评价体系研究之专利维度 / 李一鸣 余勇 黄俊杰 郭望 | 4 |
| 2 基于分布式架构打造兼顾稳态和敏态的证券交易新核心 / 王海东 | 11 |
| 3 证券行业金融科技风险若干问题思考：技术视角 / 廖倡 | 16 |
| 4 安全多方计算与证券行业数据生态建设 / 王毛路 俞枫 | 23 |

实践探索 Explore

- | | |
|---|----|
| 5 运维智能服务台的探索和实践 / 杨超 吴浩 耿锋 李沛臻 | 34 |
| 6 基于业务调度框架的证券业信息系统应用研究 / 周靖 任荣 王东 王洪涛 | 42 |
| 7 运用深度学习技术处理投行项目富文档的探索与实践 / 陈贇 晏强 周朝阳 吴波 罗平 | 49 |
| 8 证券公司信息系统用户权限集中管理的研究与实践 / 刘国杨 王海航 | 56 |

行业观察 Observation

- | | |
|----------------------------------|----|
| 10 全球交易所行业区块链最新发展及思考 / 徐广斌 牛壮 徐丹 | 65 |
| 11 券商证券交易系统质量评估框架 / 何志东 朱立 邓博 | 77 |

信息资讯采撷 Information

- | | |
|----------|----|
| 监管科技全球追踪 | 89 |
|----------|----|



H 本期热点 Hotspot

- 1 智能科技评价体系研究之专利维度
- 2 基于分布式架构打造兼顾稳态和敏态的证券交易新核心
- 3 证券行业金融科技风险若干问题思考：技术视角
- 4 安全多方计算与证券行业数据生态建设

智能科技评价体系研究之专利维度

李一鸣 余勇 黄俊杰 郭望 / 上交所技术有限公司

科创板的推出为准确评价企业科技实力带来了新的挑战。传统评价体系大多基于通用指标，缺少对核心技术的精准研判，而科创板定位于突破关键核心技术的创新企业，投资者需要对种类繁多的技术领域、日新月异的技术前沿有深入的了解。为解决科技评价现存的困难，我们提出建立多维度的智能科技评价体系，本文仅针对“专利维度”重点介绍。专利维度以专利评价为基础，同业对比为核心，辅以技术前景、专利对比、专利核验的多维度分析，全面展现企业在细分行业的真实竞争力。实验显示在行业排名任务上，算法输出与人工判断基本一致，且具有较强的可解释性。



1 引言

2019年7月科创板在上交所正式开市，首批25家企业来源于信息技术、高端装备、新材料、新能源、节能环保、生物医药等高科技行业。科

创板旨在支持“主要依靠核心技术开展生产经营”的创新企业在境内上市，准确的企业科技评价对投资决策有重要意义。

近年来，以BAT为代表的科技公司为我国经济发展注入了独特的活力，企业科技评价得到

了更多研究，现有评价体系大多从四个方面的指标分析创新企业成长性，包括：

(1) 技术创新

研发人员比例、研发投入效率比、新产品收入、专利数、奖项等。

(2) 人力资本

雇员人数变化、管理者素质、员工素质等。

(3) 市场能力

市场占有率、目标市场增长率、营收平均增长率等。

(4) 财务能力

偿债、营运、获利、成长等。

然而科技公司立足于关键技术的突破，通用指标体系缺少对核心技术量体裁衣的考量。为了准确评价企业在专业领域的竞争力，投资者需要掌握大量专家知识，深入了解日新月异、门类繁多的技术前沿，准确分辨披露信息中夸大其词、避重就轻的部分。

为了解决科技评价现存的困难，我们提出建立多维度的智能科技评价体系，包括专利维度、研发力量维度、财务维度、商业模式维度、产品服务维度、管理运营维度等，本文仅就专利维度重点介绍。科技评价的专利维度以专利评价为基础，同业对比为核心，辅以技术前景、专利对比、专利核验的多维度分析，全面展现企业在细分行业的真实竞争力。专利是公司核心技术成果的体现，具有客观、真实、准确、完整的特点，借助专利大数据与自然语言处理对企业知识产权进行全面分析，监管审核人员可以迅速了解企业的技术方向分布和核心技术实力，降低对专家知识的

依赖。在专利评价的基础上，将技术方向相似的同业公司聚类 and 排名，可以清晰定位科技公司的相对实力。

本文将在第 2 节中介绍专利评价算法，在第 3 节介绍同业对比算法，包括行业分类和行业排名方法，在第 4 节进行实验验证，在第 5 节介绍其他四个分析模块，最后进行总结。

2 专利评价

专利评价算法基于国家知识产权局《专利价值分析指标体系操作手册》，结合科创板申报行业特点，从专利质量、技术价值、市场价值、社会效应 4 个方面为每个专利的技术与经济价值打分（图 1）。

具体来说，专利质量包括权利要求属性、技术影响力、稳定性等指标，技术价值包括创新程度、技术宽度、生命周期等指标，市场价值包括投融资情况、许可转让情况等指标，社会效应包括法律属性、焦点事件、政策法规等指标。

给定一家科技公司，专利评价模块计算出所有专利的评分（图 2 中），输出评分的分布情况（图 2 左），并给出每项指标的具体分数和评价标准（图 2 右），提供充分的解释性。

3 同业对比

科技评价要结合企业所在行业进行分析，选择合适的参照系才能准确定位企业发展现状，预测未来走势，了解产业发展规律，但是传统的人

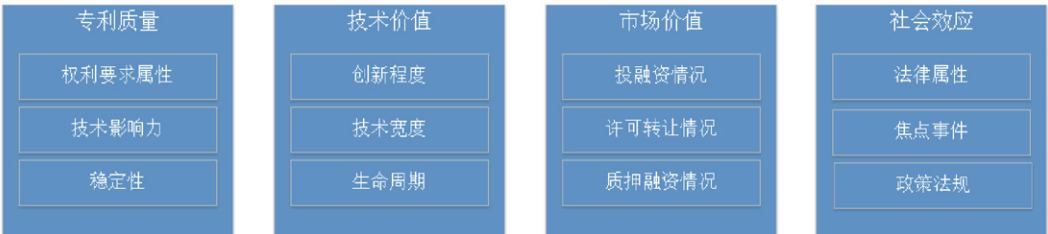


图 1 专利评价指标

重要专利



图 2 专利评价示例

工行业分类方法存在许多局限性,不能满足科创板高新技术领域的需求。为了精确评价科技公司在细分领域的相对实力,我们在专利评分和专利分类的基础上,通过智能行业分类算法(3.2)动态构建目标公司的同业公司池,并对企业科技实力评分(3.3),最后结合两者得到行业排名(3.4)。

3.1 行业分类现状

精确的行业分类是同业对比的前提,对于有限数量的企业,传统方法大多通过人工进行,典型的分类标准包括证监会分类和申万行业分类。然而在经济高速发展的今天,新兴产业不断涌现,高新技术日新月异,通过人工跟踪海量企业的技术方向和发展动态面临许多困难:

(1) 效率低

人工分类需要对企业 and 所属行业有深入的研究,而中国工商注册企业数量达到千万量级,人工分类杯水车薪。现阶段应用广泛的证监会、申万行业分类仅对 3000 多家上市公司进行了分类。

(2) 粒度粗

现有行业分类最细的仅有数百类,随着社会分工不断细化,科技竞争大多发生在极窄的行业范围内,传统行业分类已经无法适应高新技术领域的需求。

(3) 跨行业问题

许多企业的经营范围涉及多个行业,传统行业分类不能清晰确定企业技术方向的分布和核心

技术。

(4) 实时性低

科创企业引领着产业技术前沿的进步,根据研发和市场情况,企业的发展方向比传统行业变化更快,人工分类难以及时跟踪企业技术方向动态。

3.2 智能行业分类算法

为了解决传统行业分类存在的困难,我们提出基于专利大数据的智能行业分类算法,实时分析企业在多个技术方向上的专利布局情况,基于 6000 多个技术方向聚类同业公司,为准确评价企业在细分领域的相对实力打下基础。

具体来说,给定目标公司,首先统计所有专利的 IPC 分类,汇总得到公司的技术方向分布作为企业特征向量(3.2.1)。接着通过两轮筛选构造同业公司池,第一轮基于专利相似度粗筛(3.2.2),第二轮基于技术方向分布相似度细筛(3.2.3)。

3.2.1 技术方向分布

将技术方向分布表达为一个高维向量 w_k ,每个维度对应 IPC 分类的一个大组,共计 6000 余维。向量中的元素为每家企业该 IPC 分类大组的专利数量与该企业专利总数的比值,即:

$$w_k = [\frac{p_{1,k}}{tp_k}, \frac{p_{2,k}}{tp_k}, \dots, \frac{p_{l,k}}{tp_k}]$$

其中, w_k 为企业 k 的技术方向分布, $p_{i,k}$ 为

企业 k 在第 i 个 IPC 大组的专利数量, tp_k 为企业 k 拥有的专利总数。



图3 技术方向分布（某长三角企业）

以某长三角企业为例（图3），通过基于专利大数据的技术方向分布计算，可以得出核心技术集中在“控制来自独立光源光属性的器件装置 / 指示器连接的控制装置 / 电性能测试、电故障探测装置”，可见该企业主营业务为电子屏幕测试。而国民经济行业分类为“专业设备制造”，wind 行业分类为“电子设备和仪器”，由于分类过粗无法确定企业定位。

3.2.2 基于专利相似度粗筛

在粗筛阶段，以相似专利为桥梁将目标公司与同行业公司沟通起来。粗筛与细筛结合的模式，可以减小细筛阶段的数据规模，降低计算开销。由于在专利评价算法中也需要计算所有专利间的相似度，粗筛并不会增加额外开销。

(1) 构建专利特征向量

专利特征向量通过统计专利文本中的关键词计算。具体来说，将文本分词建立词库，使用 TF-IDF 评估每个词的重要性。单词 i 在专利 j 中的 TF-IDF 值由 TF 和 IDF 两部分构成：

$$\begin{aligned} w_{i,j} &= \text{TF} \times \text{IDF} \\ &= \frac{tf_{i,j}}{n_j} \times \log\left(\frac{N}{1 + df_i}\right) \end{aligned}$$

其中 TF (Term Frequency) 为词频, $tf_{i,j}$ 表示单词 i 在专利 j 中出现的次数, n_j 表示专利 j 的总词数, 词频越高单词 i 越重要。IDF (Inverse

Document Frequency) 为逆文档频率, N 表示全球专利库中的专利总数, df_i 表示全球专利库中包含单词 i 的专利数量, IDF 较高表示单词 i 在全球专利库中比较少见, 将带来较大的信息量。具体来说, 有些词在当前专利 j 中出现频率高却不重要, 例如“根据”、“特征”、“的”、“是”在整个专利库中很常见, 因此没有带来信息量, 此时 IDF 值较小。而“电性能测试”、“光学透镜”在整个专利库中比较稀少, 在当前专利中出现两三次就能带来较大的信息量, 代表了专利的技术主题, 此时 IDF 值较大。

将专利文本中每个词的 TF-IDF 值排列为向量, 得到专利 j 的特征向量：

$$V_j = [w_{1,j}, w_{2,j}, \dots, w_{n,j}]$$

(2) 基于专利相似度筛选同行业公司

在专利数据库中检索与目标公司所有专利相似的其他专利, 每篇专利会对应一个申请人 (公司), 这些公司组成了粗筛结果 A 。

专利的相似度通过余弦距离度量：

$$D_{i,j} = \frac{V_i \cdot V_j}{|V_i| \times |V_j|}$$

其中 $D_{i,j}$ 为专利 i 和专利 j 相似度, 只保留相似度大于一定阈值的公司。

3.2.3 基于技术方向分布相似度细筛

粗筛结果中存在许多与目标公司技术方向分布重合度较低的公司, 这是由于专利点对点比较的原因。以国家电网为例, 其专利多、覆盖面广, 在信息技术方面也有涉及, 但是国家电网与阿里巴巴并不是同行。因此, 在细筛阶段更关注公司间整体技术方向分布的一致性。

在粗筛结果的基础上, 保留与目标公司技术方向分布相似度高于一定阈值的公司, 相似度通过余弦距离度量：

$$S_{i,j} = \frac{W_i \cdot W_j}{|W_i| \times |W_j|}$$

其中 $S_{i,j}$ 为公司 i 和公司 j 技术相似度, 只保留相似度大于一定阈值的公司。

3.3 公司评分算法

做好行业排名的前提是准确的行业分类和公司评分，本节将介绍基于专利的公司评分算法。

3.3.1 评分计算

公司评分由技术实力（专利质量）和公司规模（专利数量）两部分加权得出：

$$C = \alpha \frac{\sum_{i=1}^{a \cdot n} S_i}{a \cdot n} + 2\beta \left(\frac{1}{1 + e^{-\frac{a}{b}}} - 0.5 \right)$$

其中 C 为公司技术分， α 为专利质量权重， β 为专利数量权重。 S_i 为第 i 篇专利的价值分， n 为专利总数， a 为参与计算的专利数量比例， $a \cdot n$ 是最能代表公司技术实力的一部分高质量专利（例如评分前 20% 的专利）。 b 为专利数量敏感度系数，该数值越大边际效应递减。

举例来说权重 α 和 β 都为 0.5，参与计算的专利比例 a 为 20%，专利数量敏感度系数 b 为 1000 时：假设公司 A 有 1000 篇专利，质量前 20% 的专利平均分为 0.5 分，则这家公司的最终得分为 0.4700 分。又如一家公司 B 有 30 篇专利，质量前 20% 的专利平均分为 0.4 分，则这家公司的最终得分为 0.2076 分。

3.3.2 参数选取

算法只选取部分专利参与计算（ a 为 20%），有两点原因：

(1) 公司核心技术实力主要体现在高质量专利，其数量较少。

(2) 为了保证核心专利安全，公司常采用星型布局，围绕核心专利注册一系列保护性专利，这些专利普遍质量不高，但是起到扩大保护范围的作用。

3.4 行业排名

3.4.1 综合排名

基于同业公司池（3.2）和公司评分（3.3），可以确定目标公司的综合排名。以某光学公司 X 为例（图 4 中第三名），公司打分为 61 分，同业公司池经过粗筛后技术相关度在 90% 以上的企

业有 9046 家，细筛后技术相关度在 70% 以上的企业有 51 家，公司 X 排名为第 3 位。

技术榜单

国内公司

上市公司

3 / 51

0.7

Q

序号	公司名称	相似度 ▾	分数 ▾
1	████████ 光学有限公司	0.91993	74
2	████████ 光 ████████ 有限公司	0.971452	62
3	████████ 光 ████████ 有限公司	1	61
4	██	0.733504	54
5	██	0.701899	54
6	████████ 光学 ████████ 有限公司	0.988723	53

图 4 技术榜单（某光学公司 X）

综合排名算法体现了公司的整体实力，但是同业公司间的技术布局各具特点，整体上落后的公司可能在某个技术方向上表现十分突出，因此我们进一步比较同业公司在每个技术方向上的排名，体现不同企业核心技术布局的差异。

3.4.2 技术方向排名

技术方向排名（图 5）显示了同业公司在每个技术方向上的实力分布，每个方向用不同颜色的曲线表示，横轴代表名次，纵轴代表公司在该方向的评分。图中可见光学元件领域在多个子方向上都由几家行业龙头把控，第一名的公司 Y 优势明显，第二三四名实力相近。

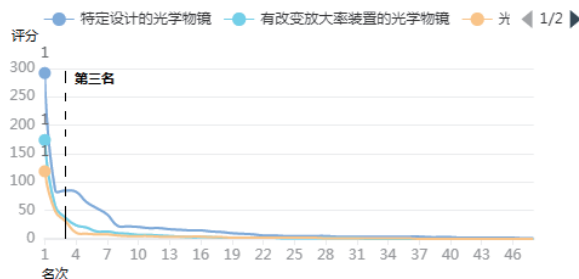


图 5 技术方向排名（某光学公司 X）

4 实验验证

4.1 人机对比

实验中从 98 家企业中抽检了 8 家企业，将算法输出与审核员主观判断比较，结果显示两者基本保持一致（表格 1）。

表格 1 预测企业排名

公司名称	同业对比算法	人工判断	一致性
公司 A	前 5%	先进	✓
公司 B	前 5%	先进	✓
公司 C	前 5%	先进	✓
公司 D	前 5%	先进	✓
公司 E	前 5%	先进	✓
公司 F	前 10%	较先进	✓
公司 G	前 40%	较差	✓

4.2 案例分析

我们从表格 1 中选取了排名不同的两家企业（公司 E、公司 G），并从专利数据角度解释算法输出。从表格 2 和表格 3 中可见，前 5% 的公司 E 主营 3D 打印，在其主要技术方向上都有较高的排名，且专利量大质高。而前 40% 的公司 G 主营网络安全，在其主要技术方向上排名均比较落后，且大部分专利质量较低。

表格 2 技术方向排名

公司 E			公司 G		
技术方向	专利数量	行业排名	技术方向	专利数量	行业排名
金属粉末制造的工件	107	前 1%	数据交换网络	11	前 30%
增材加工	20	前 2%	程序控制设计	5	前 50%
金属粉末处理	11	前 6%	保护计算机装置	2	前 50%

表格 3 专利分布

公司 E			公司 G		
专利类型	专利数量	数量占比	专利类型	专利数量	数量占比
高质量专利	14	7%	高质量专利	0	0%
一般专利	98	53%	一般专利	43	91%
低质量专利	75	40%	低质量专利	4	9

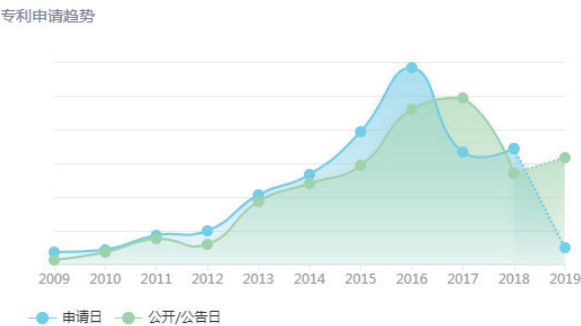
5 其他分析模块

除了同业分析，智能科技评价体系在专利维

度上还提供了包括技术前景、专利对比、专利核验多方面的分析。

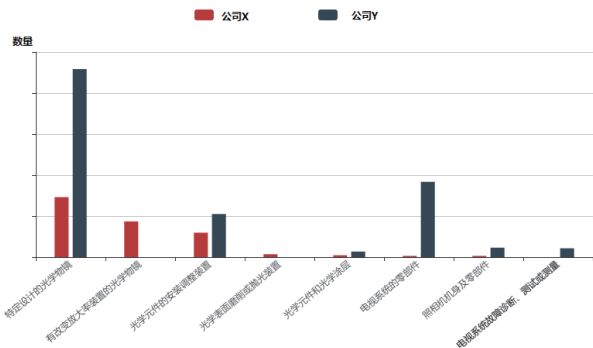
5.1 技术前景

同业分析体现了现阶段的技术水平，但是对于科技公司而言，技术发展的趋势更加重要。图 6 中某光学公司 X 的专利申请量在 2009-2016 年处于上升期，在 2016 年达到顶峰，2016 年以后迅速下降。尽管在行业排名第三，但是近年来发展放缓，创新能力下降，可能是行业技术前景不佳或企业自身发展问题。



5.2 专利对比

专利对比包括数量和质量两个维度，可以选择两个同行业公司一对一对照。我们比较了第三名的公司 X 和第一名的公司 Y（图 7 和图 8），发现公司 Y 专利数量更多，技术布局更广，在大部分技术方向上的实力都超过了公司 X，但是公司 X 在“有改变放大率装置的光学物镜”方向有明显优势。



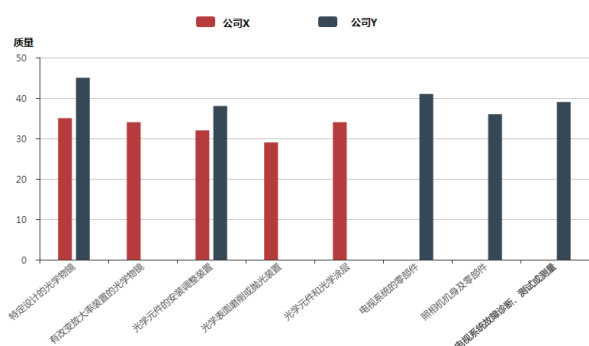


图 8 专利质量对比

5.3 专利核验

专利核验为审核人员提供招股书知识产权信息与专利数据库信息的交叉验证,包括专利数量的一致性(表格 3),专利名称、专利类型、专利号、授权时间、到期时间的一致性。

招股书是含有表格、图片和多种文字段落的富格式文档,常用非结构化形式表示数据(例如 PDF 格式),需要使用图像分析和自然语言处理技术解析文档结构,并抽取专利章节的关键词。以图 9 为例,基于目标检测和语义分析方法定位特定字段的位置并用虚线框标出,然后汇总形成结构化数据保存。

表格 4 招股书与专利数据库的交叉验证(某江苏企业)

专利类型	专利数据库(申请数)	专利数据库(有效数)	招股书披露	一致性
外观设计	6	4	2	✓
实用新型	70	58	49	✓
发明专利	107	23	20	✓

序号	专利名称	专利类型	专利申请日	专利号	专利权人	取得方式
1	一种支持定位的复合型散热器	实用新型	2015.04.21	ZL201520249360.5	麦家科技	申请取得

序号	专利名称	专利类型	专利申请日	专利号	专利权人	取得方式
1	一种支持定位的复合型散热器	实用新型	2015.04.21	ZL201520249360.5	麦家科技	申请取得

图 9 招股书专利字段抽取

6 总结

科创板的设立使企业科技评价从通用指标向个性化发展,准确分析目标企业在细分行业内的核心技术实力变得尤为重要。我们提出了多维度的智能科技评价体系并重点介绍了专利维度。专利维度基于专利大数据将技术方向相似的企业动态聚类,在细分行业精确对比,并提供技术前景、专利对比、专利核验的全面分析,帮助投资者提升效率,降低对专业领域知识的要求。

但是强调专利价值的科技评价方法仍然存在一些不足,一方面许多初创企业发展迅速,亟需资本助力进一步发展,但是专利公开存在一定滞后性,现有方法对这类企业并不公平;另一方面部分企业的主营方向发生多次调整,与企业所有专利中体现的技术方向有一定偏差。为了解决存在的问题,未来我们将探索更立体的评价体系,尝试多种专利评价模式结合,助力资本市场更公平高效地服务创新企业。

基于分布式架构打造兼顾稳态和敏态的证券交易新核心

王海东 / 东方证券股份有限公司



证券交易系统是证券行业最重要的业务系统，集中体现了一家证券公司的综合技术实力。近几年，随着个股期权、港股通、CDR、科创板等业务的不断推出，以及市场客户和交易业务量的不断攀升，对证券交易系统的业务支持速度、并发及吞吐量、处理时延、安全性、稳定性、可扩展性、敏捷性等等均提出了更高的要求。因此，打造兼具稳态、敏态的证券交易新核心是保持证券公司核心竞争力的关键。

作为国内最早关注和践行分布式技术路线的大型综合券商，东方证券从 2016 年 12 月开始规划设计，2018 年 5 月完成了兼具双态特性的证券交易新核心的构建，分布式交易系统正式投产上线。经过一年多时间的运行检验，系统运行稳定，与集中交易系统对比，分布式交易系统实际交易峰值并发量提高了 5 倍，上行全链路平均时延快 100 倍，高可用主备切换耗时缩小到百分之一，节点扩展部署速度提升 20 倍，业务开发周期缩短了一半，单节点总拥有成本不到十分之一，得到了客户和业务部门的高度认同，提高了东方证券在机构业务市场的知名度和竞争力，取得了令人满意的业务效果。

一、稳态和敏态

Gartner 最早提出了稳态和敏态的理念。所谓“稳态”，强调要保持现有各业务系统的安全稳定运行，对证券行业而言，这是所有技术工作的重中之重。所谓“敏态”，强调敏捷、高效的进行各种业务创新，这对金融科技广泛应用的今天，尤其是随着移动互联网、人工智能、大数据、分布式等现代信息技术应用场景的不断落地，显得尤为重要。

交易系统能否中长期保持稳态，除后期运维人力资源保障外，更重要的取决于它所依赖的基础架构，东方证券基于分布式技术高可用特性构建的系统，攻克了双活高可用场景下零数据丢失的难题，采用双活设计，交易节点故障切换可以做到零数据丢失，并且能在 10 秒以内完成切换，大幅提升了交易系统的高可靠、高可用水平，保证了中长期的稳态。

交易系统要保持敏态，首先要解决系统间、功能间的耦合问题，其次要解决系统架构的水平可扩展性问题，最后要解决系统对单台服务器的依赖性问题。东方证券基于分布式技术低时延特性，将业务逻辑与高可用解耦，大幅简化了业务开发难度，降低了交易业务系统风险，能够快速响应业务创新，缩短业务实现周期。

稳态和敏态虽然目标不一样，但两者并不矛盾，而是相辅相成的关系。对证券行业而言，两者缺一都难以应对激烈的同业竞争，打造支持“稳态+敏态”双态运营的系统才是王道。

二、集中式架构的痛点

基于传统集中式架构的交易系统，越来越难以适应未来证券行业的发展，存在诸多痛点。

一是集中式架构的业务和组件高度耦合、难以分拆，扩展性差、牵一发而动全身。二是集中式架构存在较大延时，处理速度一般在毫秒级，

难以满足快速交易的客户诉求；三是集中式架构相对封闭，对国外供应商的软硬件产品依赖性过强，自主选择余地小且成本高昂；四是集中式架构过于依赖内部代码逻辑，修改代价较高，且难以支持二次开发，不利于行业自主创新能力的提升。

综上，对于稳定性要求第一的证券行业，集中式交易系统的每次系统修改和升级都如履薄冰，使得系统同时丧失稳态和敏态。东方证券较早感知了集中式架构对未来业务发展的制约，较早开展了分布式技术转型，较好的融合了交易系统的稳态与敏态。

三、分布式架构的优势

分布式架构是下一代金融交易技术的趋势，这在证券行业已形成共识。分布式架构具有降成本、提速度、灵活及扩展性强、架构开放等特性，更适合中国证券市场的特点，也是实现证券公司核心业务系统自主掌控的最佳可行路径。在此，以“证券交易资金算法”的例子形象说明分布式架构的优势。

证券行业经过近 30 年的发展，尤其是 2005 后，各类业务层出不穷，进而衍生出各种交收模式。随着参与方的增多，进一步增加了交收的复杂度，使得保障投资者资金正确的资金计算方法越来越复杂，这个算法也正是经纪业务的核心业务算法之一。传统集中交易的资金算法逻辑主要是构建在资金表的资金字段上，一般会把同一种性质的资金用同一个字段表示，然后编写业务逻辑进行计算，交收的顺序和对手方的特性等都在业务逻辑中实现。每次业务修改都可能要修改相关的字段和业务逻辑，修改后可能影响所有使用该字段和逻辑的业务。为此我们从业务架构的角度上考虑改变，尝试把资金算法构建成一个即稳又敏的实现。首先，我们把资金算法从业务逻辑中剥离出来，形成

多个小核心。每个小核心实现一类资金性质的计算，比如竞价可用、港股可用、报价回购可用、银证可取等等；其次，每个小核心之间按规范进行对接，这样每类业务发生变化的时候也就只需要修改该业务的资金算法和业务逻辑，可以预见其修改的业务范围会相对小，修改也相对简单，而且在调用逻辑充分清晰的情况下比较容易分析修改并进行精准的验证。

上述案例的解决方案是典型的“小核心、大外延、分布式业务架构设计思路”，按照这种思路设计的业务架构，可以解决紧耦合带来的痛点，从而实现该业务场景对应系统功能的稳态和敏态共存。近观科创板业务的技术实现，相关的权限规则、盘后业务等逻辑，在分布式架构下都是相对独立的业务实现，完全不影响原有业务功能。而基于传统集中式架构则只能修改原有核心交易功能来实现业务，设计复杂度、业务风险、维护成本都随之上升。

四、用分布式技术支撑分布式业务

基于传统集中式架构的交易系统之所以无法实现上述案例提及的功能，与其历史发展有关。集中交易系统的业务与技术架构始于本世纪初，最初的业务及技术架构与业务发展是匹配的。但随着业务品种扩充、业务规模的飞速增长，各业务模块紧耦合的架构弊端逐渐显现。因此促使我们利用分布式技术来进行业务解耦的实现，并降低系统高可用维护成本。

业务解耦的实现主要依赖两方面问题的解决，一是业务解耦规范的制定，二是解耦后原子业务相互调用的性能提升。业务解耦规范制定的重点是能否根据业务内容整理出清晰易用可扩展的核心业务条线，能否把不同业务模块间的切面处理的干净漂亮。这在传统集中交易业务架构初期是比较难的，目前随着业务的成熟以及系统设计人员经验的积累行业内已经具

备了将业务清晰解耦的条件。第二，对于原子业务相互调用的性能来说，由于传统集中式架构依赖于数据库实现业务的完整性，且当时业务逻辑梳理不够充分，导致一个业务调用往往会有多次同步的数据持久化动作，同时，原先网络性能也相对较弱、抖动较大，因此相互调用的性能大多在毫秒，甚至百毫秒级别，并行处理能力较差，不适用于对稳定性要求高的大压力场景。现在，基于高性能网络的低延迟方案已经可以让系统间的调用运行在微秒级，单节点并发处理能力达到每秒几十万笔，完全可以适配相关要求。

对于高可用的维护成本，传统集中式架构的交易系统依赖 IBM 小型机、物理数据库、高端存储等高可用技术，一套高可用系统的硬件成本动辄需要千万级。由于成本较高，因此往往把需要高可用的功能集中部署、集中运维。从而导致运行风险集中、业务创新实现缓慢的问题，这也正是传统集中式交易系统的重要痛点。值得庆幸的是，随着自主可控浪潮的推动，各类高可用技术发展迅猛，尤其是基于低延迟高可用的分布式技术，其硬件成本通常可控制在百万级。这让高可用系统的应用范围一下就可以扩展到很多方面，从而使需要高可用特性的系统可以分开部署实现业务解耦。

五、基于分布式架构打造证券交易新核心

有了第四章提及的业务和技术架构的准备，也就有了证券交易新核心的基础。分布式技术打造的证券交易新核心相比传统集中交易系统具有明显的优势。

传统集中交易系统的现状包括但不限于四点，一是在一个相对高可用的物理数据库内存放了账户、权限、参数、费率、资产以及流水等数据；二是投资者交易指令通常比较复杂，大多会

涉及到上述所有类型的数据；三是营运业务指令相对单纯，但一般也会涉及到上述两到三个类型的数据；四是清结算相关业务最复杂，通常会反复多次用到上述所有类型数据。

这样的逻辑一开始看上去似乎并不复杂，数据和业务分开，需要什么处理什么就可以了。然而这只是表面现象，实际情况是由于所有业务都依赖这些数据，一旦有新的业务需要改动时，很难辨识对其他业务的影响，前面所举资金算法就是一种场景。如果把数据库比作证券交易系统的核心，那么所有业务使用数据库的场景就如同万箭穿心，这种情况下稳态和敏态也就成了无法实现的幻影。

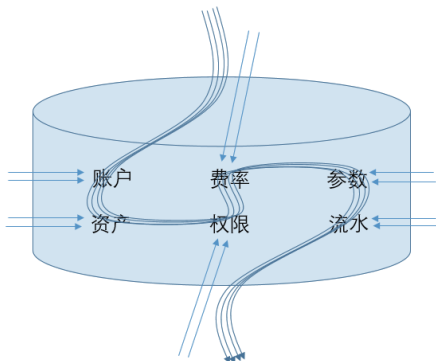


图 1

对比上述四点现状，通过分布式技术解耦后的系统状态如下：

一是业务上以营运管理为中心将系统分解成多类小核心，并实现业务核心的高内聚，避免过长的业务条线；



图 2

二是以每个小核心为中心结合相应的外延业

务形成子系统，其中引擎和服务可以根据业务分成多个，例如有竞价交易引擎、非交易引擎、大宗交易引擎等等；

三是每个子系统采用分布式的高可用技术保障稳定性，从而实现系统的稳态；

四是技术上子系统间使用分布式的低延迟高速消息总线进行数据交互，实现业务解耦，从而易于实现系统的敏态；

为更好证明业务解耦在技术层面的可行性，我们接下来谈一谈如何结合分布式技术和物理数据库实现数据一致性。这里我们在基于高性能的基础上，以资产计算解耦为例，着重从数据使用的特点进行分析，从而解决一致性问题。通常，要保障一致性往往要先解决持久化问题。本例而言，基于物理数据库所构建的高可用仍然是非常好的数据持久化工具。因此，我们这个例子中使用物理数据库作为持久化的一个场景提供业务管理和查询服务。

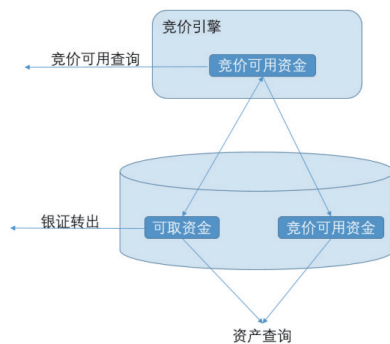


图 3

这里有两类外部指令需要特别关注，一是银证转账，一是资产查询。对于银证转出，其事务一致性要求往往较高，因此在转账时应直接从竞价引擎先行扣除再转出。而对于资产查询，如果只是针对总资产进行统计分析，一致性时延要求则不太高，约百毫秒级，这时我们使用的是竞价引擎同步下来的结果，在数据库与可取资金合并后输出的数据。此外，对总可取资金查询也可以这样实现，只有买入依赖的竞价可用查询，因为时延要求比较高，我们建议在竞价引擎中直接查

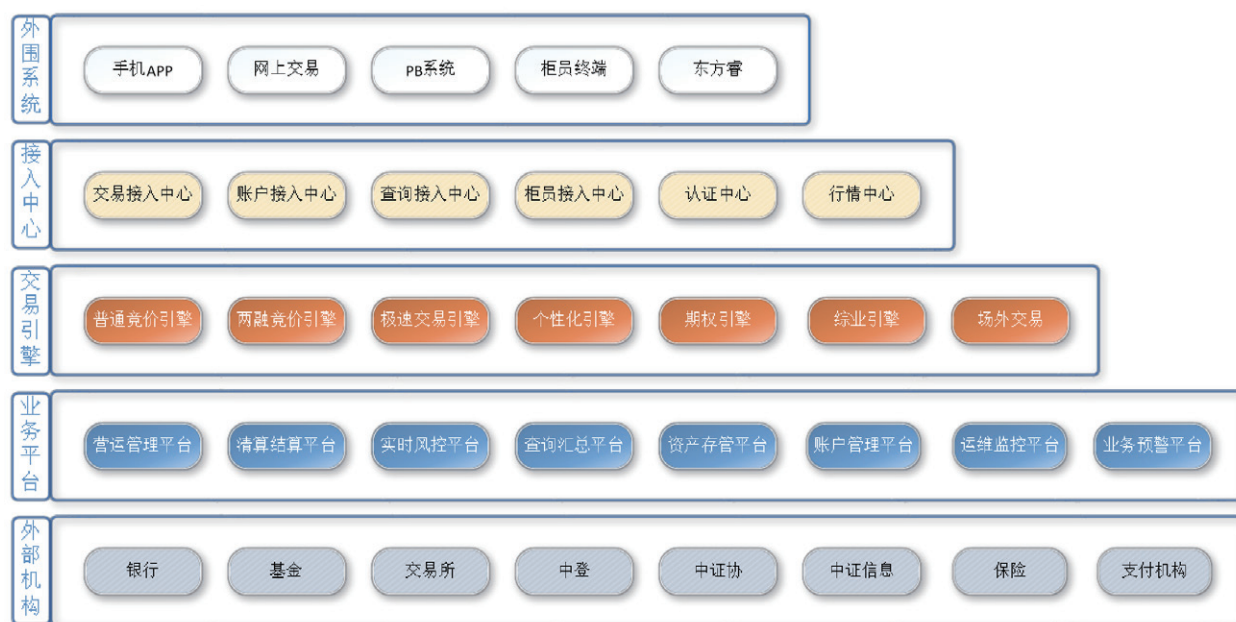


图 4

询。这种方案很好结合了分布式技术和物理数据库分场景实现了数据一致性，完成了资产的分布式计算。

六、总结

回顾系统建设过程，我们在对业内分布式架构进行了调研分析后，选择了与华锐金融技术的专业团队合作，一步一步完成了原型验证、系统设计和开发、上海节点上线、南方中心节点上线等多个里程碑，最终打造出既可靠、又快速的分布式交易系统。目前，系统已经支持了现货交易、新股发行、回购、科创板、信用交易等多项业务，并与账户系统、营运管理和清算系统之间实现了良好的交互，目前正在增

加更多的业务引擎。

除了低时延和高可用特性以外，分布式架构还具有很多优良特性，比如弹性可扩展、开放、平台化、易于自主掌控、易于国产化等。这些特点相对于传统技术而言，在系统的稳态和敏态建设中都有着不可替代的优势。

值得一提的是，东方证券分布式交易系统在 2018 年上线后，先后获得“《金融电子化》杂志评选的年度金融行业科技创新突出贡献奖 -- 开发创新贡献奖”、“深圳市 2018 年度香蜜湖金融科技创新奖 -- 优秀项目奖”等殊荣。

经过实践检验，我们深信基于低时延高可用分布式架构打造的证券交易新核心将成为解决传统集中交易痛点、实现稳态和敏态并举的优选方案。

证券行业金融科技风险若干问题思考：技术视角

廖倡 / 深圳证券交易所 证券期货业金融科技研究发展中心（深圳）

当前，金融科技正加速推动金融行业革新。然而，金融科技的发展还面临一些挑战，并存在相关的问题。本文首先回顾金融科技发展历程以及国内外关于金融科技风险若干探索，而后重点从技术角度对金融科技风险进行详细剖析。基于人工智能治理相关内容，结合金融科技若干实践，提出了技术层面相关研究和规划建议。



一、证券行业技术的发展历程与金融科技风险

本节首先总体介绍证券行业技术的进展，概述金融科技发展与相关风险，以及金融科技风险相关工作。

1. 我国证券市场技术发展历程

如图 1 (a) 所示，我国证券市场技术发展经历创立初步、快速成长、规范发展、创新发展

四个阶段。

20 世纪 90 年代，中国证券市场形成之初即抓住信息技术革命浪潮机遇，陆续实现了场内交易撮合电子化、登记结算无纸化、柜台系统电子化等业务流程，满足了市场初期的业务发展需要。

1993 年之后，我国证券市场经历了一个快速成长期。高性能的交易平台、卫星通信网络等一批新技术得以在市场上成功应用，不仅进一步推动了整个交易过程的全面电子化处理，而且也促使了全国市场及其技术体系的形成。

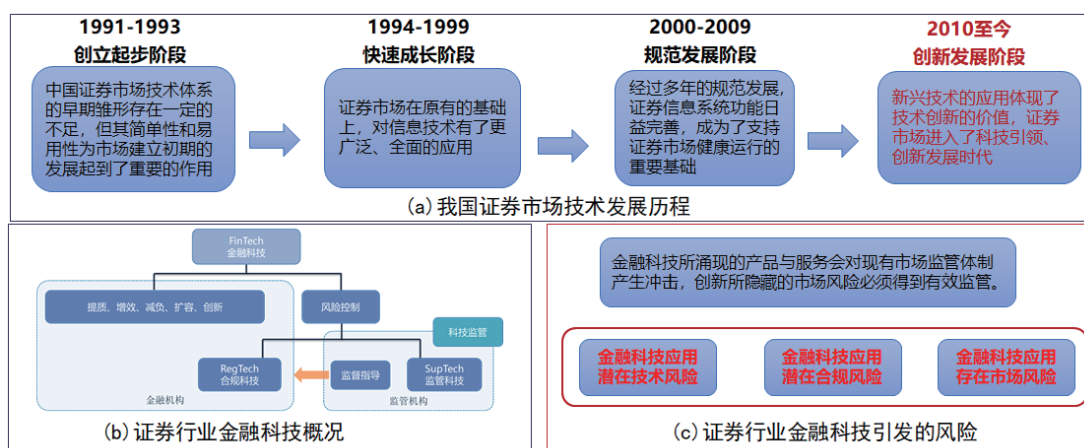


图 1：证券行业技术发展现状与金融科技风险若干探索

2000 年以后，中国证券市场开始进入规范发展阶段。多层次市场建设加快，产品与业务创新不断涌现，通过技术创新交易模式、拓展交易范畴，提高技术安全与风险控制成为这个阶段的主要任务。

2. 证券行业金融科技概况与风险概述

2010 年以来，证券市场进入科技引领、创新发展新时代，即金融科技。证券行业针对新兴技术应用开始了新的尝试，主要围绕优化业务流程、提升业务效能、降低运营成本、增强用户体验等开展创新应用。如图 1 (b) 所示，交易所等市场核心机构建设了一批行业信息化公共基础设施，为市场提供集约化云服务，降低了行业整体成本。同时，实体画像、知识图谱、文档分析等成为公司监管、市场监察、风险监测等监管业务，以及智能客服、智能投顾等经营业务的重要支撑。

金融科技作为一个跨界新兴产物，如图 1 (c) 所示，其快速扩张带来新的一系列风险^{[1][2]}，如金融科技的广泛连接带来的“网络效应”增加了金融体系的复杂性，可能会放大风险的传染性和影响面；金融相关法律法规没有充分考虑人工智能技术对金融业的影响，使人工智能算法模型处于监管真空或灰色地带；人工智能技术

直接应用于金融实际场景存在不可解释性、隐私性等问题。

3. 金融科技风险防范相关工作简述

2019 年，中国人民银行发布《金融科技发展规划》，指出要加强金融风险监测与评估。中国证监会也出台行业的金融科技指导意见，更加关注业务合规风险、信息安全风险和数据安全风险的防范要求。同时，面对蓬勃发展的金融科技产业，诸多国际组织、国家、地区陆续成立专门机构，对金融科技领域进行多角度研究，致力于探索更加完善的管理模式。新加坡金融管理局发布了一系列关于人工智能与数据分析的应用原则，以确保在金融领域使用人工智能和数据分析的公平性、道德规范、可问责性和透明度。如图 2 (a) 所示，监管沙盒¹由英国率先提出，以实现对金融科技的鼓励创新和风险防控。在这个空间里，公司或机构可以测试新的金融科技产品、服务、商业模式和交付机制，而不用承担常规监管的后果。

近期，业务层面的金融科技风险也开始逐步被资本市场所重视。如在智能投顾领域中网络的虚拟性、不同算法产生投资组合的复杂性问题，交易领域中高频交易等对传统金融交易标的、定价、流程和对价等产生巨大冲击将而引发公平性

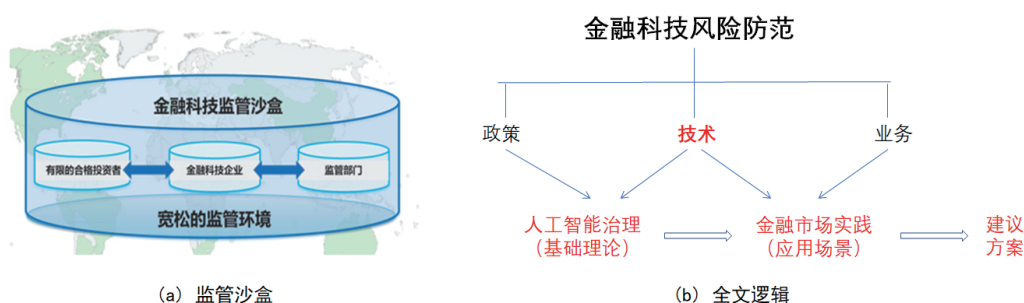


图2：监管沙盒与全文框架

等问题。此外，大型科技公司（Big Tech）涉足金融业务，有可能利用自身科技优势垄断某一项金融业务，实行混业经营，面临市场冲击时可能会出现羊群效应，加大系统性金融风险发生概率。

然而，绝大部分关于金融科技风险的工作集中于宏观层面，且大部分以政策研究为主，业务场景探讨次之，很少有工作直接从技术方面讨论如何防范金融科技风险。针对该问题，本文详细剖析技术层面的相关风险，进一步探讨通过技术手段防范的潜在可能。第二章从金融科技的理论基础出发，介绍人工智能治理研究相关工作；第三章从金融科技的应用场景出发，介绍金融场景中若干实践案例；第四章从技术角度提出金融科技风险防范的若干建议方案。全文框架如图2(b)所示。值得注意的是，本文为专注从人工智能技术治理角度防范金融科技风险，严格限定研究边界，加密资产、云计算、区块链等均不在本文研究范围。

二、人工智能治理回顾

金融科技发展的重要基础是人工智能，金融科技风险防范的相关思路可以借鉴人工智能治理相关内容。根据 Gartner 曲线²显示，如图3(a)所示，大部分人工智能技术的发展还处于萌芽阶段，但关注度极高，且人工智能治理处在最为关注的阶段。各国际组织、国家、专业协会纷纷发布相关政策准则，如图3(b)所示，致力于人

工智能治理发展研究。

欧盟委员会发布人工智能伦理准则，列出了“可信赖人工智能”的关键条件。在2018年5月25日，其出台《通用数据保护条例》（General Data Protection Regulation，简称GDPR），规定保护个人隐私，遏制数据滥用；要求一个系统必须具有“解释权”。本文第三部分将围绕这两点详细展开，并列举其在金融科技中的若干实践。

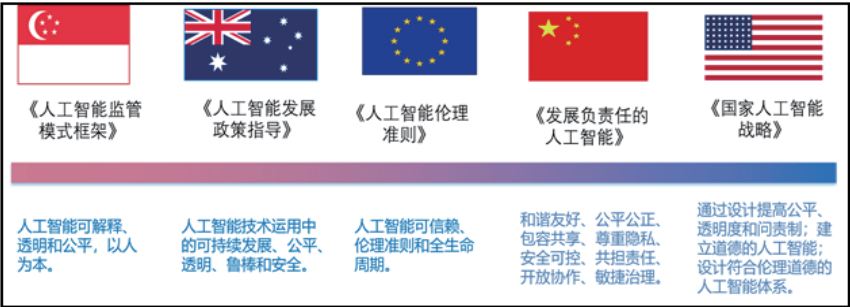
2019年，中国提出为促进新一代人工智能健康发展，确保人工智能安全可靠可控，推动经济、社会及生态可持续发展，共建人类命运共同体的若干原则。美国计算机协会针对算法歧视问题，发布关于算法透明和可责性的七条原则。除此之外，斯坦福成立“以人为本”的人工智能研究所、IEEE发布人工智能伦理标准等；亚马逊、谷歌、微软、Facebook等科技企业也设有相关部门以解决AI伦理问题，开发工具来检测算法偏见。

就具体技术层面而言，人工智能风险防范等相关内容开始被逐渐重视。如图4所示，其重要思路即建立算法风险管理体系，保证可解释性、隐私性、鲁棒性（在应用时不易受噪音或特定输入干扰）、以及公平性（不会对某类行为表现出无意识的偏好）等。近年来，在人工智能和数据科学的顶级会议（如KDD、ICML、NIPS、IJCAI、AAAI、WWW等）均设有相关的讲习班或讨论会。

与传统机器学习不同，深度学习直接从原



(a) 人工智能发展曲线



(b) 各国人工智能治理相关政策

图 3：人工智能治理回顾

始特征出发，其特征提取、选择、推理、预测等部分均为高阶复杂函数形式，缺乏明显物理含义，在实际应用中缺乏可靠性。目前，兼顾可解释性、隐私性、鲁棒性、公平性等要素的深度学习算法优化研究工作逐步增加。在模型鲁棒性方面，提出针对图深度学习模型的对抗攻击方法，是首个在属性图上考虑噪声的对抗攻击研究。在模型公平性方面，提出从“有偏差”用户反馈数据中训练“无偏差”的排序模型若干策略；总结机器学习的若干偏见，指出深度模型的普遍联系有可能进一步加剧歧视现象，并提出若干优化策略。

人工智能治理的若干研究成果可以被广泛借鉴应用于金融科技技术层面相关优化，如鲁棒的机器学习模型设计、少量的样本就能快速识别新类别（例如风险事件分类）的方法、数据去噪、偏差建模等。此外，关于建模用户行为和市场机制的多 agent 仿真建模在金融科技应用中具有广

阔空间。

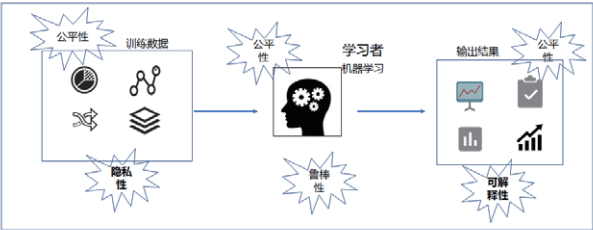


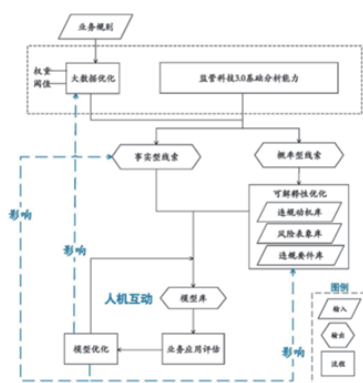
图 4：算法风险管理体系

三、金融科技风险防范若干技术层面实践

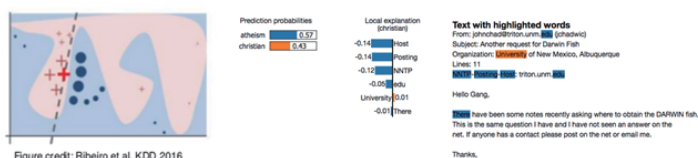
本节结合具体应用场景，重点列举在可解释性、隐私性层面若干工作实践，以及虚拟监管沙盒相关探索。

1. 算法方面的可解释性研究

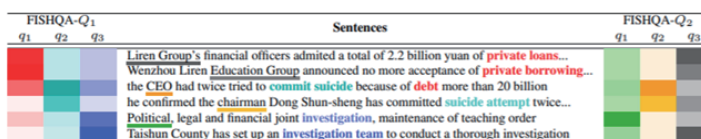
目前，兼顾可解释性要素的金融科技算法设计开始被逐步重视。由于算法的复杂性，将有可



(a) 可解释性分析在监管科技中实践



(b) LIME (local interpretable Model-agnostic explanation) 模型



(c) FISHQA情感分析模型

图 5：可解释问题相关内容介绍

能导致算法消费者、设计者、使用者以及机器之间存在信息不对称等问题。

在工程实践层面，如图 5 (a) 所示，监管科技 3.0 基于已有监管案例积累各类违法违规手法，构建违规动机库、风险特征库、违规要件库，与金融科技相关算法模型得到的概率型线索进行匹配，增强概率型线索的可解释性。例如深交所智能财务审计分析课题中，通过构建“财务舞弊动机—财务舞弊手法—财务舞弊表象”的线索链条，对财务舞弊倾向得分较高的对象进行分析，识别其违法动机、手段及表象，若形成完整的证据链，则认为该概率型线索可靠。此外，在某监管机构的市场运行风险监测智能化研究课题中，其试图整合相关信息对市场运行进行预判，同时拟结合事理图谱、事件链条等对发生异常的原因进行解释。

在学术研究层面，强调理解预测结果的重要性，提出 Local Interpretable Model-Agnostic Explanations 方法 (LIME：局部可解读的与模型无关的解释)，如图 5 (b) 所示，通过将输入值在其周围做微小的扰动，观察模型的预测效果，衡量各部分输入对预测结果产生贡献。提出一种基于双向循环神经网络的深度学习算法 FISHQA，引入层次化结构模型以及自动问答注意力机制 (Hierarchical Query-driven Attention)，如图 5 (c)

所示，有效对金融舆情的正负面情感进行识别，且对输出结果进行详细解释，有效克服情感判断依据不明确的问题。提出基于深度学习的深层次用户画像，在保留深度学习特征提取的优点的同时，结合监督/无监督学习方法实现用户分类/聚类，强化画像的特征可解释性，并能够进行后续分析。

2. 数据方面的隐私性研究

证券市场数据层面的工作主要集中于数据规范和标准化工作上，金融科技应用过程中围绕数据层面存在的隐私性、小样本性等方面要素相关工作都相对较少。同时，金融科技监管中数据的隐私保护开始逐步受到重视。

在学术研究层面，提出数据自治开放框架，如图 6 (b) 所示，通过在数据盒中基于区块链等相关技术封装数据防泄漏、数据权益保护和数据访问监控等机制，使之具有独立性、可用性、可控性。提出联邦学习框架，如图 6 (c) 所示，基于同态加密等相关理论建立把数据聚合在一起的虚拟模型，允许在不损害隐私的情况下进行知识共享，其提供一种证券行业经营机构与数据合作企业间的可能联合营销模式，确保数据在双方不出域的情况下实现联合建模，如推荐算法训练、客户分类，实现精准营销、产品推荐和权益定价等业务。

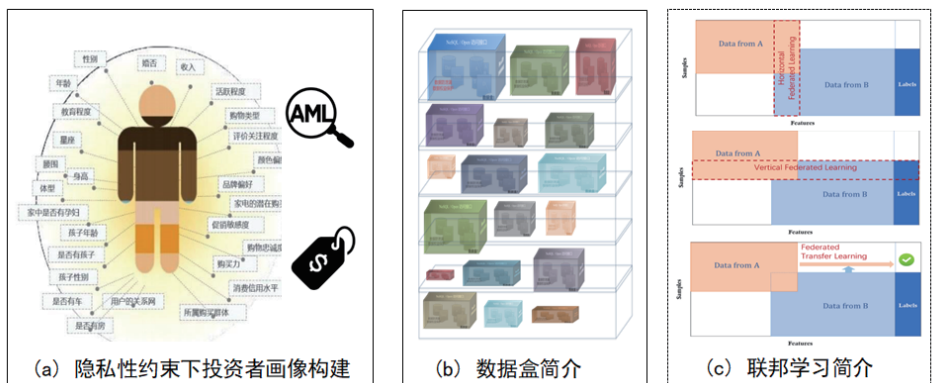


图 6：隐私性问题相关内容介绍

监管科技 3.0 常常面临数据属性维度不够高、数据样本不够多等问题。以“老鼠仓”的监管为例，通过异常交易检测发现某股票在多个交易日内存在异常波动定位关联账户，而后需要整合投资者其他相关属性信息形成对应线索链条。其中部分投资者数据属性需要来源其他部委、互联网科技公司等，如图 6 (a) 所示。针对此类存在的数据安全合规问题，一个可行的思路通过加密交换的手段，将各自拥有的同样用户的不同特征作为输入，迭代地进行训练模型、交换参数，通过建立一个数据“联邦”，可以让参与各方都获益，而又保证数据隐私。又如在某监管机构的企业画像项目后续规划中，试图基于联邦学习的思路，将存储在不同地点、不同机构的相关数据（如海关数据等）利用起来，对企业相关风险标签或缺失属性值进行预测或补全。

3. 虚拟监管沙盒相关内容

为检验相关金融科技运用对市场的影响，防范潜在风险，从技术角度出发，提出构建虚拟沙盒平台。针对金融科技创新的产品与服务，基于人工智能的模拟仿真测试环境被提出，以用于验证评估金融科技创新项目对现有市场带来的风险。虚拟沙盒的思路即基于计算实验金融学，如图 7 (a) 所示，是一种基于仿真的“情景 - 应对”型风险管理思想。国内中证监测在融资融券政策实施之前，曾借助仿真的手段开展了相关研究。深交所基于仿真技术开展“T+0”研究，依据 What-If Analysis 得到对相关市场影响分析结论。

虚拟监管沙盒的基本架构如图 7 (b) 所示，主要包括两个方面：(1) 模拟市场的各类模型，包括各类投资者模型、交易撮合模型、外部信息模型等；(2) 辅助仿真实验开展的工具，包括仿

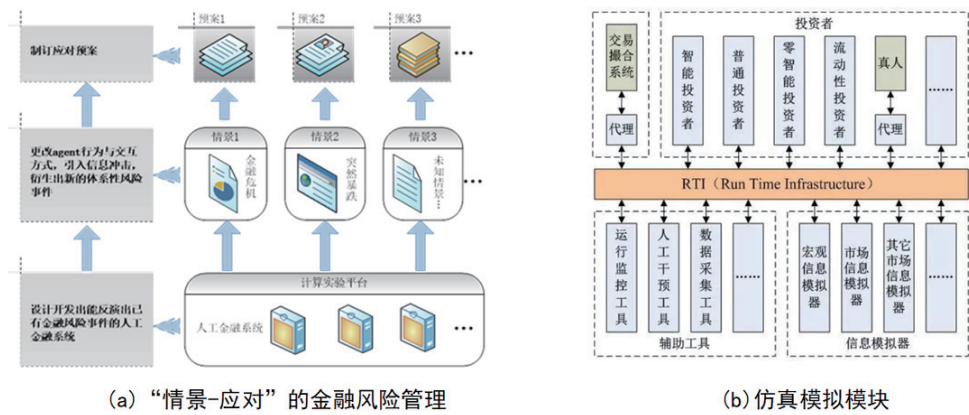


图 7：虚拟监管沙盒

真运行监控工具、人工干预工具、数据采集和管理工具以及数据分析工具等。虚拟仿真沙盒的目的即分析特定行为和交易机制要素对市场总体的影响，以及市场系统的动态变化规律。通过整合行业内外数据开发仿真模型，试图回答诸如制度设计、监管措施的效果分析、新产品的可行性分析、算法交易影响测试等等问题。

四、金融科技防范总结与建议

总体而言，金融科技在证券市场的应用已取得初步成果，但防范手段的发展仍落后市场的发展。这里从技术规划和研究角度，提出优化资本市场金融科技发展的路径：

1. 模型设计和优化方面的问题：

一是借鉴人工智能算法中的因果推断相关成果。研究因果推理在金融科技中相关应用，避免机器学习算法中简单的关联性导致决策不公平性的出现。

二是借鉴算法黑箱问题优化相关成果。研究兼顾决策过程理解的金融科技算法模型，同时加强金融业务知识在可解释性中应用，如完善资本市场执法证据链条等。

三是建立模型设计与优化生态。吸引监管机构、地方政府、技术企业、高校与研究机构、行

业组织等参与，研究金融科技的安全边界与创新路径。

2. 数据处理和融合方面的问题：

一是采用可行的技术性纠偏方法。针对主观数据（偏差数据）使用较多的业务场景或模型算法中，研究使用修正度量或相似度计算，将系统输出与期望的公平行为进行比较。

二是数据融合等方面的问题。针对数据易复制、价值难定量、渠道难管控等问题，研究如何利用区块链等相关技术使封闭的数据开放流通，研究数据自治开放理论、定价理论。

三是小样本学习研究。目前金融领域的机器学习运用普遍存在样本稀疏的问题。可以研究尝试小样本学习、零样本迁移学习方法等在具体金融场景的实践。

3. 虚拟监管沙盒方面的问题：

一是规划构建大规模自主 agent 的计算实验模型。包括但不限于信息、资产、投资者类型、交互方式、交易机制、时间维度等，研究金融系统复杂演化规律。

二是规划研究市场各主体复杂行为、机制建模。例如在仿真平台中，嵌入固定策略和智能策略的虚拟交易者，以及精确建模资本市场市场信息披露模拟器、控制器等部件。

注释：

¹ <https://www.fca.org.uk/firms/innovation/regulatory-sandbox>

² <https://www.gartner.com/document/3953603?ref=solrAll&refval=236757670>

作者简介：

廖倡，复旦大学博士，现就职于深圳证券交易所，研究方向为数据科学、监管科技等。

另注：本文某些观点来自互联网，限于篇幅限制，无法一一列举参考来源，特此表示感谢。此外，文中仅代表作者个人观点，仅供研究参考，文责自负，不代表所在单位观点。

安全多方计算与证券行业数据生态建设

王毛路 / 北京共识数信科技有限公司
俞枫 / 国泰君安证券股份有限公司



数据作为数字经济的核心，其基础性战略资源地位正日益凸显，通过数据开放共享来挖掘数据潜在价值已成为社会各行业的共识。证券行业数据生态建设仍处于探索阶段，“数据割据、数据孤岛”问题严重。为提升证券行业智能化服务能力，促进跨机构、跨界创新业务开展，亟需建立行业数据生态，实现行业范围内的数据协作共享。

然而由于行业数据生态建设需要共享最为核心的金融数据，这牵扯到各方利益，因此建立行业数据生态必将面临诸多困难和挑战，主要包括四个方面：一是数据共享理念尚未建立，二是缺乏平等互信和激励机制，三是缺乏敏感数据保护机制，四是数据主权难以有效保障。数据生态建设面临的困难，传统中心化系统难以全部解决。

本文来源于国泰君安证券股份有限公司和北京共识数信科技有限公司共同承接的中国证券业协会 2019 年重点课题研究成果。结合当前行业数据生态建设面临的痛点，以区块链技术为载体，结合安全多方计算、分布式身份标识 DID、可验证授权凭证等技术，构建一个可信的行业数据生态系统，用于促进数据安全流通，释放数据价值。

一、概述

1.1 数据生态建设的背景

目前我国已进入数字经济时代，数字经济已成为带动中国经济增长的核心动力，据《中国数字经济发展与就业白皮书（2019年）》数据，2018年我国数字经济规模达到31.3万亿元，占GDP比重为34.8%。推进数据资源开放共享符合技术发展趋势，大数据产业生态体系建设已经初见规模。虽然当前我国在大数据发展和应用方面已具备一定基础，数字生态正在加速形成，但是数据割据、数据孤岛现象依然严重，数据开放共享程度仍然不足。

数据生态（Data Ecology）是结合生态体系和大数据两个不同领域的特点所产生的新概念。如图1，在证券行业中，数据生态体系主体包括了监管机构、交易所、证券公司、登记结算公司、资管期货公司、基金公司等；涉及到的关键数据包括客户数据、用户数据、产品数据、交易数据、监管数据、市场数据、营销服务数据等；数据的生命周期又通常包括了数据的产生、采集、存储、流通、使用、销毁等；从数据不同阶段的使用角色上看，又涉及到了数据源方、数据所有人、数据需求方、数据服务方等。这些不同主体，不同角色围绕着数据这个核心资源相互制约、相互促

进，进而形成证券行业数据生态的有机整体。

1.2 数据生态建设面临的挑战

1.2.1 数据共享理念尚未建立

当前证券行业金融机构普遍未深刻领会数据开放共享带来的价值，都希望获取外部数据，但是不愿意共享自身数据资源。比如在场外衍生品业务中，金融机构需要知道交易对手在全市场的相关业务集中度以进行信用定价和风险控制，然而由于信息共享平台和信息共享理念的缺失，无法进行行业内的集中度信息共享，最终导致盲目授信，出现了同时爆仓的情况，造成了重大利益损失，究其根本，还是缺乏进行数据共享以互利互惠的理念。

1.2.2 缺乏平等互信和激励机制

一是，金融机构担心加入信息共享系统后自身数据被系统的建设中介或其他行业竞争对手获取，从而造成自身核心利益的受损。二是，中小型金融机构因自身行业地位及IT建设能力薄弱等原因，亦担心加入信息共享系统后处于不平等的弱势地位。三是，针对信息共享的价值，缺少市场化的激励机制，而该激励机制的设计、实现、执行、兑付难以由传统中心化系统来完成，一旦发生信息交易记录被恶意篡改的情况，将引发信用危机。

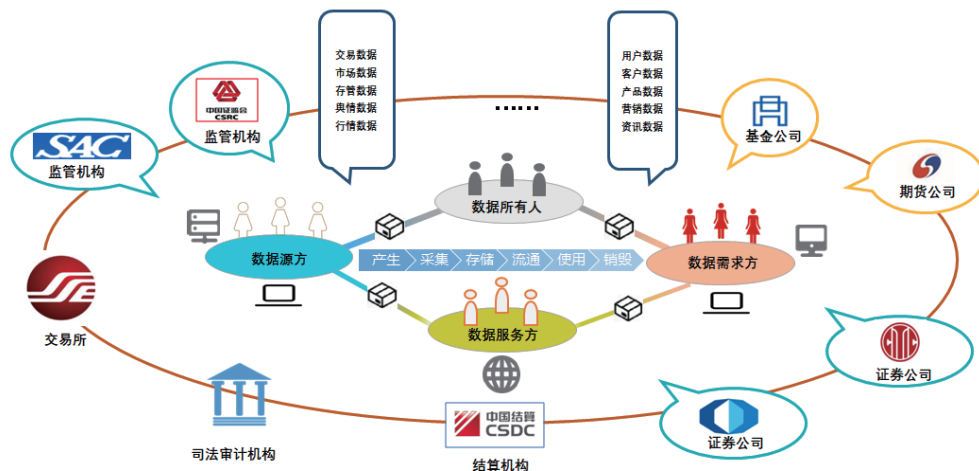


图1 证券行业数据生态体系

1.2.3 缺乏敏感数据保护机制

数据脱敏是目前应用最为广泛的传统隐私保护技术，它是按照一定规则对敏感数据进行变形或替换处理，以达到隐藏敏感数据的目的。但是在数据生态场景中，将脱敏后数据完全共享暴露仍然是不安全的，本文提出使用安全多方计算技术来解决多源数据计算场景下数据的隐私安全问题，后文将详细阐述。

1.2.4 客户的数据主权难以有效保障

对于个人信息保护和向第三方传输数据的合法条件，我国在多个法规条例文件中早已有所表述。无论是国外还是国内的相关法律，其内容具有一致性：经用户授权以及安全风险评估后，数据服务方可以向第三方传输相关数据。然而，现实场景中，用户的数据托管在机构后，机构就间接获取了用户数据的主权，使得在用户未授权的情况下，机构滥用客户数据的情况屡有发生，行业数据生态建设旨在汇集整合行业各机构海量数据，海量数据意味着平台未来潜在的巨大价值，这对系统的安全性提出了更高的要求。

二、数据生态系统相关技术简介

2.1 区块链技术

区块链本质上是一个分布式账本，以区块链结构和分布式存储保证数据不可篡改以及无法抵赖，使用密码学技术保证访问和通信安全，通过共识机制保证多方数据的一致性。现在已经从单纯的技术探索走向了应用落地的阶段，典型的应用涉及数字货币、跨境支付、登记结算、数据存证、供应链、物联网等多种场景。

区块链技术体系中的分布式存储、密码学原理、共识机制、P2P 网络等多重技术特性保证了区块链上的信息既公开透明又能保护隐私，通信信道即使暴露在外也能保证数据的不可篡改，参与方既可以共同参与决策又可以保护自身利益不受侵犯，极大地提高了价值交互的效率。

2.2 安全多方计算

数据生态与数据安全犹如一体之两翼、驱动之双轮，只有彻底解决数据隐私保护问题，数据孤岛问题才能根本解决，数据共享的创新业务场景才能真正实现，下面结合区块链及密码学相关技术，论述区块链相关隐私保护方案。

目前市场上针对区块链隐私保护问题，提出了多项技术，其中最主流的有：数据脱敏、零知识证明 + 同态加密、SGX 可信执行环境技术、安全多方计算技术。

表 1 隐私保护技术

技术路线	描述	主要问题
数据脱敏	对敏感信息通过进行变型、替换。	可以通过统计分析、数据挖掘等技术进行潜在的关联关系分析。
零知识证明	不了解某一秘密的前提下验证另一方确实拥有秘密	验证速度慢，可扩展性差
SGX 可信执行环境	通过计算机的核心处理器 CPU 来构建可信执行环境 TEE	严重依赖硬件厂商，存在侧信道攻击可能，可能受贸易战影响
安全多方计算	无可信第三方情况下，共同执行某一规则，执行过程隐藏各方原始数据	无法保证参与方的诚实性；效率较低

安全多方计算 (MPC : Secure Muti-Party Computation) 研究由图灵奖获得者、中国科学院院士姚期智教授在 1982 年提出，多个持有各自私有数据的参与方，共同执行一个计算逻辑（如：求最大值计算），并获得计算结果。但过程中，参与的每一方均不会泄露各自数据的计算，被称之为 MPC 计算。

安全多方计算所要确保的基本性质就是：从计算执行期间发送的消息中不能推断出各方持有的私有数据信息。

在本文中，安全多方计算相对于其他技术具有以下几点优势：一是数据不出域，更符合应用场景；二是适用于证券行业的半诚信场景；三是国内研究深入，目前已有一些通用解决方案。

下面以黑名单共享应用为例，说明安全多方计算技术在证券行业数据生态系统中的应用原理。

使用 Paillier 概率公钥加密算法，实现数据隐私保护，Paillier 满足加法和数乘同态，其实现

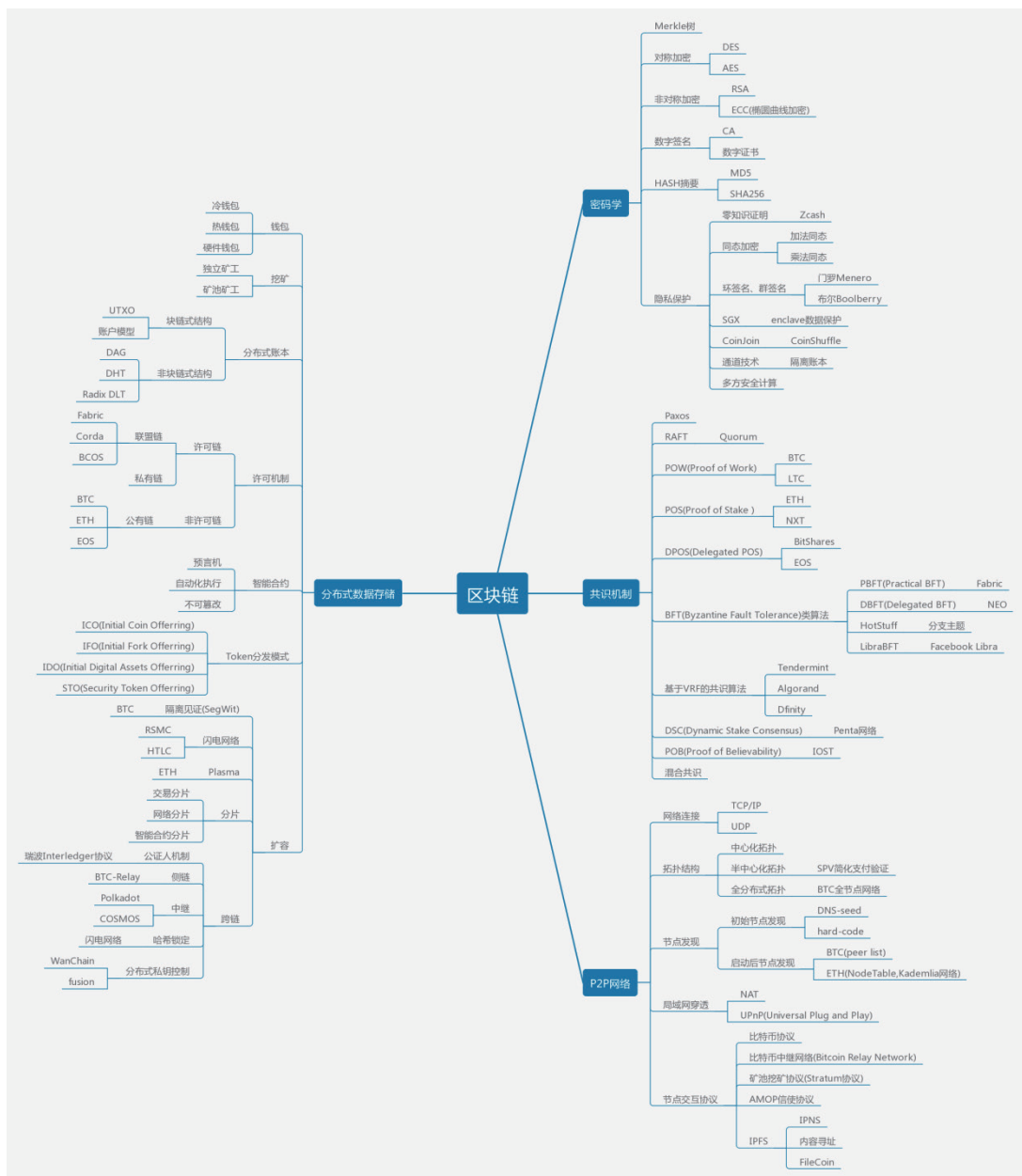


图 2 区块链技术思维导图

过程如下：

假设 A 是黑名单数据查询方，B 是数据拥有方，A 有一对 **paillier** 的公私钥，公钥记为 **pub**，私钥记为 **priv**。

A 查询的黑名单数据是 **a**，**b** 是 B 里面的某一条数据，数据 **a** 和 **b** 按下列步骤进行比对：

一、A 将 **pub** 和 **c = pub (a)** 发给 B。**paillier** 算法有随机因子，每次运算 **pub (a)** 结果都不一样，B 无法猜测查询数据。

二、B 设置两个随机数 **x**、**y**。计算 $C_{B1} = \text{pub}(ax+y)$; $C_{B2} = \text{pub}(bx+y)$ 。

三、B 将 C_{B1} 、 C_{B2} 发送给 A。因为 **paillier** 同态是满足加法和数乘同态，故 $C_{B1} = \text{pub}(ax+y) = \text{pub}(ax) + \text{pub}(y) = \text{pub}(a) * x + \text{pub}(y)$ 。

四、A 用自己的私钥 **priv** 对收到的 C_{B1} 、 C_{B2} 进行解密，得到的结果是 **ax+y** 和 **bx+y**，比较这两个数即知道数据 **a** 与数据 **b** 的是否相等。相等的话返回“=”，不相等的话返回“≠”。

五、若比较结果为“=”，则表示 B 的数据库中
存在数据 a，终止比对过程。若比较结果其
他，继续比对 B 的其他数据，直到结果为“=”，
或者所有数据都经过了比对。

若结果为“=”，则表示数据 a 和数据 b 相同，
即说明 A 所查询的数据在 B 的黑名单数据库中
同样存在，停止 a 和 B 其他数据的匹配工作。否
则，继续执行数据比对。

为了提升黑名单共享的数据比对效率，使用
此算法进行安全多方计算时可以使用并发、批量
数据对比的方式。例如，每次对 1000 条数据进
行批量比对，以此减少数据包发送的频率，降低
对网络的使用频率。查询时，可以同时等多家券
商发起数据查询申请，降低黑名单信息查询花费
的总时间。

paillier 分析：

1. 复杂度分析

Paillier 与其他密码算法的加解密复杂度对
比入下图所示：

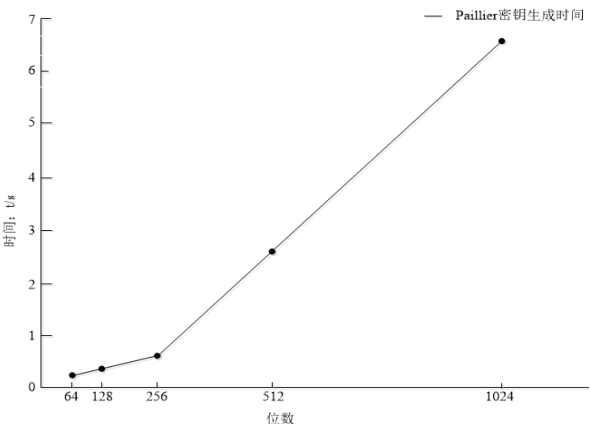
	RSA(n,e)	ELGamal	Paillier
	$ n , p =512, e=2^{16}+1$	$ n , p =512$	$ n , p =512$
加密	17	1536	5120
解密	192	768	768

说明：表中数值为进行加解密运算所需要模 $(|n|=512)$ 乘运算的次数

虽然 paillier 的加解密复杂度较高，但是密
文的同态操作减少了大量的加解密次数，所以次
数不多的加解密耗时是可以接受的，且密文的同
态操作避免了读取用户明文形式的密钥份额。当
有多个节点离开或者加入时，基于同态技术的批
处理密钥分发方案以其较少的加解密开销将更具
有优势。

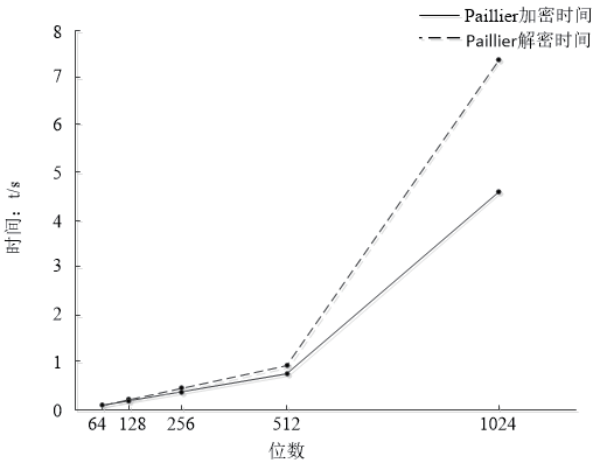
2.Paillier 公钥密码体制密钥生成效率分析

Paillier 公钥密码体制公钥为 (n,g) ， $n=pq$ ，
私钥为。仿真分析中，分别生成了 n 为 64 位，
128 位，256 位，512 位和 1024 位密钥，相应的
产生 g 和 λ ，从而生成整个加解密过程中所需
要的公钥和私钥。由于生成的密钥 n 的长度逐渐
增长，所以生成过程中所需要的运算量也逐步增
加，而且增加的速度也逐渐上升，如图所示：



3.Paillier 公钥密码体制算法加密解密性能分

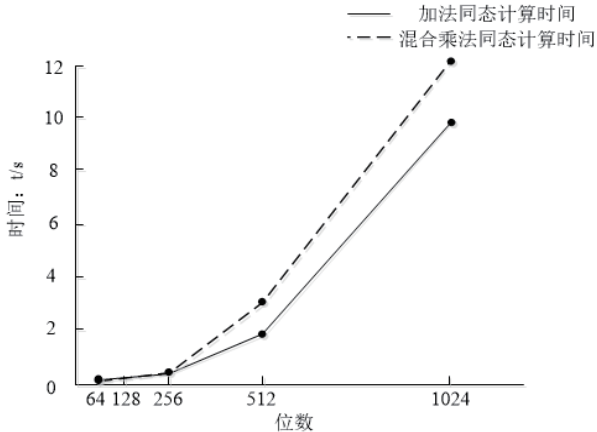
析



上图给出了在密钥 n 为 64 位, 128 位, 256 位，
512 位接 1024 位的情况下加密 32bits 数据所消耗
的时间，随着密钥 n 长度的逐渐增加，加密和解
密过程中所消耗的时间也逐渐增大，加密和解密
过程中的运算量逐渐增大。

4.Paillier 公钥密码体制加法同态特性效率分

析



如图所示,在 n 分别为 64 位, 128 位, 256 位和 1024 位时,已知 $E(x)$ 和 $E(y)$ 的情况下,通过 paillier 公钥密码体制的加法同态特性可以求出 $x+y \bmod n$, 另外当已知 x 和 $E(y)$ 时,也可以求出 $E(xy \bmod n)$, 其中 x 和 y 均为 32bits 数据,消耗时间如上图所示。

5. 黑名单场景分析:

在黑名单场景中,数据源方和数据查询方会交互多次,数据查询方会一直进行比对,直到结果为“=”,或者所有数据都经过了比对。而每一次的比对都会利用 paillier 算法,故最后黑名单场景整体的复杂度是单次 paillier 比对的 n 倍,即 $n \cdot O(n)$, 这里的 n 代表的是数据源方数据的数量,或者是出现“=”时的已经匹配过的个数,其性能与 n 的大小成负相关,即 n 越大,性能越低。

2.3 分布式身份标识 DID

在金融、政务等重点领域,身份认证环节有着举足轻重的作用,由它完成用户物理身份和数字身份的转换。随着人脸识别、区块链分布式技术的成熟,分布式身份认证系统得到了发展,本文所使用的实体身份认证方案是基于区块链技术建立底层平台,使用人脸识别技术,接入权威的身份认证节点(例如公安部一所的人脸识别区块链节点),可以自动、快速、合法的完成实体身份的认证,区块链技术可以对身份认证进行全流程记录,以此达到身份认证的合法合规性,方便后续开展业务。而随着用户对身份控制和自我保护意识的增强,数字身份先后产生了四种形态:分别是中心化身份、联盟身份、以用户为中心的身份和去中心化身份(DID)。由于前3个阶段的数字身份或多或少在安全性上存在隐患,由此DID就应运而生了。

去中心化身份(DID)利用区块链技术实现让数字身份真正为用户所拥有并支配,就像我们把身份证、护照、户口本这些纸质文件放在自己家里小心保存,只有在需要的时候再拿出来一

样,不再有任何中间人(即使是DID技术供应商)接触拥有控制用户的身份和数据。

为了保证DID的唯一性,DID是按照身份证信息进行计算的,通过加密算法避免反向破解导致用户信息泄露的风险,为增加证券行业数据生态系统的公信力,DID由中立的第三方机构负责执行具体运算过程。基于DID可实现分布式用户信息标识共享、可信数据交换、安全多方计算等一系列应用,还可实现用户对数据共享进行授权管理。

DID具有如下特点:

一、具有唯一性,它是由不可更改的用户身份信息按照一定的规则生成,在数据共享系统内是唯一的。

二、可以避免反向破解,首先计算规则只有第三方中立节点知道,只能由第三方中立节点根据身份信息计算出DID,但不可以由DID反向推出用户身份证号码。

三、具有公开性,DID在链上是公开的,用作信息上链、信息存储、信息查询。

四、各券商基于用户DID进行信息的共享和维护,用户可以对个人数据权限进行控制。

2.4 可验证授权凭证

本文使用的可验证授权凭证属于自定义凭证,是用户管理个人信息的工具,用户通过授权文件的签署、验证达到管控个人信息的目的。它是利用区块链不可篡改的特性,将用户授权的过程记录的数据Hash上链,再加上授权的目的、数据申请方的签名,确保授权的合法性、真实性、可验证性。

本文参考了微众银行Welidentity关于可验证凭证的思路。可验证授权凭证是基于W3C规范生成的,可以把用户DID、授权时间、授权方式、授权对象、数据源方、数据使用方、数据类型、目标用途等与授权有关的数据标准化和电子化,通过将原始数据的Hash上链,附上数据使

用方的权威签名之后,即可生成可验证、可交换、不可伪造的授权凭证(Credential)。在各券商之间,授权凭证可以互通互认,从而实现用户对个人信息全平台的授权管理。

WeIdentity 是一套微众银行自主研发并完全开源的实体身份标识与可信数据解决方案,可承载实体对象(人或者物)的现实身份与链上身份的可信映射、以及实现实体对象之间安全的访问授权与数据交换。

2.5 区块链、安全多方计算、DID 的作用与配合

基于区块链安全可信的特性搭建行业数据可信共享平台,以实现数据的可信流通及安全管理,该平台可提供数据可信记录。在 DID 场景下区块链记录数据包含实体身份认证过程、DID 生成过程、用户数据、授权过程,在安全多方计算场景下区块链记录数据包含数据调用过程,数据计算过程等。区块链提供基础算法服务,如加密算法(RSA、ECDSA、SM2、SM3、SM4 等)、计算函数(求和、求余等)等。

DID 和区块链配合:

基于区块链实现分布式身份识别系统,使用人脸识别技术,接入权威的身份认证节点(例如公安部一所的人脸识别区块链节点),可以自动、快速、合法的完成实体身份的认证,区块链技术可以对身份认证进行全流程记录,以此达到身份认证的合法合规性,方便后续开展业务。

用户通过权威机构完成实体身份认证,区块链智能合约自动生成用户唯一数字身份标识(DID),该 DID 保存与区块链分布式账本中。通过 DID,用户可以向访问的系统提供有限的信息,该过程称为‘用户授权’,用户授权过程基于区块链智能合约实现,授权过程在区块链账本中进行可信记录。

本系统结合了实体身份认证,DID 身份唯一标识,实现了用户数据的分布式标识体系,系统

可以兼容传统数据库,实现数据的检出与上链。各券商可以通过区块链数据生态系统进行数据查询,数据访问申请,用户可以通过刷脸等方式基于 DID 对个人信息进行可验证授权凭证的管理。在用户信息共享的基础上,数据生态系统还支持券商进行安全多方计算,在保护用户隐私的前提下,通过智能合约的协作,在各券商之间进行数据共享,提升数据价值。

安全多方计算和区块链配合:

安全多方计算的实现依赖于特定的算法及执行流程。随着应用场景的不断丰富,安全多方计算算法的不断研发,可根据实际的数据交互需求,开发并部署不同的区块链智能合约,基于区块链智能合约实现更多的隐私数据共享方式。

隐私数据在本地进行处理,处理后的数据发送的区块链系统,调用区块链智能合约,智能合约完成数据的验证和交互,数据需求方获得数据提供方提供的处理后的隐私数据,在本地完成后续计算过程。通过该种方式形成一个去中心化的,可信任的数据隐私交易市场。该市场中所有数据交互(交易)过程,均会在区块链分布式账本中进行完整、可信的记录。

三、基于区块链的数据生态系统

3.1 系统架构设计

区块链数据生态系统是在原有业务系统的基础上工作的,可以和传统系统无缝对接。区块链网络由区块链节点组成,包括 Peer 节点、Orderer 节点和 Kafka 节点。区块链网络以块链式的分布式账本实现对数据指纹和数据交互过程的记录,并运行智能合约以规范上链数据的格式和数据交互的流程,并实现权限控制。

系统架构图如图 3 所示。

3.2 系统业务流程

数据生态系统的主要流程包括数据上链、数

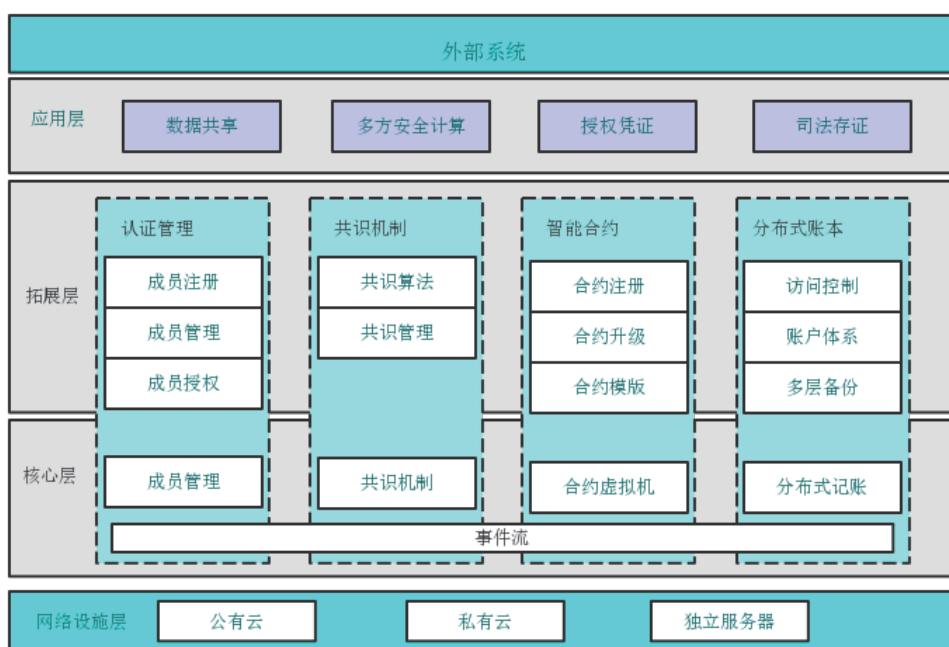


图3 基于区块链的数据生态系统架构图

据共享以及安全多方计算，结合了实体身份认证、DID，实现了分布式用户数据共享。生态系统的参与券商可以通过此系统进行数据查询，数据申请，用户可以基于 DID 对个人信息的共享进行授权。在数据共享的基础上，还可引入相关算法，通过安全多方计算，提升证券行业整体数据的价值。比如黑名单信息共享可以帮助各个券商充分识别风险，降低业务损失的风险。

数据上链。数据共享与协同过程中，需要在信息泄露最小的情况下，保证数据的可信度。通过数据指纹上链的方式，可实现对数据内容的有效锚定。数据指纹通常是计算原始数据的 hash 值，其特性是同一条数据内容计算得出的数据指纹是唯一的。

信息查询。通过区块链节点提供的查询服务对 DID 进行查询，数据使用方可以查看通过区块链检索到的数据，并从中选择需要发起数据共享请求的部分。

用户授权。数据使用方应向用户发起信息访问授权申请，用户通过扫脸识别、指纹等手段验证身份并进行授权，全程通过区块链进行记录，

并生成可验证的授权凭证，加上数据使用方的机构签名之后，向数据持有方发起数据请求。

3.3 节点准入控制

在数据生态系统中，确认各参与方的身份及权限是十分重要的。首先，对数据发布方的身份必须进行有效的验证，才能保证数据的可信度和有效性，否则存在数据质量不合格的风险，后续数据交互也无法有效开展。而如果对数据使用方的身份不进行严格的验证，则无法控制数据的使用范围，有可能导致数据泄露，造成更大的损失。

因此，本系统使用区块链技术作为数据共享平台底层的账本机制，只有获得指定 CA 证书的参与方，才可以加入网络，并按照证书许可的权限范围，参与数据共享与数据协作的过程。

3.4 司法存证功能

尽管区块链上不可篡改的记录可以理清数据共享与协同过程中的各方权责，但随着数据的重要性日益突显，对数据真实性引起的纠纷也是不可避免的。为了能够有效地处理这类纠纷，在建

设数据共享与协同系统时，预先考虑了与司法体系对接进行数据存证。

目前，经过认证的区块链数据可以作为包含时间的、确保内容完整性的证据。在具有司法效力的机构如公证处、司法鉴定中心等机构间搭建一条具有公信力的证据链，并将业务链（如数据交互区块链）上发生的所有交易的数据指纹锚定在证据链上。在发生争议时，业务链上的交易记录可以对数据内容、发生时间等进行证明，而证据区块链上的数据指纹可以为业务区块链上的数据提供完整性背书，并根据参与节点在司法体系中的职能，提供公证书、鉴定报告等易于被司法程序所采信的证据。

3.5 安全多方计算功能

本系统将基于区块链打造安全多方计算网关，允许各个参与方在不暴露原始数据的情况下，获取基于联合数据的计算结果。各个参与方在安全多方计算中的关系如图 4 所示。

安全多方计算的实现依赖于特定的算法及执行流程。随着应用场景的不断丰富，安全多方计算算法的不断研发，可根据实际的数据交互需求，

开发并部署新的智能合约，实现更多的数据共享应用方式。

四、总结

本文创新性的提出了使用区块链技术来搭建行业数据生态，并同时使用了安全多方计算、分布式身份标识 DID、可验证授权凭证等技术来解决传统数据生态建设面临的数据共享、数据协作、隐私保护等难点问题，具有重要研究和实践价值。

使用区块链技术来实现不同机构、不同角色在无需中间代理机构的情况下大规模、安全的进行分布式数据共享。这种开放式、无中介、分布式的数据生态体系对行业的发展具有重要的意义：

对个人来说，数据主权有史以来第一次真正掌握在自己手中。通过结合分布式身份标识 DID 和可验证授权凭证技术，只有经过用户认证并授权的个人数据才可以被第三方共享，同时数据源可以对用户授权进行验证，用户成为个人数据的真正控制者和受益者，不但将数据主权归还给用户，还有可能蕴育出全新的分布式商业模式。

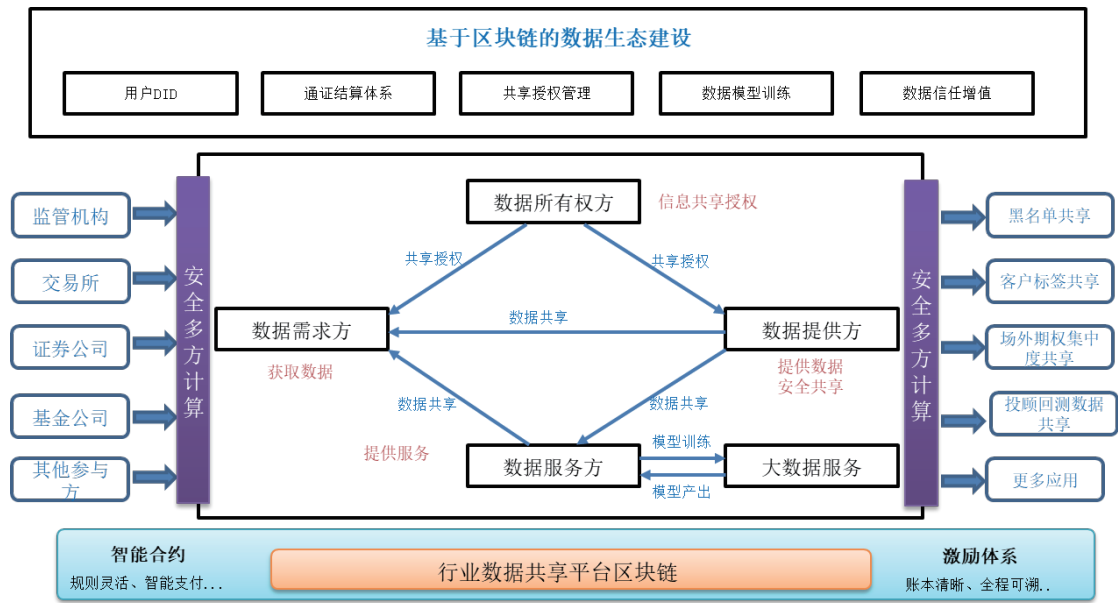


图 4 安全多方计算架构图

对金融企业来说，数据共享有利于跨机构、创新性业务开展，从而为客户提供更加多元化产品和服务。区块链是一种分布式对等网络，每一个参与方都可以拥有一个对等节点并有一份完整的数据拷贝，各方之间的数据通过建立 P2P 加密数据信道传输，使得各参与方能够在一个平等互信的网络环境中进行数据价值流通，释放数据红利。一方面大型、小型机构都是基于机构内部已有客户数据来完善自身客户的信息完善度，各方平等参与。另一方面，按机构贡献的程度设计合理的激励机制，促进大家对有价值的数据进行共享，既提升了数据价值，又促进了跨机构合作等创新业务开展。

对监管层来说，数据共享有利于减少信息不对称带来的市场风险，降低了因为封闭带来的金融垄断，有利于行业整体高质量发展。主动使用

新技术，抢占战略制高，坚持金融开放，推动行业数据生态建设和金融开放，将成为助力证券业高质量发展的重要战略引擎。

最后需要指出，随着数字经济和大数据的不断深入发展，数据共享不但是证券行业的“刚需”，同时也是银行、保险等其他金融行业甚至是全行业的广泛需求，本研究也因此具有广泛和普适的应用价值。

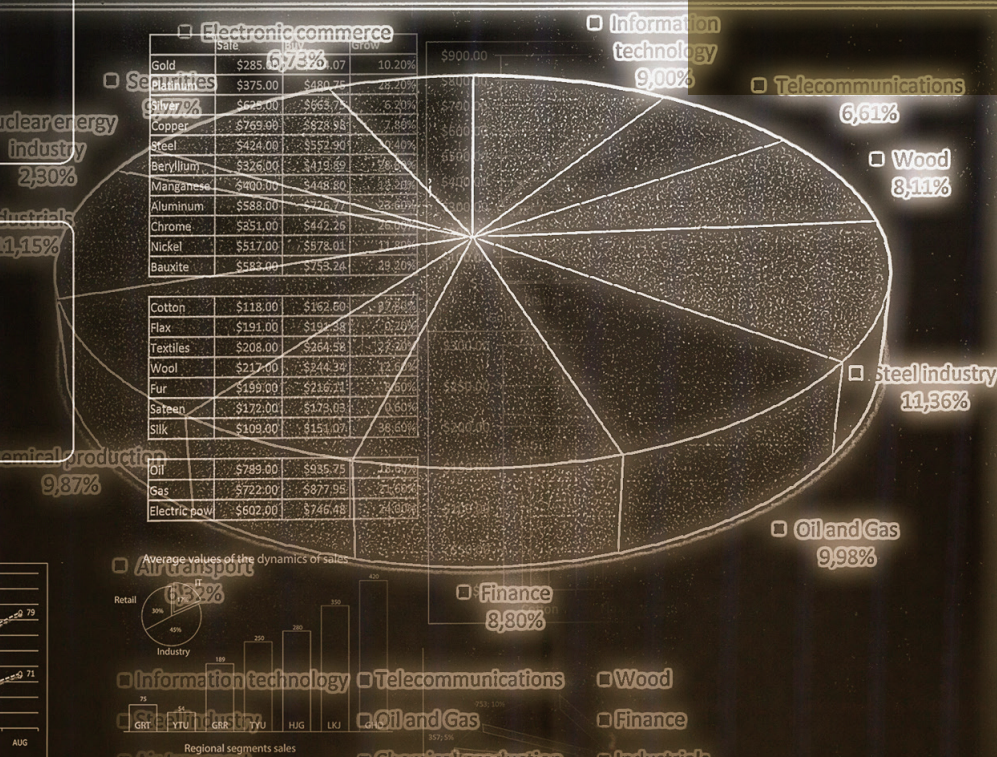
为充分发掘区块链在行业数据生态建设中的价值，让各方都从繁荣的行业数据生态中真正受益，本文提出了以下几点建议：

一是制定政策，明确数据生态建设及区块链应用监管底线。二是支持搭建行业链，打造区块链基础设施。三是制定行业数据治理标准，提高数据质量。四是加强人才培养，打造行业数据生态标杆应用。



E 实践探索 xplore

- 5 运维智能服务台的探索和实践
- 6 基于业务调度框架的证券业信息系统应用研究
- 7 运用深度学习技术处理投行项目富文档的探索与实践
- 8 证券公司信息系统用户权限集中管理的研究与实践



运维智能服务台的探索和实践

杨超、吴浩、耿锋、李沛臻 / 光大证券股份有限公司 wuhao@ebscn.com



运维服务台提供了统一对外的服务窗口，但存在沟通成本高、人员流动性大，服务请求的整体交付周期普遍较长，难以保证服务质量的问题。同时，知识库不完善等问题也限制了服务台作为一线服务的工作能力，往往需要通过二、三线人员的专业知识解答问题，构建完整的运维服务闭环。因此，为了缩短运维响应时间、提升服务质量，同时降低沟通成本、缩短服务路径，光大证券信息技术总部建设了运维智能服务台，重点聚焦运维服务知识库的构建及搭建智能服务台。首先创新地使用“知识众包”形式构建运维知识库，然后将 fastText、注意力机制等人工智能技术应用在识别用户意图，快速定位问题并提供解决方案中。通过智能服务台的成功应用，多种渠道对外提供月访问服务次数约 700 次，有效问题一次解决率约 58.6%，其中在服务营业部的云桌面领域问答一次解决率高达 83%，大大提高了运维服务的质量和效率。

前言

随着国内金融市场的快速发展，技术的迭代升级掀起了新一轮的金融科技创新浪潮，生产系统的双态运维压力也在不断增加，在传统运维思路中，系统管理员集中于一线大量繁琐重复运维工作的工作模式已不符合当前的发展要求。为提升 IT 服务能力及效率，光大证券信息技术总部提出了“服务标准化、流程自动化、管理数字化、决策智能化”的运维要求，结合 ISO20000 国际标准建设的 IT 服务管理体系，建设了光大证券信息技术总部运维服务台。服务台提供了统一对外的服务窗口，但存在沟通成本高、人员流动性大，服务请求的整体交付周期普遍较长，难以保证服务质量。同时，知识库不完善等问题也限制了服务台作为一线服务的工作能力，往往需要通过二、三线人员的专业知识解答问题，构建完整的运维服务闭环。因此，IT 运维服务的服务台更应关注缩短运维响应时间和提升服务质量，降低沟通成本和缩短服务路径，提高用户满意度。

基于大数据、云计算和深度学习等领先的人工智能技术，不断的向产业互联网应用赋能，进一步提升行业的各项能力，服务台领域也在发生着快速的变化，大多传统型服务台向智能化转型。人工智能在赋能服务台领域的过程中，已经可以实现自主问答、故障引导等一系列基础任务，满足大部分的应答需求，快速高效的引导用户解决问题。结合市场调研和运维的实际工作，参照对比多家智能客服解决方案，IT 运维服务台通常需要负责多套系统的运维工作，存在着系统覆盖面广、知识分散、解决问题实时性要求高等特点，目前市场上很多智能客服领域的产品不能满足运维智能服务台的需求。

本文立足于人工智能等技术，对运维领域的智能服务台进行了积极的探索，通过构建扁平化的数据处理模式和建立环状的系统管理员交互机制，力求可以构建一个证券业标准化、可推广的

智能服务台体系。

1、设计思路

基于服务台运行现状和需求分析，光大证券调研多种前沿解决方案、技术及算法，从运维服务的维度出发做了一系列的探索，主要包括运维服务知识库的构建及搭建智能服务台。

1.1 构建运维知识库

结合目前运维工作实际，运维知识分散性较大，具有多而广的特点。为有效建立运维服务的智能知识库，充分调动运维服务梯队的力量，创新性提出“知识众包”的形式，即由运维一线根据工作实际，充当知识贡献者角色提交知识，由运维二线的系统管理员充当专家角色审核知识。

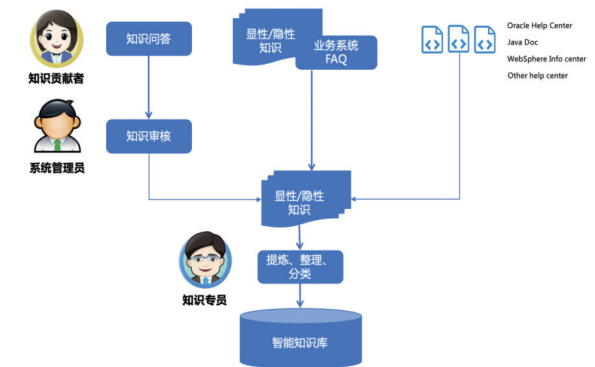


图 1 运维知识库数据流

如图 1 所示，构建运维知识库，主要从数据收集、数据清洗、数据标签等三个方面重点开发：

数据收集。主要从以下三个渠道维度进行收集显性和隐形知识：

- 充分利用运维系统现有的 QA 知识。收集现有各运维系统已有的 QA 知识，进行数据梳理，导入知识库中作为智能学习模型训练基础；
- 第三方接入形式进行数据采集。通过对接现有 ITSM 平台等知识库，采集 CMDB 数据，实现知识库的动态更新；
- 通过爬虫形式扩充数据采集点，获取 Oracle、操作系统等运维专业文档的数据采集。

数据清洗。构建多种数据清洗因子实现对现有知识库的切分、格式化、去除和修正，达到数据清洗的目的；

数据标签。根据现有知识库建立自动化学习模型，通过传统服务台模拟真实场景进行模型训练，构建数据标签，区分意图理解错误、答案质量问题、未登录知识点、持续运营、训练提高知识库命中率。

1.2 搭建智能服务台

完善好运维知识库后，智能服务台基于机器学习等算法对问题进行分析 and 聚类、要素提取、模糊识别、算法匹配等数据处理，帮助用户快速定位问题并提供解决方案。

智能知识库的对外输出渠道具备多样化，整个交互过程为：终端用户通过 QQ、微信、网页等多渠道形式提问，智能服务台通过智能语义识别后，根据机器学习等算法进行置信度匹配知识库答案实时返回给用户。若未匹配到合适答案，智能服务台则对外吐出通用性回复，这时则需传统服务台介入干预。其目的一方面是充当智能服务台的有效补充，提供运维闭环服务；另一方面，

进行智能知识库的完善与补充。

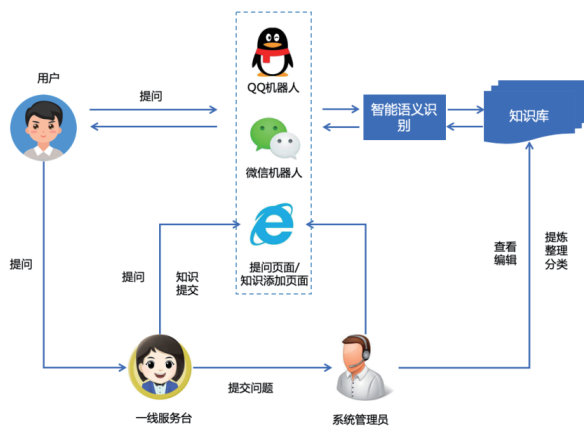


图2 智能服务台的服务渠道框架图

智能知识库的完善程度和语义识别准确性是智能服务台支持力度的关键因素。综合考虑到运维解决问题的紧迫性，人机辅助相结合的形式是智能服务台推广初期的有效解决方案，有利于高效训练智能学习模型，提升运维服务的解决效率。

2、核心算法技术

智能服务台的核心服务模块是问答引擎，该

表1：三种分词方法的对比表

分词方法	说明	优缺点
基于字符串匹配的分词方法	又称机械分词方法，它是按照一定的策略将待分析的汉字串与一个“充分大的”机器词典中的词条进行匹配，若在词典中找到某个字符串，则匹配成功。	优点：速度快 缺点：未登录词的效果差
基于理解的分词方法	计算机模拟人对句子的理解，达到识别词的效果。其基本思想就是在分词的同时进行句法、语义分析，利用句法信息和语义信息来处理歧义现象。	由于汉语语言知识的笼统、复杂性，难以将各种语言信息组织成机器可直接读取的形式
基于统计的分词方法	在给定大量已经分词的文本的前提下，利用统计机器学习模型学习词语切分的规律（称为训练），从而实现对未知文本的切分。通常使用分词词典来进行字符串匹配分词，同时使用统计方法识别一些新词，即将字符串频率统计和字符串匹配结合起来	既发挥匹配分词切分速度快、效率高的特点，又利用了无词典分词结合上下文识别生词、自动消除歧义的优点。

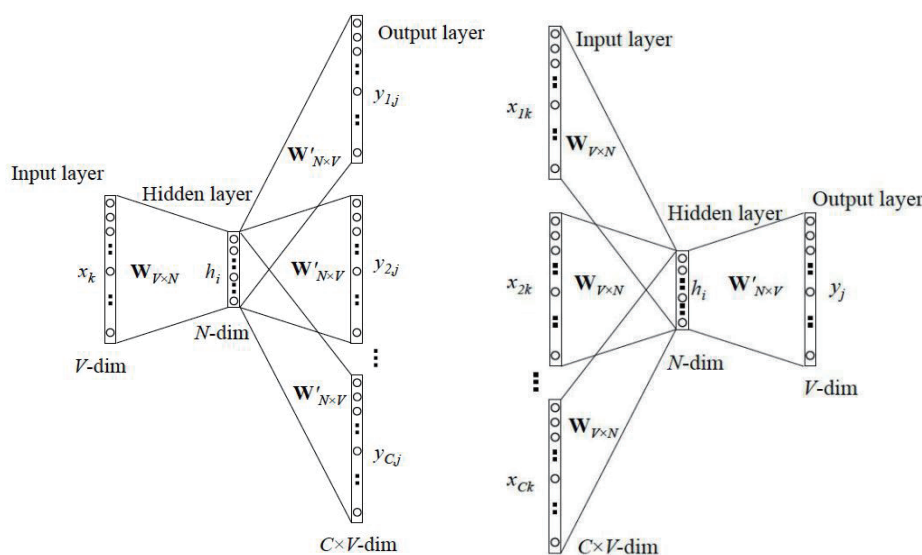


图 3 Word2vec 算法模型图

引擎用于识别用户意图。其中的关键技术包括：分词、词语相似度（词嵌入模型）、分类预测、相似问句计算。

2.1 基于统计算法的领域分词技术

词语组成句子，句子再组成段落、篇章、文档，因此分词和词性标注是中文自然语言处理的基础。分词技术已非常成熟，目前主流的可分为三大类：基于字符串匹配的分词方法、基于理解的分词方法和基于统计的分词方法。

运维领域知识特性及业务术语具有较强的专业性，所以分词这一自然语言处理的基础服务必须要具备较强的领域敏感性。特别考虑到运维领域的专有名词多等现象，为达到更好的自然语言研究效果，需要更精准的领域分词结果。经过综合考虑，本课题采用基于统计的中文分词方法，可以加载默认分词词典和自定义分词词典，发挥匹配分词切分速度快、效率高的特点，又利用了无词典分词结合上下文识别生词、自动消除歧义的优点。

2.2 基于 Word2vec 的词语相似度

基于运维系统说明书和知识库中的知识点等大量语料，我们采用 Word2vec 这种词嵌入工具

训练出证券运维领域的词向量。通过词向量计算词语相似度，也是数据处理中句子相似度的基础。

Word2vec 分两种模型：CBOW (Continuous Bag of Words) 和 Skip-gram。CBOW 是拿一个词语的上下文作为输入，来预测这个词语本身；Skip-gram 刚好相反，用一个词语作为输入，来预测它周围的上下文。

这两种方法都利用人工神经网络作为它们的分类算法。起初每个单词都是一个随机 N 维向量。经过训练之后，该算法利用 CBOW 或者 Skip-gram 的方法获得了每个单词的最优向量。训练过程如图 3 所示。

本课题的研究目标是证券行业的 NLP，选用的训练语料为运维系统说明书、知识点和问答记录，基于上述的分词算法对语料进行分词后，通过 Gensim 工具包训练词向量 Word2vec。

2.3 基于 fastText 的分类算法

在识别意图时，我们需要对产生的问题进行预分类，确认该问题所属系统，缩小运维问答领域，提高运维准确度。

词嵌入算法会为每个词语生成一个向量，这可能忽略了词语内部的形态特征，比如：“光大证券”和“光大”，两个单词有较多公共字符，

即它们的内部形态类似,但是在传统的词嵌入中,这种单词内部形态信息因为它们被转换成不同的id丢失了。而 fastText 使用了字符级别的 n-grams 来表示一个词,这样因为 n-gram 可以和其它词共享,对于低频词生成的词向量效果会更好。同时, fastText 在输出时采用了分层 Softmax,大大降低了模型训练时间。

本课题在研究过程中,也经过多次试验,尝试了 TF-IDF、fastText、TextCNN、Transformer、Bert 等多种算法,在业务日志历史数据测试集下的分类效果如下:

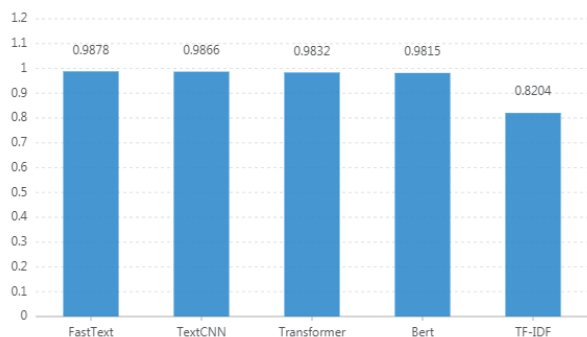


图 4 各分类算法的训练结果

如图所示: TF-IDF 算法表现不佳; TextCNN、Transformer、Bert 均为深度学习算法,语料少的情况下容易过拟合,训练时间相对较长,并且对硬件要求较高; fastText 在短文本处理上效果好,收敛快,预测快。

综合考虑到样本集的数量要求、硬件要求、和训练时长,对运维系统简单的分类均采用 fastText 算法。

2.4 基于注意力机制的句子相似度计算

随着深度学习技术的发展,基于神经网络的算法可以从大量的语料库中自动学习文本特征,训练得到文本分类模型,实现中文文本自动分类。同时,注意力机制的引入,已成为一种有效的策略用于动态学习不同特征对特定任务的贡献程度,在自然语言处理中也获得优异的效果。本质上句子相似度计算也是一个分类问题,即对于

问题 Q , 如何在已知的 $Q_1 \sim Q_n$ 中, 寻找最相似的 Q_k , $1 \leq k \leq n$ 。

在通用领域,我们可以获得大量数据,但是行业领域尤其是运维领域,往往没有足够多的语料,因此为了更好的学习到问句特征,获得较快的收敛,本课题采用了基于领域注意力机制语义嵌入模型,将领域术语作为特征输入。其模型架构图如下:

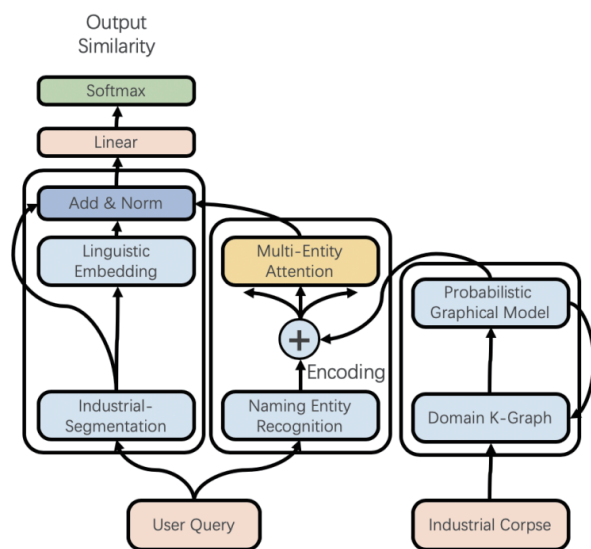


图 5 领域注意力机制语义嵌入模型

如上图所示,首先将领域知识体系构建为领域特征知识图谱,建立领域概念的属性及关系拓扑结构,并基于领域语料知识进一步训练知识之间关系的概率图模型,作为知识相关性因子。对用户的问题首先进行业务领域分类和业务实体识别,并将业务分类和实体识别的结果作为领域注意力特征加入到用户意图理解的特征表达,并在后续训练过程中持续迭代,经过实测该模型能够较快的达到收敛,有效抓住用户问题的核心业务诉求。

该算法不但可以有效提升意图识别的准确率,而且提高了中台运营的能力,业务人员可以通过补充领域术语来快速实现领域意图理解效果的调优。

3、功能展示

运维智能服务台采用自然语言处理和机器学习

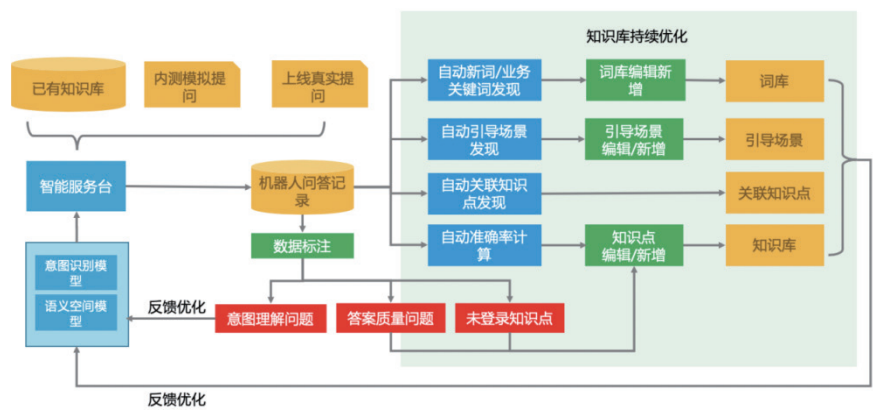


图 6 智能服务台的数据运营展示图

表 2 智能服务台的问答规则及动作表

匹配置信度	触发动作
95%以上	自动给出答案
80-95%	给出近似答案，同时给出关联问题
60%-80%	给出关联问题
60%以下	不给出答案，可给出关联系统负责人或者引导传统服务台；

习技术，将公司全面运维的生产系统的知识点纳入知识库体系，构建了扁平化的数据处理模式，完善运维服务维度的知识库标准体系结构，如图 6 所示。

通过用户不断的提问和使用，基于服务台内置的新增知识点、自动新词 / 业务关键词发现、自动引导场景发现、自动关联知识点发现、自动计算准确率等功能，不断的反馈优化，建立环状的系统管理员交互机制，实现了智能问答、问题

引导等功能。

3.1 智能问答

智能服务台已经在内部试运行，通过 QQ 机器人、微信机器人、网页端等渠道对外提供服务，其中 QQ 机器人服务 600 多个用户，每个月问答约 570 多次。问答范围覆盖网上交易、云桌面、三方存管、现金宝、协同 OA、数据服务等多个



图 7 IM 端智能问答展示

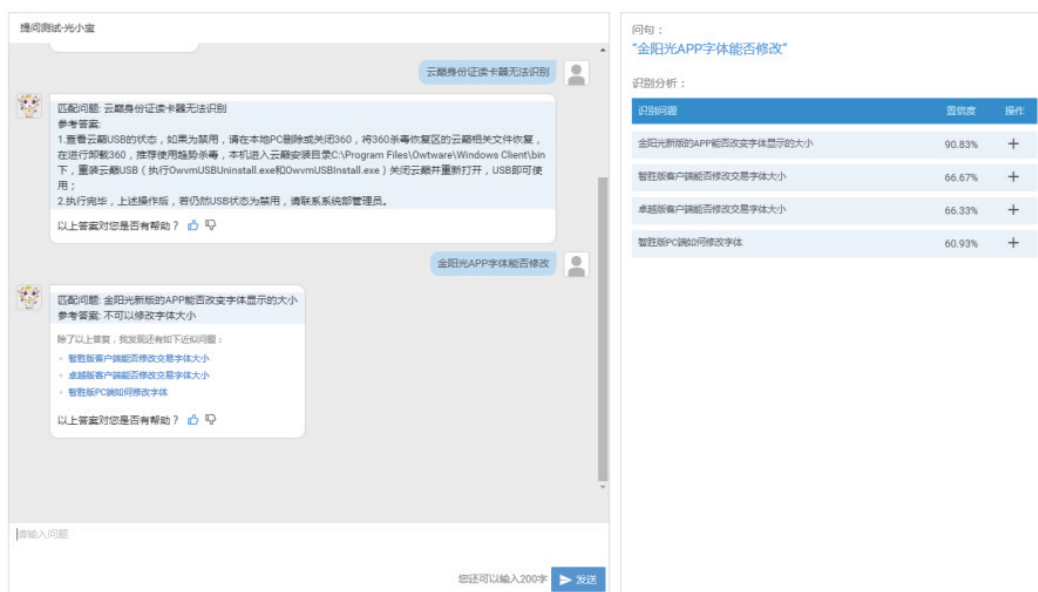


图 8 网页端智能问答展示

领域。智能问答规则如表 2。

以 QQ 机器人为例，实现的功能如图 7 所示，当用户 @ 光小宝的时候，我们会根据上述的智能问答规则进行自动匹配，反 @ 用户触发动作回答知识库匹配的答案。

网页端主要是针对运营人员和系统管理员等用户开发，主要是通过提问测试等功能实现问题

的场景测试、辨别新增知识点、增加扩展问题等，如图 8 所示。

3.2 问题引导

用户在提问时，经常会出现无法完整表达意图，或未抓住根本原因的情况，即意图不清。对于智能服务台来说，无法从语义上完全理解这类

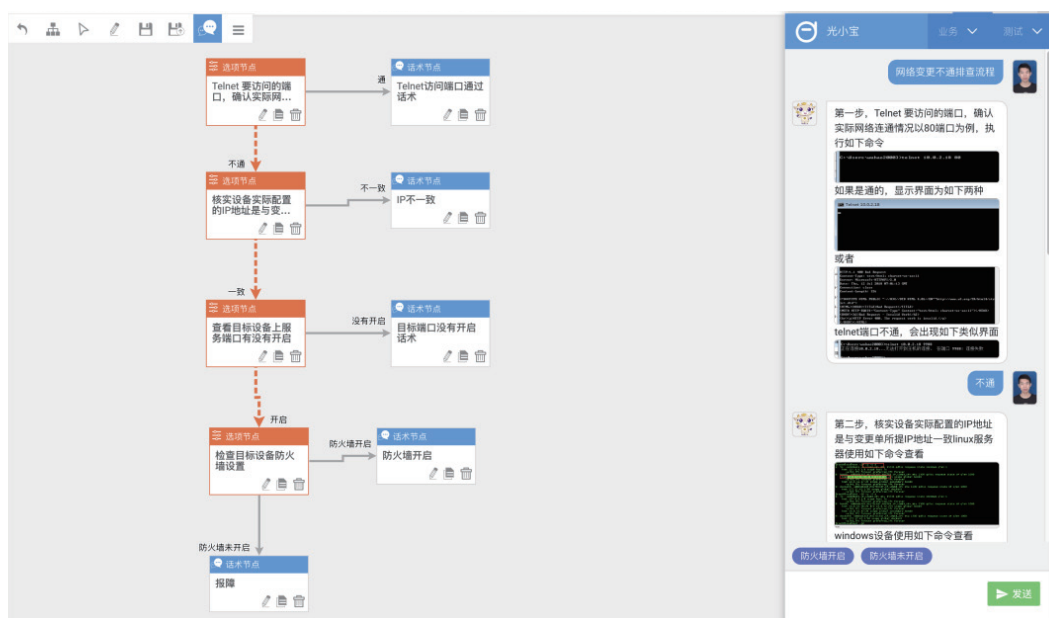


图 9 问题交互引导流程图展示

问题，因此考虑了模拟人工交互的过程，通过多轮交互式的引导，帮助用户定位问题，并引导用户解决。如图 9 所示。

4、总结

本课题是在智能服务台的深度探索，为面向运维服务的传统服务台起到了很好的辅助作用。

通过智能服务台的成功应用，通过多种渠道对外提供月访问服务次数约 700 次，有效问题一次解决率约 58.6%，其中在服务营业部的云桌面领域问答一次解决率高达 83%，大大提高了运维服务的质量和效率。通过智能服务台的持续建设，将帮助 IT 运维优化资源分配，打造一二三线专业服务梯队。本课题为行业内此类业务提供一定的借鉴和参考。

基于业务调度框架的证券业信息系统应用研究

周靖 任荣 王东 王洪涛 / 海通证券股份有限公司信息技术管理部

随着信息技术的快速发展，企业信息系统不断迭代更新，业务的办理形式也愈加多样化。外部环境竞争越来越激烈，促使业务部门产生快速迭代新业务、高频调整现有业务逻辑的诉求。但是，证券行业的业务往往逻辑异常复杂、业务系统耦合度非常高。信息技术管理部门在这样笨重且复杂的业务系统上进行持续迭代开发，势必会面临回归测试及业务监控困难等问题。而且上线迭代经常会引入新的生产问题。这些问题都说明，传统的应用系统开发模式已经满足不了精益模式的软件开发需求。

为了应对业务的快速变化，同时保障系统的高速稳定运行，本文提出一种基于业务调度框架的证券业信息系统设计方法。业务调度框架的核心原理是 Pipeline 管道模式，包含管道驱动机及状态机两个部分。业务调度框架可以统一编排调度各项业务流程，且具有完备的调度监控、业务执行监控能力。根据证券行业集中运营的特点，证券信息系统在设计的时候可以引入业务调度框架。依靠业务调度框架，可以将网上开户、一柜通业务办理等中的各业务生效步骤抽象成业务原子。由业务调度框架负责业务原子的编排生效，并保证业务原子执行一次且唯一一次。通过调度框架将复杂、重复的编排任务封装起来，统一调度、统一监控。开发人员可以集中精力在业务原子的开发上，减少无效、重复的劳动，企业也可以用比较小的成本适应快速的业务调整。

■ 注释：

业务调度——将系统的业务原子根据一定的规则进行编排并执行

Pipeline ——一种连接模式，将原本独立运行的节点连接起来

业务原子——将业务按业务数据 / 类型拆分，拆成相对独立可共用的原子能力

状态机——根据一定的规则，在有限的状态之间进行转移和动作

一、引言

互联网的快速发展，带动着企业业务系统不断优化。其中证券行业的业务系统不仅面临业务逻辑及用户体验优化的驱动需求，还受到监管政策的强制要求。监管需求往往时间短、任务重且要求高。如何在保证系统质量的情况下满足监管需求是证券信息系统的重大挑战之一。传统开发模式中，研发人员通过硬编码的方式固定死业务原子的执行顺序。如果需要修改或者更换业务原子步骤，对源代码的改造量较大，且容易产生不易发现的漏洞。本文介绍的业务调度便是基于这一背景下产生。通过引入业务调度框架，软件研发人员可以专注于实现业务逻辑。业务变化不再会造成整个业务系统流程的混乱，更不会对现有业务系统造成严重的冲击，系统运行成本也随之缩减。

业务调度框架的核心是可以调度编排的业务原子集，因此业务调度框架必须是能够支持一定顺序配置，以支持业务原子灵活顺序执行，同时业务调度框架必须提供一个统一的任务原子沙盒，让研发人员可以在业务原子沙盒中实现业务逻辑，而不用去关心沙盒与沙盒之间的执行逻辑。

本文提出基于调度框架的证券信息系统设计方法，采用 Pipeline 管道模式提供业务原子沙盒的框架结构 (Node)，同时采用轻量级的流程引擎（也称状态机）实现业务原子集执行顺序可配置化。

二、业务调度框架简介

介绍业务调度之前，先介绍下 Pipeline 管道模式。Pipeline 本身是一种设计模式，这种模式的最早发明者是 Unix 系统的创建者之一，Pipeline 模式在很多场景都有非常丰富的使用，从工业的生产流水线，比如富士康的工厂生产线，到我们日常非常熟悉的软件容器，比如大名鼎鼎 Tomcat，便是从 Engine 到 Host 再到 Context 一直到 Wrapper，都是通过同一个责任链，来传递请求；再到非常著名的通讯软件 Netty，也是基于 Pipeline 框架，框架中增加了 Handler 处理机制，由 Pipeline 驱动 Handler 进行流水线式的程序执行。

本文的业务调度框架便是基于 Pipeline 设计的，Pipeline 本身是种非常成熟的设计模式，在此基础上结合了 Netty 的 Handler 的设计思路，自主研发出一套轻量级流程引擎驱动，如下为业务调度的核心设计图。

业务调度框架主要包括两大功能：管道节点驱动，业务流程管控。两大功能对应不同的业务场景，可酌情选型使用。管道节点驱动功能，适合业务处理节点复杂繁琐的场景，可将业务拆分抽象成独立的功能单元，并进行一定顺序的编排，后续业务原子单元可以被其他业务组合复用。业务流程管控功能，适合带有流程审核的业务场景，可以保证节点的结果幂等一致性。在海通证券有很多适合运用它的场景，客户的证券开户、一柜

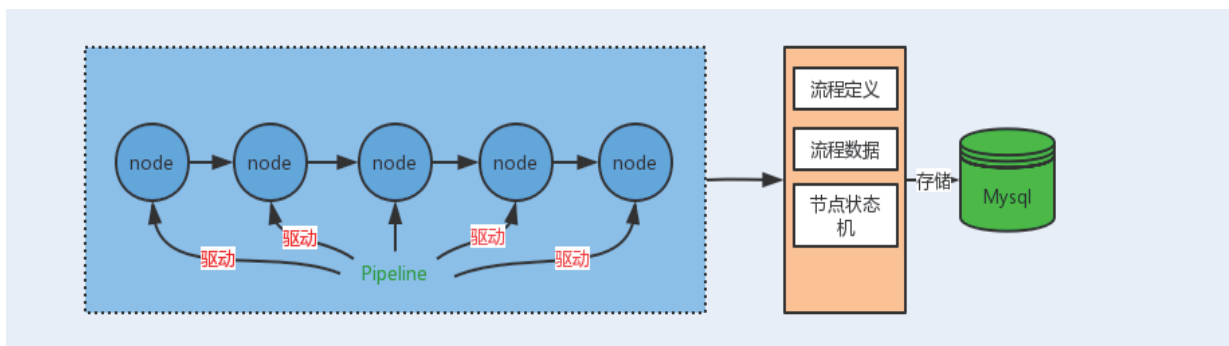


图 1 业务调度整体设计示意图

通业务等均可以使用此框架进行实施。

三、与传统应用开发模式的对比

传统的开发模式下，在一些复杂系统中，假如存在某个业务对象需要被进行繁琐的逻辑处理时，一般是在一个组件下统一处理，这确实也达到了业务效果，但方式简单粗暴，随着业务迭代的持续发展，如要在繁重无比的业务逻辑中做一些修改时，看似简单的改动可能便会无从下手，望而止步，最终系统整体缺失良好的可扩展性和可重用性，代码臃肿，可维护性逐渐变低。遵守一定的开发规范，建设可扩展性好、可维护性好的应用架构，可以有效的规避传统模式下遇到的一些问题，归纳抽象、代码精简、风格统一、流程可编排、能力可扩展是一个应用架构良好的表现，这也是本次业务调度框架的主要价值。

下面通过几个维度的对比，以列表的形式展现了传统开发模式与业务调度框架各自的特点。

四、业务调度的应用架构定位

图 2 简要的描述了下业务调度框架与业务系统的关系，以及在整体应用架构层面的定位。业务调度旨在结合具体业务场景在实际编码时提供更高效率的开发方法，提升开发效率以及后续的系统维护性。业务调度框架以二方包的形式集成到业务系统，业务调度本身自嵌入轻量级的流程引擎（支持扩展实现自定义），为业务调度实现状态保持，业务调度的集成方式是非常轻量级的，业务系统的接入改造非常灵活，框架提供了便捷的工厂方法，可以在各场景灵活使用。目前支持两种使用方式：1) 显式代码调用工厂方法直接创建并驱动；2) 与 spring 集成，通过 bean 的方式进行管理。

五、业务调度框架的设计

本节会重点阐述一下业务调度框架整体的设计与实现，业务调度框架从功能实现上来说，有

表 1

维度 \ 开发模式	传统开发	业务调度
编程风格	自由发挥，风格不一	必须遵守框架协议实现，风格统一
编程难度	前期开发自由，但系统庞大后，耦合严重，开发人员容易落入大量的程序陷阱，导致业务实现异常困难	前期需学习并掌握拆分出来的子业务开发方法，但系统庞大后，业务调度带来的灵活性，反而会极大促进业务的复用
业务执行成功率保证	依靠程序编写顺序，依靠程序对各种异常进行控制，极容易因为异常的漏判，导致业务不完整，并需由人工介入	开发人员无须编写业务的执行顺序，只需实现业务中的子业务逻辑，业务调度由框架实现并保证成功率
上下文(Context)跟踪	无上下文，依靠程序代码反推跟踪	统一管理上下文，统一输出上下文，无须开发人员刻意注意
日志调度链分析	靠开发自主使用统一日志组件，并做好埋点	框架自动记录执行上下文，统一日志风格，并提供了统一链路 traceId，作为链跟踪标示
扩展能力	无法轻易个性化扩展	业务调度框架提供了很多 Plugin 的入口，可以很容易嵌入，比如权限认证，自定义编排逻辑等功能。

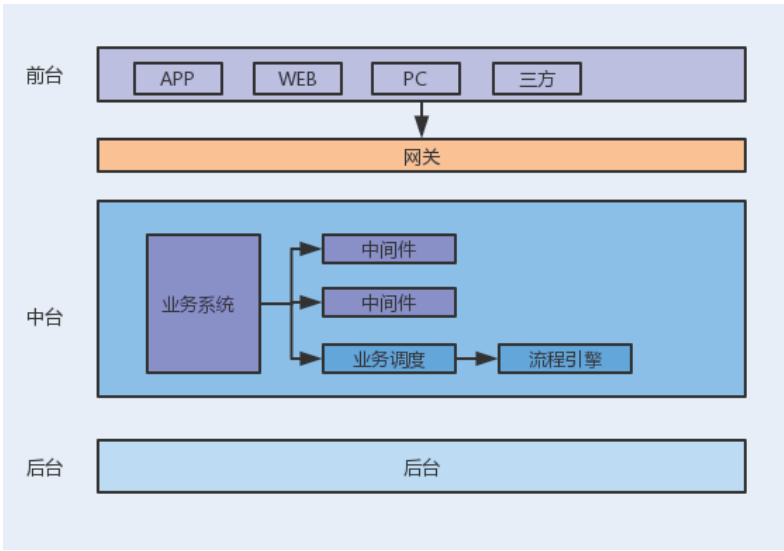


图 2 业务调度在应用架构中的定位

两大核心的模块：管道驱动、管道状态机，下面将针对这两大模块的设计方案展开描述。

1) 管道驱动模块

首先介绍几个核心概念：Executor(执行器)、Node(节点)、Context(上下文)、Listener(监听器)。一次完整的管道驱动的运转逻辑如下：管道执行器 Executor 模块，将有序的 Node 集纳入到执行器中，执行器有统一的 Context 上下文，上下文中存放着业务对象数据以及环境信息数据，Context 贯穿整个执行器的生命周期，由所有的 Node 节点共享，每一个 Node 作为执行器的基础执行单元，通过上下文获取业务参数后执行自己的原子业务逻辑，Node 的执行状态也可在状态机中保存，每个 Node 还可以挂接自定义 Listener 监听器实现前置数据检查、状态检查、权限认证、异常处理等，待 Node 节点全部执行完毕由 Executor 执行器返回业务结果。

整个管道驱动的设计见图 3。

• Executor(执行器)

Executor 是业务调度框架的整体执行器，推动整个业务节点的执行及结果处理。不同的业务场景可以定制不同的执行器，例如网上开户流程

作为一个完整业务场景，有一个针对开户业务的执行器；临柜开户全流程是另外一种完整业务场景，有另外的执行器。每种执行器均有一个管道状态机流程定义，业务调度框架通过执行器将 Node(节点) 的执行情况登记在状态机中。执行器负责整体业务节点的调度与执行，扮演的角色是整体业务进行的推动者。

• Context(上下文)

Context 上下文是由 Executor 使用，由 Node 集进行读 / 存的内存数据集。Context 采用 Key/ Value 进行存储，根据需要可再细分为各种不同业务场景的子上下文。Context 存储的内容一般为各 Node 的输入、输出，可用于 Node 之间消息通讯。可见 Context 上下文扮演的是一个数据通道的角色，是管道执行器中业务信息的载体。

• Node(节点)

Node(节点) 是业务调度框架执行的原子单元，每个 Node 代表一个相对独立的功能单元，代表一次完整的业务原子操作。为了更形象的说明 Node 的具体含义，以证券行业开户流程为例，基本资料保存、资金账号开户、三方存管绑定、风险测评保存是整个开户流程中的一系列业务操作逻辑，将其独立成单独的业务原子节点：基本

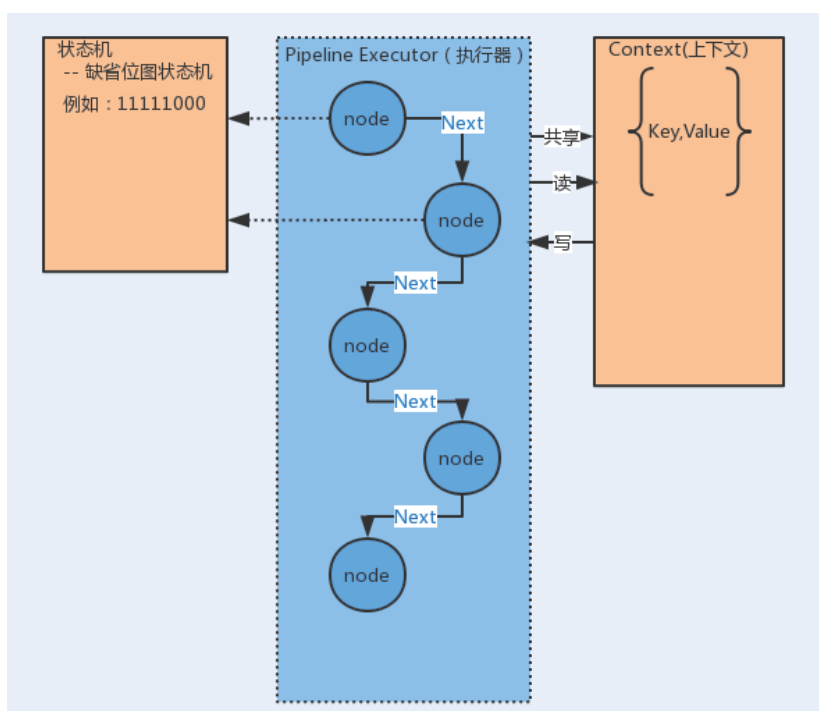


图3 管道驱动模块设计

资料保存 Node、资金账号开户 Node、三方存管绑定 Node、风险测评保存 Node。通过上述方式的细分，业务开发可以将原本耦合在一起的整体业务逻辑代码，分拆到多个独立的节点处理器中，每一个节点处理器是一个独立的类，将代码独立分隔成原子单元，这样对于后续功能的修改与维护也会非常方便，因代码修改而引入的业务风险也会大大降低。

• Listener(监听器)

Listener 是对原节点业务的进一步细分而衍生出的产物。Node 是业务调度的基础单位，是一个相对独立的一块完整业务逻辑，业务逻辑在实现的过程中，会有很多的业务判断，比如判断业务执行权限、业务执行的前置条件是否满足、业务逻辑是否在遇到错误时需要中断执行等。根据以上需求，以 Node 为宿主，衍生出了 Listener(监听器)。Listener 根据职责不同，进一步可以划分为前置条件监听器、异常处理监听器、后置监听器。

如下简单罗列了下 Node 节点大体结构的抽象类代码实现，仅供参考：

```
public abstract class AbstractHandler<T extends
Context> implements Handler<T>{
    private String handlerId;
    private boolean __flag__=false;
    public void handle(T ctx){
        __flag__=false;
        setHandlerId();
        try{
            onPreListen(ctx);
            beforeExecute(ctx);
            doHandle(ctx);
            onPostListen(ctx);
        }catch(Exception e){
            onException(ctx,e);
            if(__flag__){
                throw e;
            }
        }
    }
}
```

简要描述：

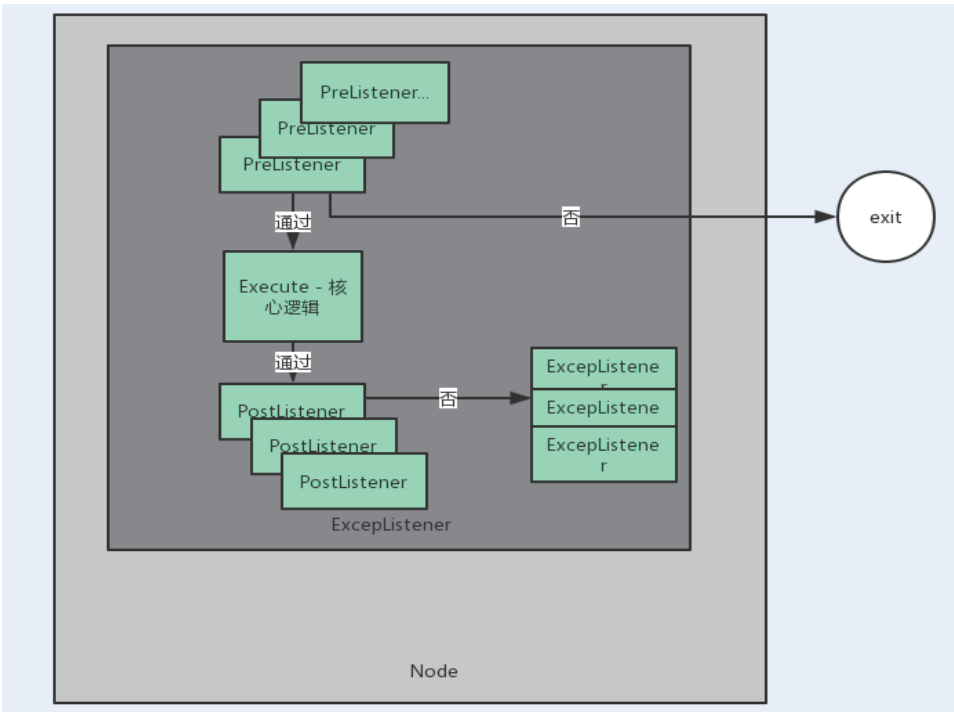


图 4 Node 内部结构设计

- 1、业务定义 Node 节点时需要继承此抽象类，对应实现 doHandle 方法。
- 2、Context 在前置监听器、执行前准备、执行过程、后置监听器、异常监听器中共享。
- 3、每个节点 Handler 都有一个唯一 Id, 该 Id 会受调度编排调度，用于控制节点的执行顺序。

2) 管道状态机模块

管道驱动模块只负责各业务节点 Node 的调度、编排，各节点在节点链中具体的执行状态情况，则由状态机负责存储，状态机通过状态控制为节点的编排提供执行的幂等保障，保证每个节点执行一次且只会执行一次。

状态机的基本流程如图 5。

本文不深入介绍状态机的实现，如下作简要说明：

- a) 上述图中共包含 ABCD 4 个节点，ABCD 为按顺序依次执行的 4 个节点
- b) 如果状态机最终状态为 1111 时，代表节点链执行完毕
- c) 4 位数字分别代表 ABCD 的执行位图，0 代表未执行，1 代表已执行
- d) 状态机可以通过对位图的判断，知道应该在哪个节点续做，也称为断点续做

六、实际运用

业务调度主要适用于拥有相对固定执行顺

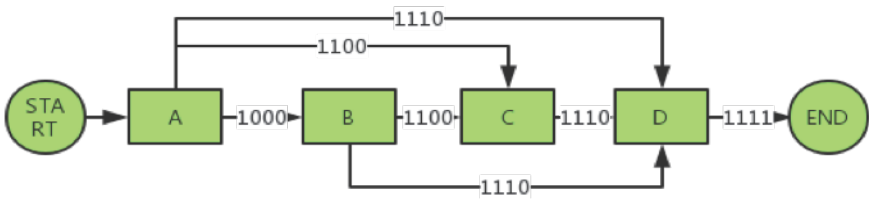


图 5 状态机流程示意图

序的业务子集的业务场景，并且这些业务子集在一定范围内有大量复用、组合的场景，这非常符合证券行业的业务规则，证券行业的业务模式相对比较固定，并且大量的业务会有类似、组合的特点，如下从实际运用角度介绍几个常见的场景。

● 网上开户流程（见图 6）

业务场景描述：

1、客户通过手机、PC 浏览器进行证券账户开户的场景，客户一般需要依次做手机号码验证、营业部选择、身份证上传、资料确认、视频见证、协议签署、风险测评、密码设置等步骤。

2、客户在中间的任何一步出现中断之后，比如客户退出 app，则下次客户继续开户时，需要自动识别客户已完成的步骤，并引导客户继续完成未做完的步骤。

3、通过多个业务步骤收集客户的信息，在后续审核环节时，则根据业务顺序将客户的各种资料依次落地生效，并最终开出证券账号。

解决办法：

（一）开户过程

a) 将如上的手机号码验证、营业部选择、身份证上传、资料确认、视频见证、协议签署、风险测评、密码设置分别抽象成一个 Node 原子节点业务，并单独实现。

b) 将如上子业务集抽象成 Node 节点类，统一纳入业务调度框架中，并配置一个状态机，用于记录用户已完成的步骤。

（二）审核环节

a) 审核生效环节需要依次将信息同步更新至后台系统，将这些环节分别抽象成子业务，封装

成对应的 Node 节点并单独实现。

b) 将如上原子业务集纳入到业务调度框架中，并配置一个状态机来记录执行状态，用于识别哪些子业务已执行，哪些未执行。

c) 业务调度通过状态机记录的状态，确保业务执行结果的幂等性，保证各个原子业务节点按照指定的顺序执行一次且唯一一次。

● 一柜通业务办理

一柜通业务办理是指证券公司各营业部统一受理客户的开户、资料变更、交易权限开通等各种业务，一柜通业务的具体受理模式与网上开户类似。营业部业务人员在现场将客户的资料、影像、视频、密码等逐个验证完之后，由集中审核岗负责业务生效，业务生效按照业务的类型，将资料、影像、权限同步各后台系统。

在此背景下，一柜通业务通过将各业务生效步骤抽象成业务原子，由业务调度框架负责业务原子的编排生效，并保证业务原子执行一次且唯一一次，通过调度框架将复杂、重复的编排任务封装起来，让业务系统开发人员可以集中精力进行业务原子的业务开发。

七、总结

业务调度框架作为基础组件，可以被运用于业务环节比较固定，业务环节执行状态需精确控制的场景。业务调度的引用在特定场景可以大量降低开发人员的重复开发工作量，并减少业务控制导致的异常，基于统一业务调度的系统在日志、调度链、上下文留痕等方面都可以有非常良好的体验。

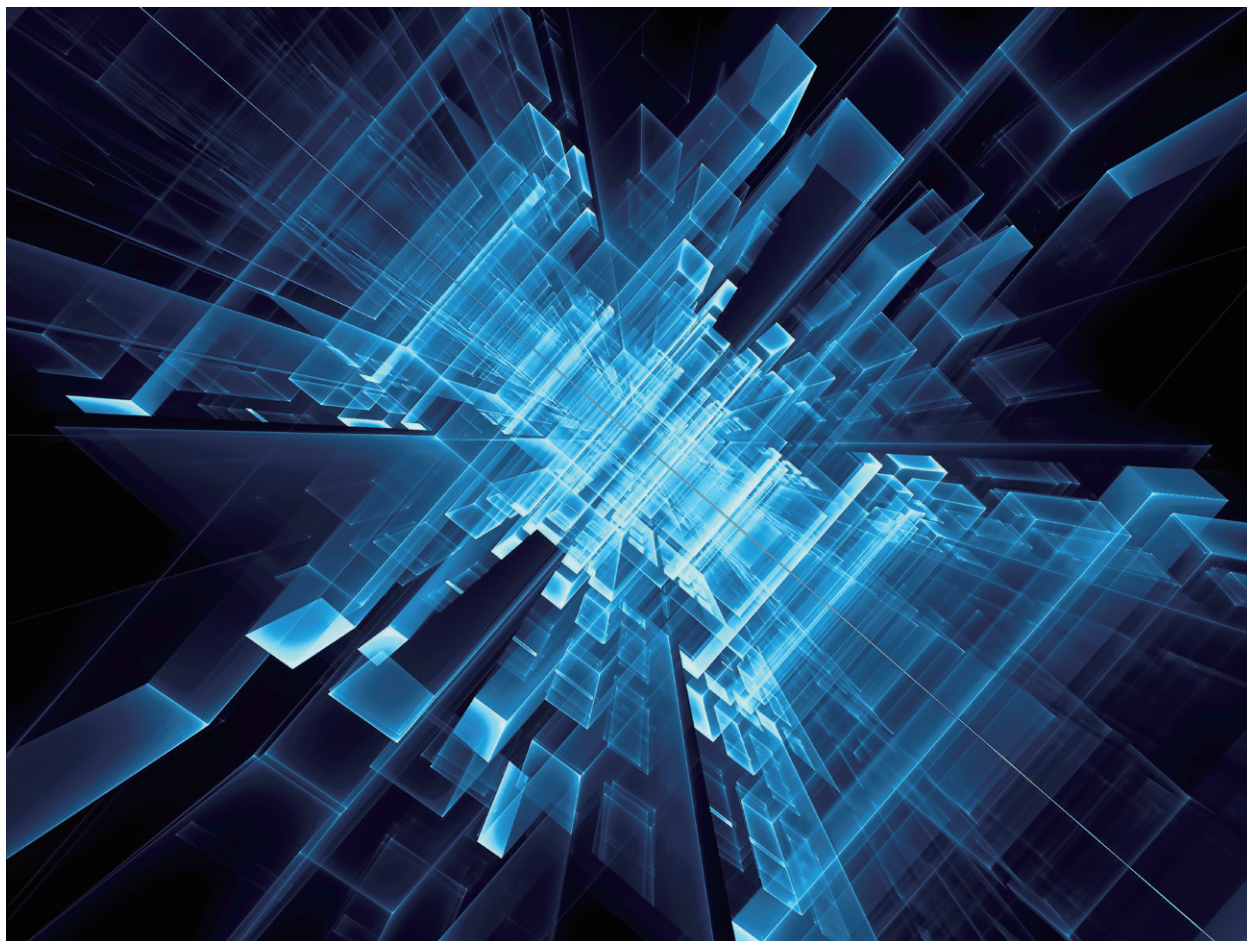


图 6 网上开户流程

运用深度学习技术处理投行项目富文档的探索与实践

陈贇 /yunchen@ebcn.com; 光大证券股份有限公司 晏强 /yanqiang@ebcn.com,
周朝阳 /zhouzhaoyang@ebcn.com、吴波 /wubo@ebcn.com、庖丁科技罗平 /Lp@paodingai.com

本文根据光大证券在金融文档智慧服务平台中的建设经验，阐述对于深度学习技术的研究认知及处理富格式文档的探索实践经验。该平台利用海量的数据信息，建立平台“数据标注 - 模型训练 - 模型评估”周期迭代方式，在模型自主学习改善复核效果的基础上，将人工审核与机器审核相结合，以提高金融数据质量监控的效率和准确度。同时，基于平台的主动周期性迭代方式，保证平台有较强的动态适应性与灵活性。



一、概述

证券行业在高速的成长发展过程中，各个业务板块都积累了非常多的信息文档，形成的文档类型和文档版本也是不尽相同。如何让计算机自动或半自动地理解自然语言文本，深度挖掘金融文档的潜藏价值，切实提升证券业员工在面临浩瀚文档资料时的工作处置效率，已成为公司内部管理方面的重要研究课题。而投行作为典型的文档密集型业务板块，涉及到的文档类型复杂、多样，且多为富文档，即包括但不限于文件扫描件、跨平台文件格式（PDF 文档、Word 文档等）和网页等，它们是标题、文字章节和段落、表格和图像等组成成分的复合有机排版。本文将阐述如何运用深度学习的技术对投行富文档进行格式化的处理，将以银行流水识别系统与投行文档智能审核系统两个运用场景出发，对其中涉及的技术做出深度剖析。

二、国内外文献综述

深度学习是近年来机器学习领域发展最为迅速的领域之一，它是基于深层神经网络的机器学习方法的别称。事实上，深度学习并不是一个全新的概念，深度学习最早可以追溯到 1943 年由神经科学家沃伦麦卡洛克和数学家沃特皮兹提出的人工神经网络^[1]。随后在 20 世纪 50 年代，感知机和自适应线性单元的提出对神经网络发展起到极大促进作用，使得神经网络的研究进入第一个发展热潮。然而这一阶段的神经网络模型为线性模型，它无法表示异或函数，这也导致了神经网络的研究热潮逐渐消退。在 1980-1995 年间，随着联结主义（connectionism）方法的流行，人工神经网络进入一个全新的发展阶段。联结主义的基本思想是：如果将大量简单的计算单元连接在一起，神经网络便会更加智能^[2]。基于这一思想神经网络在这一阶段涌现许多优秀的成果，使

得神经网络进入新的研究阶段。在以上研究的基础上，深度学习的概念在 2006 年被正式提出，它以 Geoffrey Hinton^[3] 提出的深度信念网络为开端。自此以后，得益于更强大的计算机、更大的数据集和能够训练更深网络的技术，深度学习的研究进入爆发阶段，在这一阶段研究人员能够训练以前难以训练的较为深层的神经网络，同时也在理论上说明深度的重要性。经过 10 年左右的发展，深度学习迅速席卷计算机视觉、语音识别和自然语言处理等领域，取得了举世瞩目的成就^[2]。

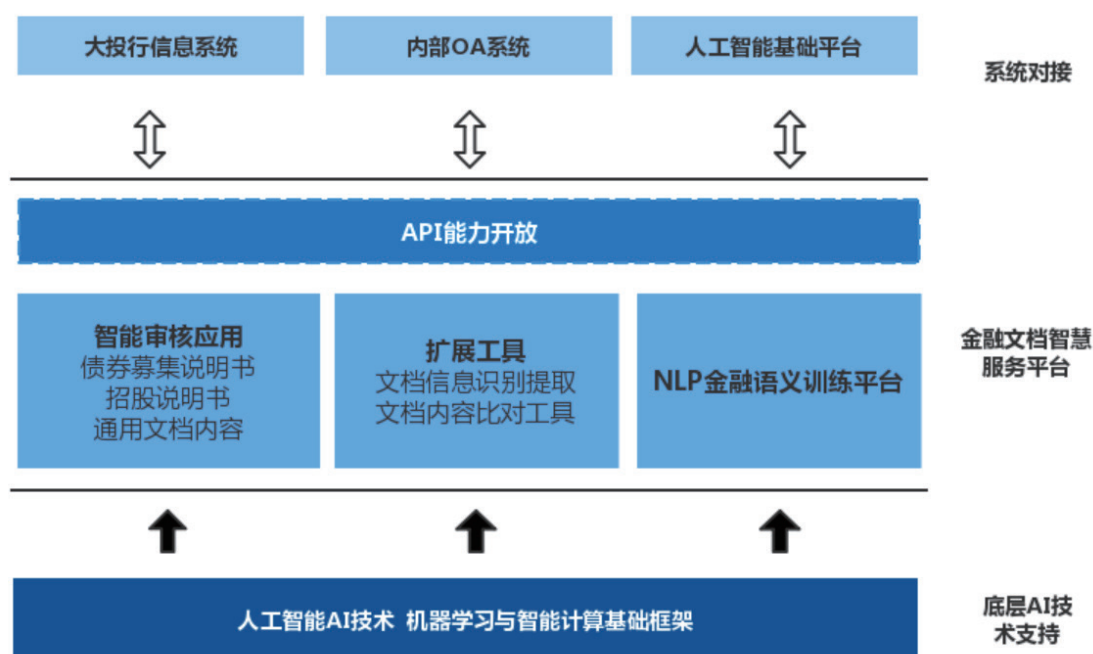
三、投行项目富格式文档的格式化处理在光大证券的运用

（一）项目背景

光大证券金融文档智慧服务平台是结合光大证券内部人工智能平台的统一建设要求，以“促进相关业务，营建智能平台”为目标，结合热点需求，在解决业务痛点的基础上，向公司内部提供简便、高效的文档综合处理平台，为公司不同部门、不同业务阶段的文档提供智能化处理工具，以利于提升工作效率，改善文档质量，规范工作流程，实现工作文档全环链向电子化方向发展。本平台选取投行业务领域的文档审核场景作为试点并积累项目建经验，优先满足标准化程度较高且使用频繁的“债券募集说明书”、“IPO 招股说明书”审核需求，后续将视平台运营以及业务发展情况，逐步扩展平台的文档审核范围。

（二）日常工作及业务现状

根据对投行板块人员的日常工作调研，发现存在大量的人工审阅工作。以债券募集说明书为例，项目人员依靠发行人和主承销商提供的材料，按照协会公布的文件报备格式要求编写募集说明书，这其中涉及的各类文件内容繁杂，极易出现像笔误、财务数据填写错误，财务指标公式编写



图表 1

错误等问题；而专职负责文档审查的员工则需要全文通读项目人员制作的项目文档及工作底稿，从几百页的文章中检查其中可能存在的笔误及各式逻辑错误；同时还需要检查多个文档之间针对相同内容的表述是否一致。上述重复且巨量的工作任务对员工是个非常大的考验，不但投入的人力时间较多，而且一旦工作状态下降，文档的审查质量也会急剧下降。

（三）建设方案介绍

1、项目建设内容

根据本期项目规划（如图表 1 所示），项目建设范围主要包括如下三个方面：

- 构建面向债券募集说明书、IPO 招股说明书的智能审核应用；配套提供智能提取、智能比对的共享工具应用。
- 搭建底层完整的智能训练及计算框架。
- 对外发布文档智能服务能力接口；支持与现有投行相关系统、协同办公等系统的对接。

2、主要采用的技术路线及创新亮点

日常工作中人们最常见、容易获取的就是文本信息，文本格式多种多样，Word、PDF、Excel、JPG 图片等等，不仅内容多样且内在格式复杂，我们称之为“富文本”。由于自身格式、内容具有高度复杂性且数据来源复杂，计算机在识别上具有很大的难度。

在 Word 文档转换为 PDF 文本之前，Word 文档中存储了内容及格式信息，我们可以通过这些信息，进行 Word 文档的编辑修改；转换为 PDF，PDF 文本是不可编辑文本，转换过程中抹除了原始格式信息，取而代之的是位置信息，即字符和字符的位置信息，与之前 Word 文档内储存的信息大不相同，我们需要将单纯的文字及文字对应位置信息，转换成机器可读懂的结构化信息，即将 PDF 转换成 Word 文档，这个过程就叫做文档结构处理。

金融文档智慧服务平台底层计算基础框架，利用深度学习及自然语言处理技术，针对富文本，进行结构化处理：

1)、文档的结构处理：

当有客观证据表明可供出售金融资产发生减值时，原直接计入股东权益的因公允价值下降形成的累计损失予以转出并计入减值损失。对已确认减值损失的可供出售债务工具投资，在期后公允价值上升且客观上与确认原减值损失后发生的事项有关的，原确认的减值损失予以转回并计入当期损益。对已确认减值损失的可供出售权益工具投资，期后公允价值上升直接计入股东权益。	
对于权益工具投资，本公司判断其公允价值发生“严重”或“非暂时性”下跌的具体量化标准、成本的计算方法、期末公允价值的确定方法，以及持续下跌期间的确定依据为：	
公允价值发生“严重”下跌的具体量化标准	期末公允价值相对于成本的下跌幅度已达到或超过50%。
公允价值发生“非暂时性”下跌的具体量化标准	连续12个月出现下跌。
成本的计算方法	取得时按支付对价（扣除已宣告但尚未发放的现金股利或已到付息期但尚未领取的债券利息）和相关交易费用之和作为投资成本。
期末公允价值的确定方法	存在活跃市场的金融工具，以活跃市场中的报价确定其公允价值；如不存在活跃市场的金融工具，采用估值技术确定其公允价值。
持续下跌期间的确定依据	连续下跌或在下跌趋势持续期间反弹上扬幅度低于20%，反弹持续时间未超过6个月的均作为持续下跌期间。

图表 2

利用模型对文档进行结构化处理：

(1) 切分段落和表格、画表格内线：

输入需要处理的文档输入系统，进行内容元素块识别，指将文档内容包括正文、图表、表格进行给定精确位置，一般是用矩形来划分，即使用左上角坐标及长宽。

运用模型将文档中每个段落及表格进行切

分；针对每个表格，结合视觉及 NLP 自然语言处理技术，对每个单元格进行划分，从而将文档切分为各个元素块。如图表 2 所示，识别出了文段、表格的精确位置。

(2) 目录结构识别：

通过文档结构化处理，将生成的一个一个的元素块内的内容，进行语义理解，例如元素块中

六、报告期内债券持有人会议的。 七、报告期内债券受托管理人履。 八、截至报告期末和上年末（或。 九、公司逾期未偿还债项。 十、报告期内对其他债券和债务。 十一、报告期内获得的银行授信。 十二、报告期内执行公司债券募。 十三、报告期内发生的重大事项。 十四、公司债券是否存在保证人。 ▼ 第十节 财务报告 一、审计报告 ▶ 二、财务报表 三、公司基本情况 ▶ 四、财务报表的编制基础 ▶ 五、重要会计政策及会计估计 ▶ 六、税项 ▶ 七、合并财务报表项目注释 ▶ 八、合并范围的变更 ▶ 九、在其他主体中的权益 ▶ 十、关联方及关联交易 ▶ 十一、承诺及或有事项 ▶ 十二、资产负债表日后事项 ▶ 十三、其他重要事项 ▶ 十四、母公司财务报表主要项目。 ▶ 十五、补充资料 第十一节 备查文件目录	公允价值、估值技术主要包括市场法、收益法和成本法。 4. 金融资产转移的确认依据和计量方法 金融资产所有权上几乎所有的风险和报酬转移时，或既没有转移也没有保留金融资产所有权上几乎所有的风险和报酬，但放弃了对该金融资产控制的，应当终止确认该项金融资产。金融资产满足终止确认条件的，将所转移金融资产的账面价值与因转移而收到的对价和原直接计入其他综合收益的公允价值变动累计额之和的差额部分，计入当期损益。部分转移满足终止确认条件的，将所转移金融资产整体的账面价值，在终止确认部分和未终止确认部分之间，按照各自的相对公允价值进行分摊。 金融负债的现时义务全部或部分已经解除的，则应终止确认该金融负债或其一部分。 5. 金融资产减值 以摊余成本计量的金融资产发生减值时，按预计未来现金流量（不包括尚未发生的未来信用损失）现值低于账面价值的差额，计提减值准备。如果有客观证据表明该金融资产价值已恢复，且客观上与确认该损失后发生的事项有关，原确认的减值损失予以转回，计入当期损益。 以成本计量的金融资产发生减值时，计提减值准备。发生的减值损失，一经确认，不再转回。 当有客观证据表明可供出售金融资产发生减值时，原直接计入股东权益的因公允价值下降形成的累计损失予以转出并计入减值损失。对已确认减值损失的可供出售债务工具投资，在期后公允价值上升且客观上与确认原减值损失后发生的事项有关的，原确认的减值损失予以转回并计入当期损益。对已确认减值损失的可供出售权益工具投资，期后公允价值上升直接计入股东权益。 对于权益工具投资，本公司判断其公允价值发生“严重”或“非暂时性”下跌的具体量化标准、成本的计算方法、期末公允价值的确定方法，以及持续下跌期间的确定依据为： <table> <tr> <td>公允价值发生“严重”下跌的具体量化标准</td><td>期末公允价值相对于成本的下跌幅度已达到或超过50%。</td></tr> <tr> <td>公允价值发生“非暂时性”下跌的具体量化标准</td><td>连续12个月出现下跌。</td></tr> <tr> <td>成本的计算方法</td><td>取得时按支付对价（扣除已宣告但尚未发放的现金股利或已到付息期但尚未领取的债券利息）和相关交易费用之和作为投资成本。</td></tr> <tr> <td>期末公允价值的确定方法</td><td>存在活跃市场的金融工具，以活跃市场中的报价确定其公允价值；如不存在活跃市场的金融工具，采用估值技术确定其公允价值。</td></tr> <tr> <td>持续下跌期间的确定依据</td><td>连续下跌或在下跌趋势持续期间反弹上扬幅度低于20%，反弹持续时间未超过6个月的均作为持续下跌期间。</td></tr> </table> 9、应收款项 (1) 单项金额重大并单独计提坏账准备的应收款项	公允价值发生“严重”下跌的具体量化标准	期末公允价值相对于成本的下跌幅度已达到或超过50%。	公允价值发生“非暂时性”下跌的具体量化标准	连续12个月出现下跌。	成本的计算方法	取得时按支付对价（扣除已宣告但尚未发放的现金股利或已到付息期但尚未领取的债券利息）和相关交易费用之和作为投资成本。	期末公允价值的确定方法	存在活跃市场的金融工具，以活跃市场中的报价确定其公允价值；如不存在活跃市场的金融工具，采用估值技术确定其公允价值。	持续下跌期间的确定依据	连续下跌或在下跌趋势持续期间反弹上扬幅度低于20%，反弹持续时间未超过6个月的均作为持续下跌期间。
公允价值发生“严重”下跌的具体量化标准	期末公允价值相对于成本的下跌幅度已达到或超过50%。										
公允价值发生“非暂时性”下跌的具体量化标准	连续12个月出现下跌。										
成本的计算方法	取得时按支付对价（扣除已宣告但尚未发放的现金股利或已到付息期但尚未领取的债券利息）和相关交易费用之和作为投资成本。										
期末公允价值的确定方法	存在活跃市场的金融工具，以活跃市场中的报价确定其公允价值；如不存在活跃市场的金融工具，采用估值技术确定其公允价值。										
持续下跌期间的确定依据	连续下跌或在下跌趋势持续期间反弹上扬幅度低于20%，反弹持续时间未超过6个月的均作为持续下跌期间。										

图表 3

公司的流动负债主要由短期借款、应付账款、票据结算、一年内到期的非流动负债和其他流动负债构成,2015-2017 年及 2018 年一季度公司流动负债的主要构成情况如下表:

表 6-20: 发行人 2015-2017 年及 2018 年一季度流动负债结构情况表

单位: 万元, %

项目	2018 年 3 月末		2017 年末		2016 年末		2015 年末	
	金额	占比	金额	占比	金额	占比	金额	占比
短期借款	2,345,796.04	28.59	2,387,109.29	31.74	409,000.00	5.73	6,937,865.43	27.94
应付票据	0.00	0.00	0.00	0.00	40,000.00	0.56	77,081.23	1.11
					1,527.59	18.35	2,147,931.43	17.51
2017 年末	发行人流动负债金额	短期借款	2,387,109.29 万元, %		1,049.06	11.85	216,319.70	10.33
					1,616.83	6.18	257,055.44	3.71
应付账款	187,069.45	2.38	248,868.84	1.98	23,021.41	1.72	459,822.11	0.72
应付利息	85,634.63	0.43	80,634.63	0.63	85,634.63	0.33	85,634.63	0.33

图表 4

存在“第一章”，即一级目录，模型会将第一章及第一章下的元素快作为同一级目录元素快，并打上第一级目录标签。

运用上述目录识别模型进行文档内各级标题的识别，自动生成文档目录，如图表 3。

综上，文档结构识别输入为富格式文档，包含的各种类型的内容元素块的序列作为输出，并保持正确的阅读顺序。

2)、语义理解

基于上述对富文档的结构识别，针对切分出的不同元素块中的数据，运用模型抽取我们所需要的信息。

(1) 二维表格

根据文档结构识别出的表格内外线划分出的单元格，进行信息抽取，每个单元格为一个数据，抽取该数据对应的行、列信息，作为表格“三元组”，从而确定该单元格数值所对应值的意义。

(2) 文本段落

识别出的文本段落，将段落中需要抽取数值的定语、属性进行对应抽取，组成一个完整的三元组。

综上所述，通过语义理解抽取出的信息，都是基于文档结构识别，正确的信息抽取都是基于准确的元素块进行划分。

3)、实践应用

(1) 文档智能提取工具

文档智能提取工具可提供文件上传模块，针对已上传的文件进行分析，提取内部的表格对象，而用户可对提取的表格进行复制等操作。同时，该工具的提取能力应通过 API 接口的方式暴露出来，供其他系统调用。

复杂排版的表格内容语义提取任务主要分为如下几个步骤：

- 根据文档页面内容，找出每一个表格区域。
- 根据表格区域，识别出表格内部的结构信息，即单元格的排布顺序以及单元格内部的具体内容。
- 根据表格的结构信息，抽取单元格的函数关系和语义关系。

在银行水单的识别应用方面，通过 OCR 技术，系统可以智能识别银行流水单扫描件中的数字以及所属类目，并进行自动分类和整理，配合

产减值准备 162,725.34 万元，具体情况如下：

表 6-18: 发行人 2017 年末固定资产累计折旧及减值准备情况表

单位: 万元

项目	账面原值	累计折旧	减值准备	账面价值
房屋、建筑物	1,409,317.23	473,294.23	888.92	935,134.08
机器设备	471,658.69	304,323.33	-	167,335.36
运输工具	212,551.07	158,086.74	6.20	54,458.13
自置飞机	11,274,128.69	4,955,575.65	149,467.34	6,169,085.69
融资租赁飞机	9,241,168.12	1,692,241.65	-	7,548,926.47
其他	2,195,328.12	1,159,894.21	12,362.88	1,023,071.03
合计	24,804,151.92	8,743,415.82	162,725.34	15,898,010.77

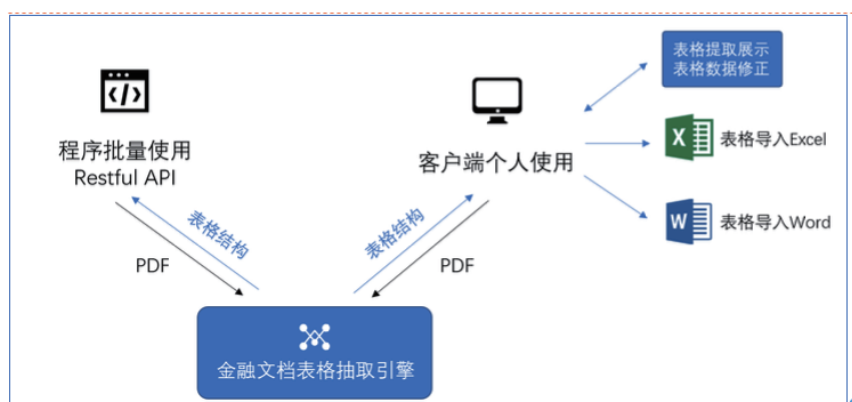
3、在建工程

截至 2015-2017 年末及 2018 年 3 月末，发行人在建工程余额分别为 196.13 亿元、290.17 亿元、315.51 亿元和 329.72 亿元。公司 2016 年末的在建工程余额较 2015 年末增加 94.04 亿元，主要是由于 2016 年自购引进新飞机导航，公司 2017 新办公大楼及北京新机场南

2018 年 3 月末 在建工程余额 329.72 亿元

4、无形资产

图表 5



图表 6

“水单异常交易识别”引擎，可以自动化的识别、提示异常交易。

在通用单层 PDF 识别应用方面，通过 OCR 技术，可以将扫描或拍照的 PDF 转化为 docx、txt、双层 pdf 等格式。将识别后的文字导出后，可用于复制、检索等后续用途。

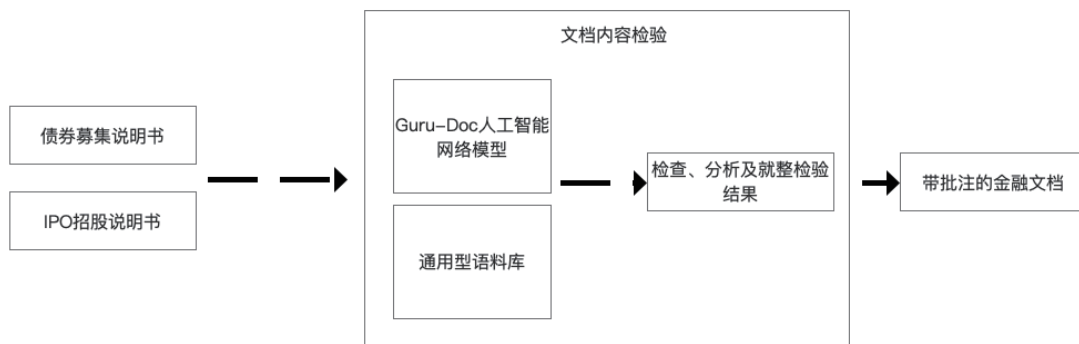
(2) 文档智能审核应用

用户可上传 Word 或 PDF 格式的《XXX 债券募集说明书》和《XXXIPO 招股说明书》，文档传输至后台运行的系统中，基于文档结构处理及信息抽取后，经过 Guru-Doc 人工智能模型网络和通用型语料库交互处理，自动锁定文档中可能存在的错误和冲突，逐条列出供市场参与者检验复核。在复核各个错误和冲突后，通过点击按钮“确认完成，生成待批注的文档”，将所有“批注到文档”的错误和冲突，以 Word

的修订批注方式，自动添加到已上传的文档。市场参与者通过点击按钮“下载文档”，即可获得带有批注的 Word 文档，支持标注后的文件导出到本地。

该功能极大程度地运用 Guru-Doc 人工智能模型网络和通用型语料库对文档内容进行检验。处理流程包括但不限于对文档中的数据 and 金融术语进行抽取和处理；对文档中笔误、财务数据填写错误，财务指标公式编写错误、数据前后对应关系等问题开展智能识别、智能溯源与智能甄别；并对文档检验中的数据进行深度结构化处理，通过对比和可视化展示，实现能够以直观的形式向用户提供展示和复核的功能，最终形成可导出的带批注文档。

系统提供错误反馈功能，用户在使用文档复核功能的过程中，可以对误报为错误的情况进行



图表 7

反馈，经过一定周期的累积，该数据可以用于优化模型训练，提高检查的质量。

（四）项目推广及运营

项目上线后，为帮助投行板块员工尽快熟悉系统功能及具体操作步骤，逐步适应文档智能审查推出后的全新工作方式，项目建设小组采取了一系列的推广运营举措：

1、在系统首页模块中内置使用说明，帮助内容包括文档及视频形式，方便用户在线观看学习。

2、邀请中科院计算技术研究所副教授兼庖丁科技联合创始人罗平博士来到现场就系统使用方法、行业发展动态、未来演进方向开展培训交流。

3、组建微信即时聊天交流群，通过实时答疑、应用技巧小贴士等形式，帮助解决业务人员在日常使用中遇到的问题。

另外，项目建设小组十分重视用户的实际使用感觉，在系统内设计了报错机制，允许用户上报系统或者业务错误；维护人员则主动收集用户上报反馈，定期开展问题诊断，同时在即时聊天交流群当中及时反馈上报问题的处理计划，让用户更加放心地使用系统；后续则通过针对性的标注及训练过程，持续优化模型，实现版本的迭代升级。

四、实践总结及展望

总结本次项目实践活动，文档智能提取应用及智能审核助手，均从文档结构识别开始，先将文档进行结构化，再通过深度学习模型提取文段表格三元组。依据三元组模型、表格识别模型、文档目录结构识别模型等模型抽取的结果，结合关联模型及算法进行相同含义的数值关联，去找出勾稽关系、一致性等冲突问题。由此可见，模型的及时迭代更新将对系统保证高准确率及召回率起到重要作用。

当前，金融文档智慧服务平台中的文档智能审核助手主要用于缓解业务部门针对募集说明书及招股说明书的一些人工审核工作量，帮助找出其中可能存在的错别字和笔误、标点及格式问题、数据勾稽问题、比例计算问题、财务指标计算问题及底稿一致性问题等；而文档智能提取工具主要针对银行流水进行表格抽取及识别，将 PDF 格式的表格扫描件，转化为可复制的表格，并支持导出相应的表格 Excel 文档，方便投行同事审核及使用。

未来，金融文档智慧服务平台审查范围将不仅限于财务方面，也会逐步增加一些行业及业务规则方面的审核；同时扩展支持更多的文档类型，成为一套更为智能、更为全面的系统。

证券公司信息系统用户权限集中管理的研究与实践

刘国杨 王海航 / 民生证券股份有限公司 E-mail : liuguoyang@mszq.com



证券公司信息系统数量众多缺乏统一管理，人员岗位调整或离职后信息系统用户权限不能及时更新，不仅影响业务开展，更成为系统安全隐患。本文对证券公司信息系统用户权限管理的现状进行了调研分析，提出了一种与证券公司现有的人力资源管理系统和办公流程管理系统整合的信息系统用户权限集中管理解决方案，并结合实际案例说明如何通过开放 API、适配器和 RPA 等多种技术实现与现有信息系统的对接。

一、概述

经过数十年的发展，信息技术已经成为证券行业发展的重要技术之一。每个证券公司内部都建立起了大量的信息化系统，涵盖了绝大多数的业务领域，如自动化办公、人力资源、财务管理、客户交易、风险控制、法人清算等等，为公司业务的开展提供了重要的保障。但是我们注意到在这种新形势下，信息系统的权限管理出现了新的问题。

首先，由于大量异构的信息系统缺少统一有效地管理，在面对公司人员岗位变动或入离职等情况时，往往无法及时对系统内部账户权限进行及时地更新，一方面影响了业务快速开展，同时也为系统的运行埋下了安全隐患。

其次，随着证券行业监管力度不断地加强，对信息系统的权限管理提出了明确的要求。要求证券公司的风险管理部，对公司内部信息系统权限进行严格的内控管理。但是由于各信息系统权限分散管理，内控人员往往无法快速地厘清不同部门、职级、岗位的人员在各信息系统中所拥有的权限，难以及时响应监管的各种具体要求。

针对上述问题，人们首先提出了统一用户认证的方案，即通过对身份统一认证来控制用户访问信息系统的权限。但是我们在实际应用中发现，大量的信息系统权限不只限于系统访问权限，还包括了大量的角色、功能、数据操作等权限控制，仅仅依赖统一用户认证来管理权限是远远不够的。另外，统一用户认证系统的使用，需要对接的信息系统进行相应的改造，包括账户数据同步、认证功能对接等等。对于大量依赖外购信息系统的证券公司而言，这些工作不仅涉及大量的协调工作以及资金成本，也会对现有系统稳定运行带来一定的影响。

因此，在本文中我们提出了建设一套权限集中管理系统，作为一种替代方案，以解决目前权限管理的问题。

二、总体思路

通过实践摸索，我们认为建设权限集中管理系统，并充分发挥其管理作用，必须从以下几个方面着手。

1) 权限集中归集

通过源头对权限数据进行归集，降低权限归集的复杂度。传统的权限数据归集一般是通过与信息系统直接同步实现。这种方法需要权限管理系统与各信息系统建立了大量的耦合关系，给归集工作带来了较大的难度。

随着今年来各证券公司 IT 管理能力的提高，权限申请和分配逐渐实现了电子化和流程化，也为权限数据从源头归集打下了基础。

2) 权限集中管理

通过开放灵活的数据模型设计，实现对不同类型、不同来源权限数据的统一管理。公司员工在所有信息系统中的账户和权限尽可能的集中存储。任何权限的变更都可以通过权限集中管理系统进行管理或者被其所感知。

3) 权限统一下发

在对权限数据进行集中存储的基础上，利用技术手段实现权限统一、自动化的下发。确保信息系统能够及时获取到权限集中管理系统中的各种权限变更。这种方式与传统的基于人工手动配置的方式很大的不同，降低了人工操作风险，也避免了跨系统、跨部门的大量协调工作，提升了管理效率。

三、系统架构

根据实际情况，我们规划的权限集中管理系统总体架构图如图 1 所示。

整个权限集中管理系统可以分为七个组成

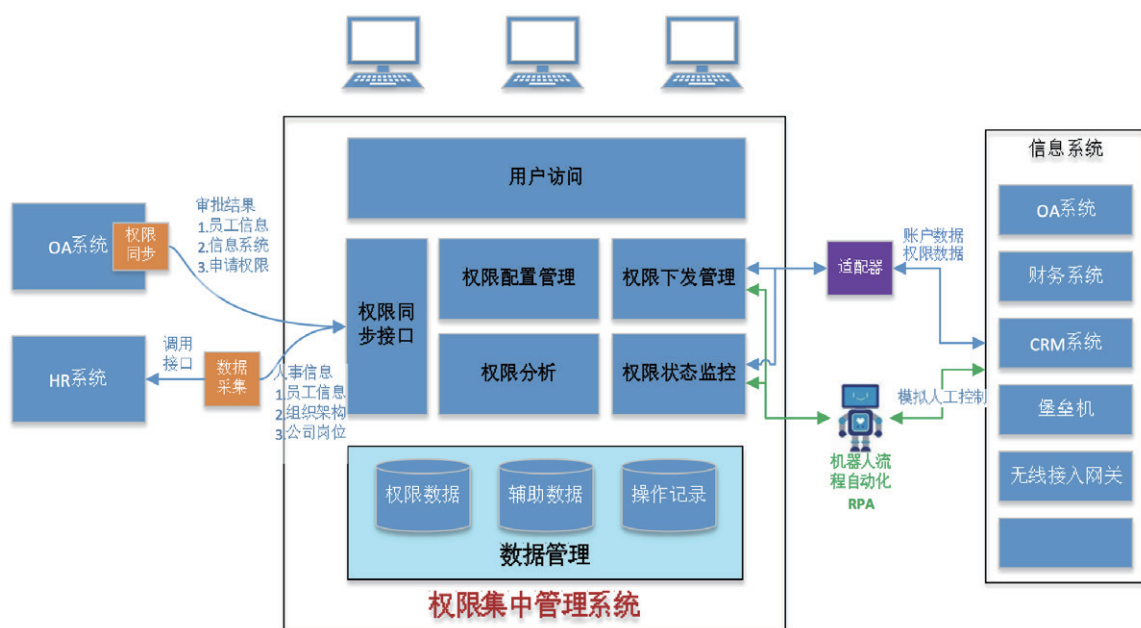


图 1 权限集中管理系统结构图

部分：

(1) 数据管理：对各个信息系统的权限数据进行规范化存储管理。在权限集中管理系统中，管理的数据包括权限数据、辅助数据、权限变更的操作记录数据等。

(2) 权限同步接口：权限集中管理系统对外发布了一组权限数据的同步接口，用以接收来自于 OA、HR 系统的权限数据。在 OA 系统中，内嵌权限同步模块（Servlet）。当权限申请流程完成审批后，审批结果（包括员工信息、所属信息系统、所申请权限等）将通过该模块调用权限同步接口写入到系统中。对于 HR 系统，单独建立了数据采集程序，定时调用 HR 系统接口，获取最新的人事信息（包括员工信息、组织架构、岗位信息等），并将这些信息通过权限同步接口写入到权限集中管理系统中。

(3) 权限下发管理：当用户的权限发生变化时，该模块将触发权限下发任务，通过调用不同信息系统所对应的适配器，将最新的权限信息写入到指定的信息系统中。不同的信息系统，根据自身的状态，将配置不同的适配器。

(4) 权限状态监控：通过多种技术手段，对

各个信息系统中员工的最新权限状态进行监控，确保与权限集中管理系统中记录的权限数据保持一致。

(5) 权限配置管理：为系统管理员提供手工配置权限的功能。

(6) 权限分析：基于集中管理的员工账户权限，进行权限状态分析监控。包括权限操作的跟踪记录、账户权限查询等功能。

(7) 用户访问：提供权限查询配置的用户图形接口。

四、相关技术研究

下面我们将对权限集中管理系统中涉及到的相关技术进行介绍。

（一）权限数据分析

1、权限数据分类

从用途角度来划分，权限数据又可以分为权限数据、辅助数据和日志数据。

权限数据用于表示指定账户在特定系统中所享有的权限。通常与具体的信息系统中的权限

数据一一对应。一般的信息系统通常基于经典的 RBAC 模型构建账户和权限功能。在这个模型下，权限的数据被分为了用户、角色、权限三个维度，基本上满足了独立信息系统绝大多数应用场景的需求。但是，从满足证券公司所有信息系统权限管理角度出发，我们建议将权限数据进一步扩展为“应用、组织、用户、角色、权限、资源、操作”七个维度。

辅助数据包括员工信息、公司组织架构信息、岗位（包括主岗和兼岗）信息、权限配置规则等等。其主要作用是辅助权限数据的管理。

日志数据记录了权限的变更情况，包括对权限的创建、变更、删除、分配等。这些信息将为公司内控管理提供有力的支持。

表 1 为权限数据分类情况说明。

2、权限数据的来源

权限数据根据产生的方式，大致可以归为两类：(1) 根据员工人事信息自动生成的权限；(2) 根据业务需要动态申请的权限。

对于第一种来源的权限，主要是根据员工的人事信息，如岗位、组织部门等信息，根据预先制定好的权限分配规则，直接为员工分配的权限。这类权限常见的有：OA 系统使用和各种流程审批权限、邮箱使用权限、办公网络接入权限等等。由于人事信息在证券公司通常存储于 HR 系统，因此，一旦人事信息与权限分配的对应关系确定后，HR 系统将是自动分配权限数据的主要来源。

对于第二类动态申请的权限，通常并不是事先约定好的，而是需要当事员工或者部门根据实际情况，依据相应制度规范，向特定信息系统的管理部门进行申请获得。例如，公司员工为办理某项业务时，向相关部门临时申请信息系统的使

表 1 权限管理数据分类

序号	大类	子类	说明
1.	权限数据	应用（系统）数据	根据不同信息系统对权限进行分类
2.		应用（系统）	根据系统来区分权限
3.		组织机构	基于公司组织机构对权限数据进行分组归类
4.		用户	信息系统操作账户，通常与员工对应
5.		角色	信息系统设定的用户角色，通常与岗位对应
6.		权限	对应信息系统的权限，实际操作中可以暂时忽略
7.		操作	代表对信息系统资源的操作，如增、删、改、查
8.		权限数据	对应信息系统的权限，实际操作中可以暂时忽略
9.	辅助数据	员工信息	包含员工的姓名、工号、联系方式等
10.		组织机构	公司的组织机构信息
11.		岗位	包括公司设定的岗位信息
12.		权限分配规则	权限自动化分配的规则
13.	日志记录	权限变更记录	对权限进行创建、删除、变更时记录日志
14.		员工权限分配记录	为员工分配权限时记录的日志

用权限。通常情况下，权限申请流程一般都在证券公司的 OA 系统中进行，因此 OA 的权限申请流程的处理结果就成为动态申请的权限的主要来源。

3、权限数据范围和粒度

我们注意到，由于各信息系统涉及的业务不同，差别是非常大的。例如，诸如堡垒机、防火墙等基础平台类系统，通常对角色的定时比较简单；但如集中交易账户系统、O32 系统等，对角色以及角色所拥有权限的却有着非常精细的划分。而且大多数下信息系统的权限只需要管到功能即可，而诸如 CRM 等类型的系统，却又要求对数据记录的进行严格的权限控制。

在实践中，建议在建立权限集中管理系统初期，对权限管理的粒度可以细化到用户、角色（岗位）一级，而对于角色下的具体功能甚至是数据权限的管理可暂时忽略。这样设计尽可能地避免了各系统因为业务的不同，导致权限管理难度的增加。经过实践证明，这种方式也满足了证券公司权限管理大多数场景的需求。

另外，除了针对权限数据进行管理以外，权限集中管理系统还需要记录详实的权限变更操作记录。这些留痕信息可以直接为公司信息技术审

计工作提供支持。

（二）权限归集

根据前节分析，权限数据来源主要由两个部分组成：HR 系统和 OA 系统。

1、HR 系统

从 HR 系统主要是采集根据人事信息，并根据预先设定好的规则，在权限集中管理系统中自动生成权限。

一般证券公司的 HR 系统都以外购为主，通常都提供数据查询接口，通过这些接口，权限集中管理系统能够及时捕获到员工人事信息的变化情况，并根据预先设定在权限集中管理系统中的权限分配规则，及时更新权限信息。

对于部分数据查询不完善的 HR 系统，可以通过部分开放后台数据库表或视图的方式，由权限集中管理系统主动查询获取人事数据。

2、OA 系统

通常在公司 OA 系统中，已经权限管理制度设立相应的权限申请审批流程。任何信息系统的权限创立、变更、废止，都是通过相关岗位、业务部门、以及风控部门进行审批。因此，可以通过捕获 OA 系统权限申请审批流程的处理结果，

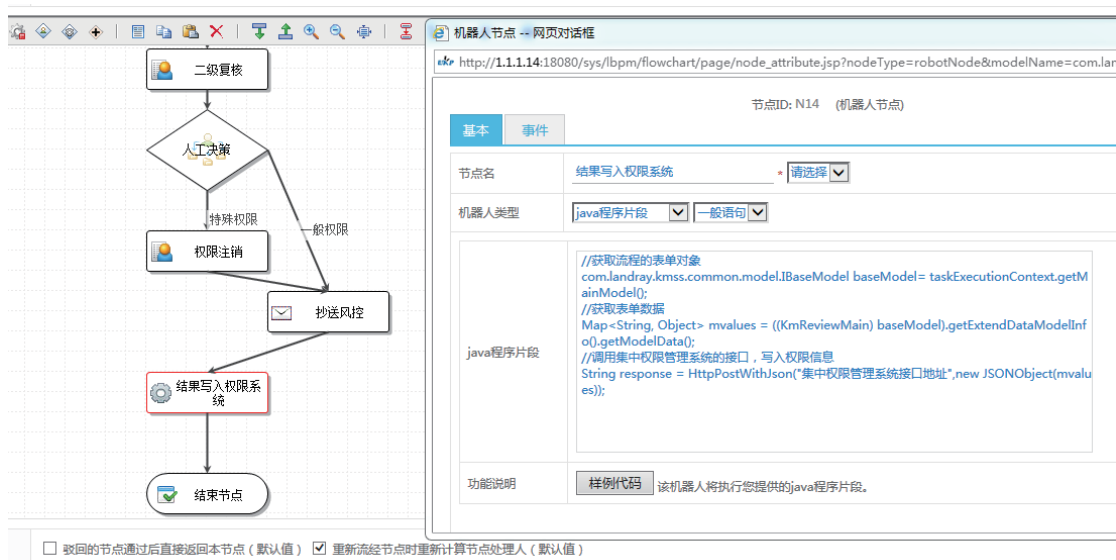


图 2 蓝凌 EKP 平台流程模板编辑

从而获取权限数据。

以蓝凌的 EKP 平台为例, 该系统的流程引擎提供“机器人”处理节点, 在权限申请审批流程结束时, 可以配置相应的“机器人”处理节点, 通过该节点调用 Java 程序片段手动获取审批流程表单中的权限数据, 并利用权限集中管理系统的接口, 将权限变更数据写入到权限集中管理系统中。如图 2 所示。

表 2 流程机器人节点 Java 处理代码

```
//获取流程的表单对象
com.landray.kmss.common.model.IBaseModel baseModel=
taskExecutionContext.getMainModel();
//获取表单数据
Map<String, Object> mvalues = ((KmReviewMain)
baseModel).getExtendDataModelInfo().getModelData();
//调用权限集中管理系统的接口, 写入权限信息。httpPostWithJson 接
String response = httpPostWithJson("权限集中管理系统接口地址",new
JSONObject(mvalues));
```

(三) 权限下发

权限在归集到权限集中管理系统后, 需要及时地下发到指定的信息系统中。通常的实现方案一般包含信息系统主动拉取和权限管理系统推送两种方式。

(1) 拉取: 主要是在信息系统端增加访问权

限集中管理系统的模块, 定时从权限集中管理系统中获取最新的权限数据, 并更新自身系统的相应权限。这种方式适合正处于规划中或者正在建设的信息系统, 要求系统进行部分改造以实现与权限集中管理系统对接。对于已经建成的系统, 由于需要协调系统供应商原有的开发团队资源, 往往实际操作困难较大。

(2) 推送: 则是指权限集中管理系统调用各信息系统的接口, 定时或实时地将最新的权限变更数据发送到信息系统中。这种方式要求权限集中管理系统针对不同的信息系统都有不同的权限数据推送模块。随着接入的信息系统数量不断增加, 会导致权限集中管理系统的复杂度大大地增加。

在实现拉取或推送时, 为避免频繁接口调用带来的性能问题, 同时降低系统与系统之间的耦合性, 我们进一步引入了消息中间件 (如 Kafka、RabbitMQ 等), 利用发布/订阅的模式来发布权限数据。这种方法混合了上述两种方式, 如图 3 所示。

无论是拉取还是推送方式, 都会对原有的系统造成较大的改动, 带来较大的成本开销, 同时也对各系统的稳定性造成影响。随着近年机器人流程自动化技术不断成熟, 为上述的问题的解决提供了一个方案。



图 3 发布/订阅模式发布数据

(四) RPA 技术的应用

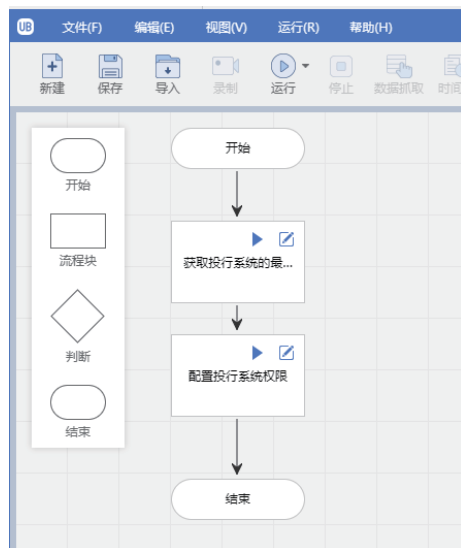
无论是拉取还是推送方式，都会对原有的系统造成一定的改动，也会带来一定成本开销。随着近年机器人流程自动化技术不断发展，该技术通过模拟人工的方式对信息系统操作，为我们解决上述问题提供了一个新的思路。因此我们也在在这方面进行了一些探索。

所谓机器人流程自动化技术 (Robotic Process Automation, RPA)，是利用驻存在电脑中的流程自动化管理软件，遵循事先拟定好的脚本流程逻辑，对于电脑中的多种业务软件进行自动操作，从而实现办公自动化的系统。

目前国外主流的 RPA 平台有：UiPath、Automation Anywhere、Blueprism 等。国内的有阿里云 RPA、UiBot、金智维等。我们在权限集中管理系统中采用了 UiBot 平台，实现了各信息系统的对接。详细的技术架构如图 4 所示。

如图所示，权限集中管理系统实时地将各个信息系统变更的权限数据发布到消息中间件。RPA 机器人通过订阅的方式，从消息中间件获取各个信息系统最新的权限变更情况，然后利用信息系统的管理员账号，模拟登录进入信息系统，并完成新变更权限的配置工作。

这里以投行项目管理系统为例，简要说明如何利用 RPA 实现新增权限配置的流程。流程主要分为两步：(1) 从权限集中管理系统中获取关



于投行项目管理系统的最新权限变更申请。(2) 使用管理员账号模拟登录投行项目管理系统，并配置相应账号的权限。下图为 UiBot 开发环境的设计截图：

图 5 UiBot 流程图

两步流程在 UiBot 中的执行脚本如图 6、图 7 所示。

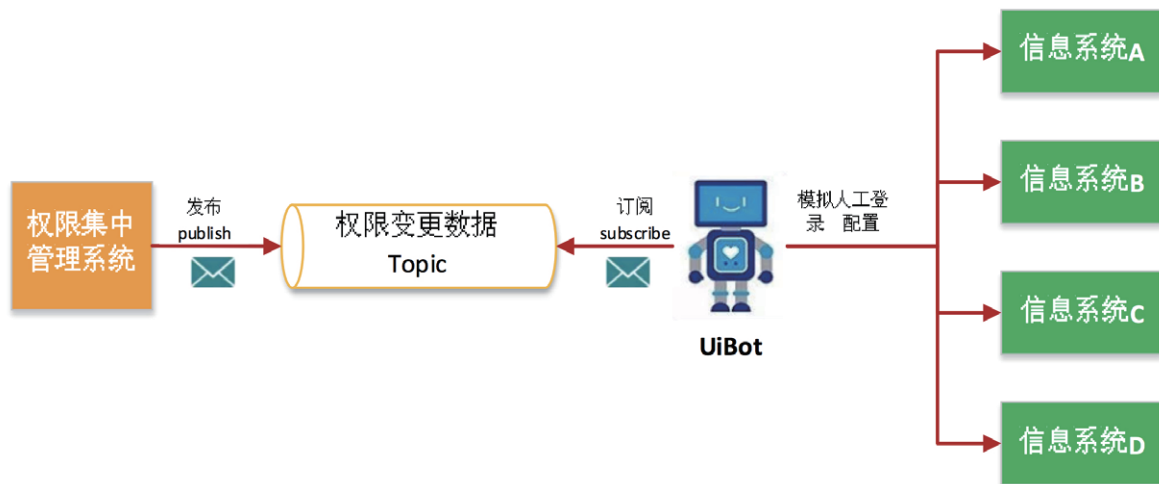


图 4 RPA 实现权限下发



图 6 从权限集中管理系统获取所有权限

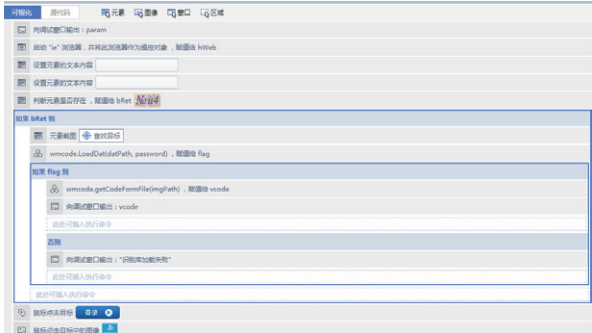


图 7 模拟操作员账户配置用户权限

根据设计完成的脚本，UiBot Worker 将会在 UiBot Commander 的调度下模拟系统管理员的操作，实现在投行项目管理系统中权限的配置更新。这种方式避免了对旧有信息系统侵入式改造，也避免了因为改造而引起的各种系统风险，从而达到节省人力、降低开发成本的目的。

不过在实际应用中，我们也发现 RPA 并不能适应所有的技术环境，例如现有版本的 UiBot 不支持基于 Flex 实现的应用系统。因此，是否采用 RPA 技术需要根据实际情况来判断。

五、总结

本文介绍了一种权限集中管理系统的建设方案，用于解决目前证券公司信息系统权限管理的问题。通过对这个方案的实施，验证了该技术可行性。这也为其他证券公司加强信息系统权限管理，提升自身内控能力提供了一种新的思路。下一步，我们还将围绕现有的研究成果，通过实践不断的完善现有的权限管理系统功能，并结合业务需求进一步展开深入的研究。



0 行业观察 bservation

9 全球交易所行业区块链最新发展及思考

10 券商证券交易系统质量评估框架

全球交易所行业区块链最新发展及思考

徐广斌 / 上交所资本市场研究所 gbxu@sse.com.cn

牛壮 / 上交所市场监察二部 zniui@sse.com.cn

徐丹 / 上交所技术公司 dxu@sse.com.cn



因比特币而兴起的区块链和分布式账册技术，因其去中心、分布式、高效、安全、可追溯、不可篡改等特性，为互联网时代创造了一种不依赖于第三方的全新价值交换方式。近年来，着眼于区块链的独特价值和交易所在市场组织方面独有的天然优势，全球各大交易所纷纷加大对区块链的投入，使区块链成为行业金融科技发展的热点之一。本文对全球交易所行业的区块链发展情况进行综述，并对我国证券行业区块链发展提出建议。

■ 本文仅代表研究人员个人观点，与所在机构无关。

相关研究和案例表明，区块链技术可以使得电子金融交易流程变得更加公开、透明，通过区块链技术，使得原本高度依赖中介和集中式处理的传统交易、结算模式，变为不受时空限制的自组织网络交易模式，这不仅使参与者的范围更广、服务更便捷、效率更高、成本更低，也使得监管、风控和合规更加智能化和自动化。

从纳斯达克 2015 年推出私募股权服务平台 LINQ 为开始，全球各大交易所纷纷搭建区块链平台、探索区块链应用。本文对境外主要交易所区块链的最新发展和应用实践情况进行跟踪，并对证券行业未来区块链的发展提出了建议。

一、区块链和分布式账册技术简述

分布式账册（DLT）是一个对等虚拟网络，由一系列按共识方式定义参与者权责的合约所构成，这使信息的点对点传递、同步和决策无需中央管控或获得第三方授权。分布式账册的数据库分布于对等网络的节点上，各节点都复制及存储与总账本完全相同的副本并独立更新，当一笔交易发生时，每个节点都将执行一次更新，同时网络中所有节点以共识机制投票决定哪个账本是正确的。一旦达成共识，其余所有节点都将按照正确副本的数据进行更新。账本共识算法包括工作量证明、持有量证明和代持量证明等类型。

区块链（Block Chain）是分布式账册的一种具体实现形式。通过构建自组织网络、分布式数据存储，记录时间有序、不可篡改的加密账本，并利用分布式共识机制，实现全网记账、共同公证，支持信息的点对点传输。区块链借助密码学串接并保护区块内容，每一区块包括了前一区块的加密散列，时间戳和交易数据（通常为—散列（Hash）值）。用区块链技术串接的分布式账册，使交易双方可以有效地记录和查验交易记录。加密数字货币则是建立于区块链上的，分散式账本的一种实际应用，比较知名的有比特币、以太坊

等，其中比特币地址对应一个账户，比特币数量相当于金额。

区块链的三种基本形式是公有链、私有链和联盟链。公有链是去中心化分布式区块链，所有人可以不受限制参与，参与者可以自由访问、发送接收和认证交易，开放性和匿名性强。私有链具有分布式特点，但仍有中心。中心控制节点负责对可以参与交易和进行验证的参与者范围进行授权。联盟链也称为广义私有链，介于公有链和私有链之间，由若干机构联合发起，按照一定规则预先指定参与者范围，具体有多种实现模式。系统内交易确认节点一般也需预先指定，并且通过共识机制确认。取决于联盟内部的信任程度和需求，交易可选择匿名或具名方式。联盟链便于进行控制权限设定，比私有链具有更高的应用可扩展性，对于跨机构、跨境的清算、交收、审计等较大的应用价值。总体上，但联盟链的跨链操作可造成效率的下降并影响互操作性。联盟链减少了人工干预和信任成本，可以大幅降低传统异地结算的成本和时间，比传统系统效率更高，是区块链目前主要的应用方向。

区块链具有以下主要特征。一是分布式去中心化。加密数据在区块链上分散保存在接入区块链各节点的服务器中，每个节点都保存一套完整的区块链总账，各节点能查看全部交易信息。每当区块链更新交易信息，链上所有节点都会同步更新相关数据。分布式结构可以实现点对点之间的交易，使证券在发行、交易、结算方面均可绕过中介机构，提高交易的透明度和效率、节约成本。

二是无需中央信任。传统的信用系统中，在交易过程中涉及的参与方都需要对中央机构有足够的信任支持，而区块链技术使用技术手段来赋能信用，通过算法实现约束，可以不依赖中央权威机构信用背书。降低了信用成本，减少了处理环节，提高了交易安全性。

三是加密不可篡改。区块链使用密码学、时

间戳等技术，每个区块严格按照时间顺序递进，这种时间上的不可逆有效地保证了任何篡改数据的行为都很容易被追溯。同时，区块链的共识机制对这些数据代码共同进行维护。当少数节点出现错误、篡改，只要大多数节点保持一致，整个区块链账本的真实性仍能得到保证。

四是易编程可扩展。区块链是底层技术，以此为基础可以实现各类扩展和去中心化、去信任的应用。当有新的机构要加入区块链成为节点时部署起来方便，不影响链上已有节点的正常运行。另外，区块链可以通过编程设置交易条件，当交易不满足预先设定的条件时，代码阻断机制可自动触发以限制交易的进行。区块链上这种可编程代码被称为智能合约（Smart Contract），可方便地提供验证及执行有条件的处理，如交互、存储、转账等，允许在没有第三方的情况下进行可信交易，可追踪且不可逆。这一技术可以很方便地满

足金融市场各类产品及业务的复杂需求。相较而言，在传统模式中，风险控制往往比较滞后，难以及时地发现风险并进行快速处置，而通过智能合约则可以满足风控和监控及早发现、及早处置的需求。

根据著名的 Gartner 咨询公司在 2019 年的判断，区块链技术的发展正方兴未艾，接下来的一个时期区块链将进入与各行各业广泛、深度融合发展的一个阶段，尤其对于泛金融行业来说，未来 5 到 10 年区块链技术与传统金融业务的快速融合将有望产生重大突破。

当然，区块链作为一种新兴技术，尽管还在不断快速演化发展中，但本身也不是完美的，特别是应用于证券交易行业时面临以下的局限。

一是全账册存储模式对系统的容量造成挑战。在典型的区块链中，若要保障记账的准确和完成对记账人的监督，网络中必须存在大量的全

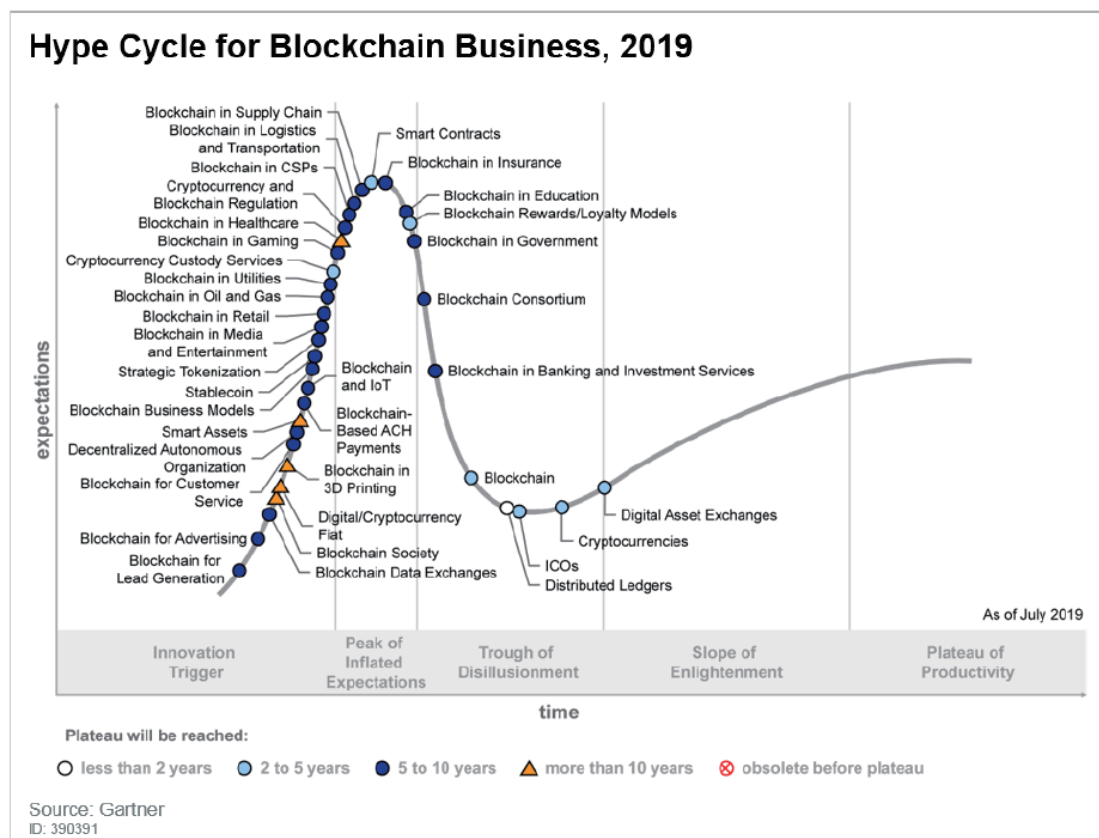


图 1 Gartner 公司 2019 年区块链技术成熟度分析

数据节点。这些独立节点上，需要记录该区块链应用初始以来的所有区块数据，即记录全部成交数据。目前，大型证券机构尤其是核心机构每日产生成百上千 GB 的海量数据，这对保存这些数据所需要的存储容量造成压力，相应的存储成本和管理成本不容小觑。

二是处理效率难以支持高性能处理需求。由于采用的分布式存储，且网络中发生的任何一笔交易其它节点均需进行认证并记录，系统达成共识的效率低，尤其对于大规模的节点、或者大数据量的情况下性能会急剧下降，最多支持每秒几千到上万笔处理性能。对于对处理速度和实时性要求高的业务，譬如证券交易所股票集中交易撮合来说，远远达不到目前证券核心交易的近似性能。需要在共识算法、硬件支持等性能关键的技术手段上获得革命性突破才具有可替代的可能。

三是安全及高可用性。作为一种全电子化网络，区块链不可避免面临传统的计算机及网络的仿冒、篡改等信息安全风险，同时，在完全去中心化架构下，由于区块链的监管依靠网络中所有的节点共同完成，因此理论如果掌握全网超过 51% 的算力就有能力成功篡改和伪造区块链数据。

二、境外交易所区块链最新发展及实践

近五年来着眼于区块链及分布式账册技术在促进数据共享、优化业务流程、降低运营成本、提升协同效率、建设可信体系等方面的优势，境外主要交易所纷纷加大投入，积极探索，推动区块链和交易所业务深度融合，促进服务价值和模式创新。

对于交易前业务，区块链主要用于公司分析和风险管理和控制。区块链使得获取真实可靠信息的效率更高、透明性更强，搜索成本更低。区块链的智能合约便利了在发生特定事件，如风险

事件时，可以自动采取相应的警示或处置。

对于交易中业务，区块链可以使传统交易所的服务时间延展至 7×24 小时 $\times 365$ 天全天候不间断，并使物理服务范围扩展至跨地域、跨境、跨时区，也支持突破分业边界，形成跨行业、板块、机构的多层次一体化交易基础设施，结算可方便地从每日盘后批处理方式转为逐笔实时进行。

表 1 区块链在交易所行业相关业务链的应用前景分析

	业务环节	主要局限	区块链的作用
交易前	公司研究分析 风险管理	信息可获取性、传递效率、及有效性	可追溯、穿透至底层资产，自动验证归属，信息有效性、不可篡改、可追溯，条件触发自动处置
交易中	订单执行及撮合	运营时间限制、处理结算	全天候、实时处理 跨机构、跨行业、跨境
交易后	清算交收	对手方风险、机会成本、融资成本	消除一对一风险降成本，高效、实时、无缝处理资产转移和变动
	托管 资产管理	操作成本高 需要人工处理及确认	交易流不可篡改，无需第三方保管账目 自动执行合规，减少操作成本

2.1 纳斯达克交易所

2015 年 5 月，纳斯达克宣布筹建 LINQ——一个基于区块链网上发行 Pre-IPO 股票的平台。同年 11 月 LINQ 正式发布，成为首个通过区块链平台对资产交易进行数字化管理的产品。它面向纳斯达克私募股权市场，为创业者和风险投资者提供便捷服务。通过，私募发行方可以获得代表所有权的数字记录，并且以完全在线的方式与私募投资者之间完成一笔交易。对私有公司而言，它是一个全新的股权管理工具，作为纳斯达克私人股权市场的一部分，为企业家和风险投资者所准备的完整解决方案的一部分。2016 年 4 月，区块链初创公司 Chain 允许私人投资者在上购买其发行的股票。

2016 年 5 月，纳斯达克宣布推出 Nasdaq Financial Framework，向全球超过 100 家市场运营者提供区块链服务。该框架协议旨在为纳斯达克金融基础设施服务的用户提供端到端的解决方案。该协议允许交易所、经纪商、清算机构和托

管机构在统一的平台上和纳斯达克协作，区块链服务是核心服务之一。

2017年初，纳斯达克创建 Nasdaq Voting—一个基于区块链技术的投票和代理委托应用。所有成员都可以通过网络前端获得会议或投票活动的必要信息，提高年度会议和股东投票流程的效率，解决目前多数据源的访问和安全问题，让公司投资者，市场基础设施和托管机构在普通会议投票中降低成本和复杂性。

2017年5月纳斯达克宣布与花旗集团合作推出基于分布式账册的一体化自动交易结算支付解决方案，支持私募市场进行跨境多方交易和实时结算，集成了由纳斯达克提供的私募市场平台、

由 Chain 提供的区块链联通基础设施和由花旗提供的跨境支付平台。该方案通过覆盖更大范围的客户群体和流水线式高效交易处理，有利于提高纳斯达克私募市场的流动性。通过该系统，由纳斯达克私募市场产生的一笔交易，将在系统中生成一笔记录，并且通过二级账目的方式通过区块链为手段，将支付信息传递到花旗的处理跨境跨币种的外汇直通处理平台 CitiConnect，最后再传递花旗的跨境多币种支付平台 WorkLink，最后完成最终的资金交割。

2017年9月，纳斯达克与北欧投资银行合作作为共同基金份额的发行和结算开发基于区块链技术的原型产品。该解决方案降低了记账和核对

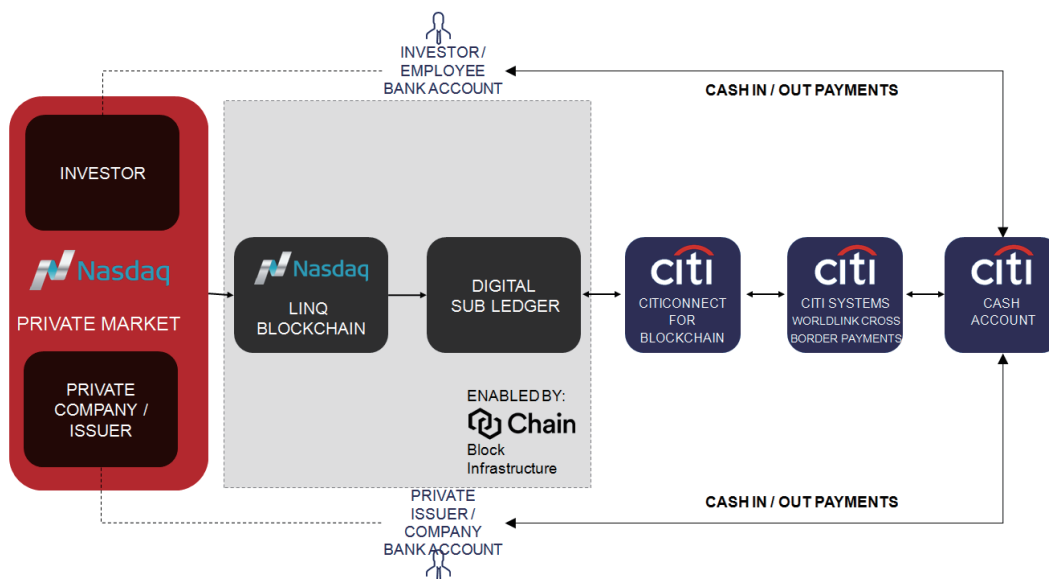


图3 纳斯达克与花旗集团合作的交易结算一体化解决方案

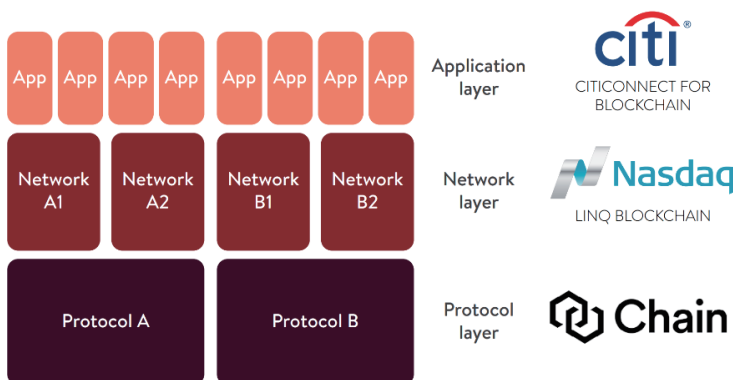


图4 纳斯达克与花旗方案的系统层次结构图

方面的复杂性，为所有各方提供了相同的不可篡改记录，增强了订单管理，结算和支付能力，提高了处理的透明度和效率。

2018年6月，纳斯达克证券交易所成功测试基于区块链的提供全天候的证券抵押概念原型解决方案，新的区块链平台由纳斯达克，ABN AMRO Clearing, Euro CCP和EuroClear联合开发，旨在解决中央交易对手方机构在证券交易所交易时间结束后进行保证金追踪时面临的困难。

2018年7月，纳斯达克交易所与印度国家证券交易所（NSE）签署技术和战略合作协议，NSE将引入纳斯达克的结算和清算技术。NSE将使用纳斯达克金融框架的“交易后”技术，同时预计将引入区块链技术应用用于股票交易。

2018年8月，纳斯达克同新加坡央行合作开发数字资产方案；2018年12月，其宣布预计于2019年上半年启动期货，并投资加密货币交易所ErisX。

2.2 伦敦证券交易所

2015年11月，伦敦证券交易所、伦敦清算所、法国兴业银行、芝加哥商品交易所、瑞银集团以及欧洲清算中心联合成立跨行业组织，探索区块链技术如何改变证券交易的清算和结算方式，该组织暂命名为“交易后分布式总账工作组”，主要探索将区块链技术应用用于交易后流程。

2017年7月，隶属于伦敦证券交易所集团（LSEG）的意大利证券交易所和IBM共同宣布，他们正在构建一个区块链解决方案，以助力欧洲中小企业（SME）的证券发行过程实现数字化。新系统旨在简化股权信息的跟踪和管理，创建一个包含所有股东交易记录的分布式共享账册，从而帮助发掘新的交易和投资机会。

2018年7月，伦敦证券交易所开始测试一个用于发行证券类代币的去中心化平台，由区块链专家Nivaura一起开发，并且致力于发展成为去中心化的交易所，可以让公司更高效，更简化

地募集资金。

2.3 德交所

德交所自2015年初就开始探索自己的区块链应用，是超级账本项目的成员之一。德国央行联合德交所共同开发了区块链原型产品，该原型以超级账本项目的代码为基础，用于转移电子证券和数字货币，还具有债券支付和到期证券的赎回功能。

2018年3月，德意志交易所宣布与HQLA^x合作创立一个以区块链为基础的证券借贷解决方案。根据合作协议，这两个机构将使用R3的Corda区块链平台，建立一个完全整合的从前台到后台运营模型，令高质量流动资产的抵押管理更加有效。德意志交易所指出，此举是为了解决全球证券系统过于分散，运营成本高，结算流动性差的问题。HQLA^x的首席执行官介绍：“我们的目标是汇集目前全球各地不同托管账户中的各种抵押品的流动性。”实际上，HQLA^x和R3已经与包括瑞士信贷银行和荷兰ING集团在内的银行巨头合作开发了一套证券贷款解决方案。经过概念验证后，上述银行已通过区块链平台完成了一笔价值3000万美元的证券交易。2019年12月，瑞士信贷和瑞银在德交所的HQLA^x平台上执行了首次实时交易。作为交易的一部分，瑞银和德国商业银行之间，交换了分别托管于欧洲商业银行和明讯银行的一篮子国债和公司债，过程中无需在托管方之间进行实物交换，相关所有权的变更都被记录于R3区块链的数字资产登记账册。

2018年5月，德意志交易所计划投资2.7亿欧元用于区块链、人工智能、大数据分析以及云计算相关领域。德意志交易所集团于2018年8月底成立“DLT、加密资产和新市场结构”的全新部门，以充分发挥区块链技术的潜力。

2.4 多伦多证券交易所

2017年5月，多伦多证券交易所（TMX）宣

布开发基于区块链的原型产品，为其全资子公司 Natural GasExchange (NGX) 提供全新服务。NGX 和数字基础设施开发企业 Nuco 合作使用分布式账本技术开发新系统来优化在美国天然气交付地区的天然气跟踪行为。该原型产品主要提供该地区天然气运输的透明信息，让参与者更准确地报告位置，从而优化 NGX 天然气结算流程。

2017 年 7 月，加拿大多伦多证券交易所运营商 TMX 集团宣布开发了基于区块链的股东电子投票原型，埃森哲提供相关咨询服务。利用区块链可以大幅度提高年度股东大会投票的效率、准确性以及股东在企业管理中的参与度，股东不需要亲自到场即可投票。

2018 年 5 月，加拿大银行、TMX 集团，以及非营利组织 Payments Canada 宣布该实时证券结算项目“Project Jasper”试运行正式完成。据悉，他们将这个最初用于比特币等加密货币的底层技术部署在了实时证券结算中，实现了现金和资产的“标记化”，继而可用于在支付网络内的实时交换。2018 年 6 月，多伦多证券交易所推出了新型区块链交易所交易基金（ETF）。

2018 年 10 月，Project Jasper 项目完成了 3 阶段的测试。该项目成员包括加拿大银行，Payments Canada，TMX 集团，以及区块链企业 R3，还有专业服务公司埃森哲，并且针对第三阶段测试发布了一份报告，认为经过一年的测试后，区块链被证明在证券清算和结算过程中是可行的。

2.5 东京证券交易所

东交所母公司日本交易所集团早在 2015 年下半年就开始研究区块链对于资本市场基础设施影响的内部研究，并得出这种创新技术将给行业带来巨大变革的结论。2016 年 2 月，该集团宣布于 3 月与 IBM 日本利用区块链技术测试其在清算、结算和其他运营领域的应用，成为 Blockchain-as-a-Service (BaaS) 的用户。

2016 年 12 月，日本交易所集团建立一个区块链联盟以在资本市场基础设施中进行区块链实验。参与联盟的包括东京证券交易所、大阪证券交易所和日本证券结算公司。三家公司牵头整个组建过程，并与 IBM 公司合作进行超级账本技术支持的概念验证测试，并试行开源分布式分类平台。

联盟运营目标包括寻求更多的其他利益相关者加入开发过程，比如各种金融机构和区块链的工程师和开发商。该集团表示：“我们会去吸纳广泛的金融机构参与者与行业专家，联盟成员将执行概念验证测试并探讨分布式账本技术在资本市场基础设施中的应用，包括技术方面和运营方面。”

此外，联盟的测试环境将利用概念验证和 IT 供应商开发的模型。金融科技和区块链开发商以及一些金融机构、交易所的共识是“验证各种应用能够帮我们强化技术和服务”。该联盟促进开发商和应用区块链技术的银行间信息共享，这种共享可以通过分布式账本技术（DLT）的工程师技术培训和金融机构的运营链培训来体现。区块链联盟的概念验证环境测试于 2017 年春季展开。

2017 年 9 月，日本交易所集团发表了一篇报告《分布式账本技术在资本市场应用探索趋势》，讨论了区块链在资本市场基础设施中的应用。其中披露 2017 年 3 月，东京证券交易所发起一项计划，广泛接纳日本金融机构开展行业合作实施区块链 PoC（概念验证）。到 9 月，有 33 家参与的金融机构可以基于分布式账本技术在联盟内部的网站上进行讨论，彼此共享信息。同时，报告表示东京证券交易所对于基于分布式账本技术的日本股票市场的交易后流程以及 KYC 流程十分关注。

2018 年 1 月 JPX 发布名为“分布式账册的交易撮合处理应用研究”报告，其中讨论了将分布式账册技术应用于机构间交易撮合的可行性。

其中，东交所机构间交易撮合传统处理方式如图（a）所示，其中卖方机构向买方机构发送订单执行的情况（交易确认），包括相应的佣金。作为响应，买方输入在账户（二级账户）中分配的结果，并将结果返回给卖方。卖方根据分配结果对佣金分仓，并将该结果返回给买方。如果双方的数据匹配无误，则撮合成功。之后，买卖双方使用日本中央托管结算（JASDEC PSMS）提供的交易撮合规程，提交（达成交易的）合约数据。一旦受托银行确认交易结果，会自动产生双向结算数据。双方的撮合匹配是按照常行结算 SSI (standing settlement instruction) 级别。作为结果，一个 DVP 结算自动生成并随之进行系统化结算处理。图（b）是使用 DLT 重构此过程的处理方式。买卖双方事前均就佣金计算表和计算逻辑达成共识，形成 DLT 上的一系列智能合约，佣金可以被自动计算。买卖双方都无需自行计算佣金，发生错配的几率大大降低。SSI 以最终分配账户为粒度执行，使得受托银行可按照每 SSI 的方式访问 DLT 上的合约数据，买卖双方的信息交换效率更高。此外，当单个基金与多卖方交易时，受托银行可以按照单个基金为操作粒度整合所有交易的信息。如果 DLT 的合约数据被认为正确，与日本中央托管结算的集成链接，将使买方与中央托管结算之间传输操作指令数据的必要

性不复存在。

2018 年 9 月，东京证券交易所上市了一家使用加密数字货币进行房地产交易的公司 Ruden Holdings，该公司正在测试利用比特币，并在不动产销售中使用智能合约。

2018 年 10 月，东京证券交易所母公司日本交易所集团网站发布报告《在 KYC 流程中使用区块链的概念测试》。

2.6 澳大利亚证券交易所

2016 年初，澳大利亚证券交易所投资 1490 万美元用于收购美国区块链初创企业 Digital Asset Holdings(DAH) 5% 的股份。

澳大利亚证券交易所自 2016 年以来提出基于分布式账册技术开发一个新系统，用于替代澳交所现行的 CHESSE 清算系统，该系统计划于 2021 年上线。Chess 系统构建于 1994 年，主要功能是为市场参与者提供清算、结算和资产登记服务。Chess 系统使用 COBOL 语言开发，使用私有的消息格式。澳交所希望新系统将提高证券结算的效率，降低系统运营成本和投资者的交易成本，并因此能提高市场参与率和成交量。

新系统可让市场参与者参与运行 DLT 节点，并使其成为 ASX 真实数据库的一个部分，而不是让多个数据库同时冗余存在。其中，DLT 使用

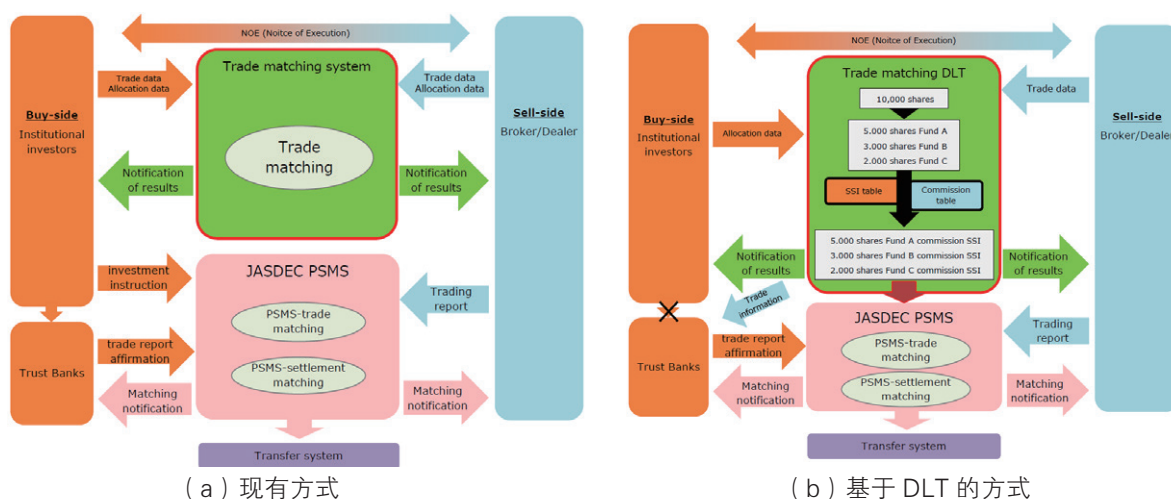
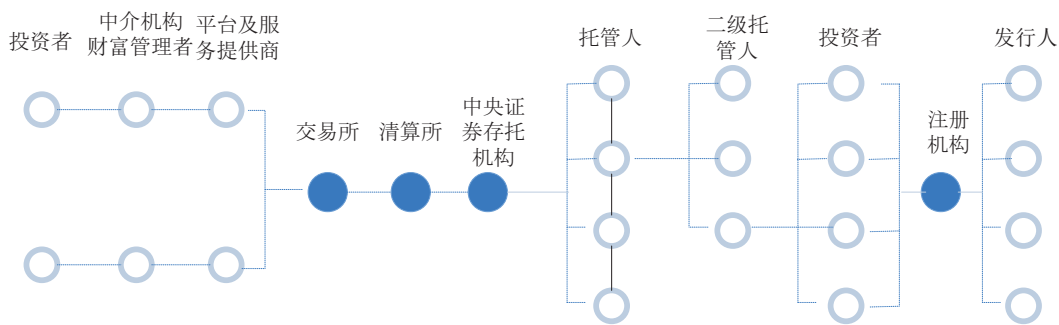


图 5 日本交易所集团基于 DLT 的机构间交易平台方案



澳大利亚证券交易所基于消息的分布式交易后处理体系

图 6 ASX 基于区块链的交易清结算示意图

的追加式交易数据库包括两个关键机制：一是共享的复制账本，二是分布式数据库同步机制。每一笔交易都会引用前一笔交易的持仓情况且被交易者数字签名。

ASX 将实施非完全公开的 DLT 解决方案，市场参与者使用需取得授权，并有逆转和纠错的功能。为确保隐私和安全，整个系统建立在 ASX 的私有的网络上。私人合约信息将被隔离而不会与所有参与者共享。新系统将使用国际通行的 ISO20022 消息格式，并采用开源的智能合约语言—数字资产建模语言（DAML）开发。

2.7 香港交易所

港交所近两年不断加速区块链等金融科技应

用布局。港交所《战略规划 2019-2021》包括立足中国、联通全球和拥抱科技三大主题，其中多次提及区块链技术应用，“灵活善用新科技，包括人工智能、区块链、云端计算基础设施和数据分析，优化营运程序，加强与客户的联系，并调动价值链各环节，提升成本效益”。

2018 年 3 月，香港证券交易所总裁李小加表示，港交所目前正与澳大利亚证券交易所 ASX 合作，以借鉴其在区块链系统中结算交易的经验，计划将利用区块链平台以低于传统结算方式的成本结算股票借贷和场外交易。

港交所在同年发布的报告“金融科技的运用和监管框架”中，对交易所行业运用区块链进行分析，重点讨论了区块链在证券交易结算、私募

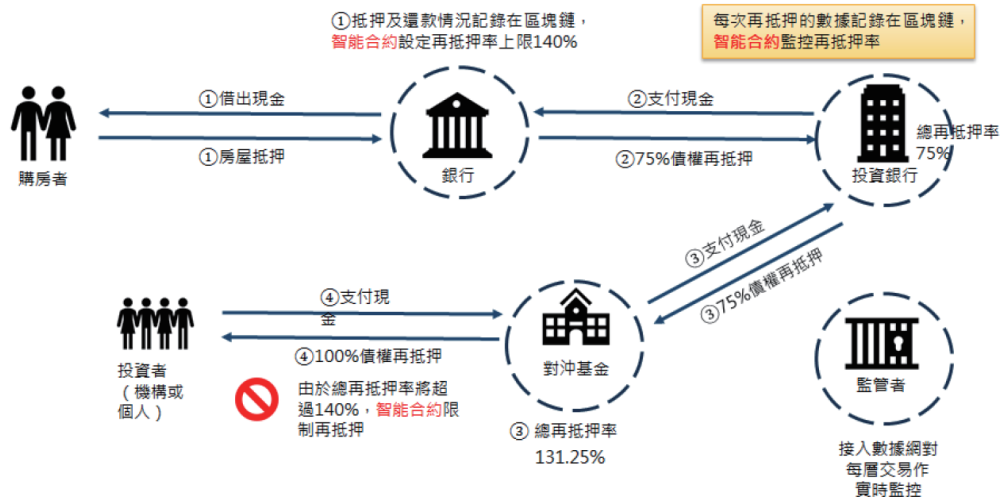


图 7 港交所基于区块链的资产再抵押业务框架

股权和资产再抵押等三大业务上的应用。如图 7 所示，以该报告对区块链应用于资产再抵押业务的考虑为例，其针对的问题是传统证券化资产抵押债券的数量众多、层层包装，对底层资产的所有权追踪困难、信用评估困难，加剧了再抵押交易对手风险和资产评估的不确定性。此外，基于资产杠杆率的约束难以实现，相应的风控措施和监管措施难以实现，不利于系统性风险的防控。利用区块链的可追踪性，可以解决再抵押业务的这些痛点问题。通过运用区块链的可追踪性，提高再抵押过程中的透明度。投资者可以方便地查看底层资产的担保价值、风险评级、所有权记录，方便投资决策。当资产池中债券发生违约或者风险事件，投资者可以及时获知并重新评估。第二，便于自动化监管。监管机构可以更清晰维护不可篡改的历史交易和再抵押记录。智能合约可以确保资产再抵押不会超过监管警戒线，最大程度满足穿透式监管要求，降低监管成本。第三，监管机构通过区块链技术进行合规约束，提升交易透明度，降低违约事件负面影响，提供金融稳定性。第四，智能合约使得机构和个人投资者对底层资产的尽职调查成本降低、效率提高。

2.8 新加坡交易所

新加坡交易所（SGX）于 2016 年参与了由

新加坡金融管理局和新加坡银行业协会共同发起，由交易所、金融机构、学界和科技业共同协作开展的金融科技项目 -Ubin。该项目作为新加坡央行数字货币应用的一个重要试点，旨在探索利用分布式账册技术变革证券和支付的清结算处理基础设施。

新加坡央行的数字货币采用基于分布式账册技术的代币化新币的实现方案，其发展分为 6 个阶段：1) 代币化新币。推出可用于跨行支付的数字货币；参与方银行将现钞存放在央行的托管账户，央行在分布式账册上为参与方银行生成等值的数字货币，相应银行一旦收到央行代币转账，即可对其自由支付或转账，该阶段已于 2016 年底完成；2) 实现基于分布式账册技术的实时跨行转账，该阶段验证也已完成；3) 实现基于分布式账册的 DVP，进行跨不同区块链平台的互联互通，该阶段自 2018 年起至今；4) 实现跨境 DVP；5) 对运营模式定位，包括评估分布式账册对于现行监管框架和业务处理流程的影响；6) 实现跨境支付和证券结算。

新加坡交易所所在该项目中，主要负责利用分布式账册技术开发面向代币化资产的 DVP 业务，以提升 DVP 的跨账册互操作性和处理效率，使全结算周期可从 T+3 缩短至 T+2 甚至更短。如图所示，为提供跨现货账册和证券账册的交互，

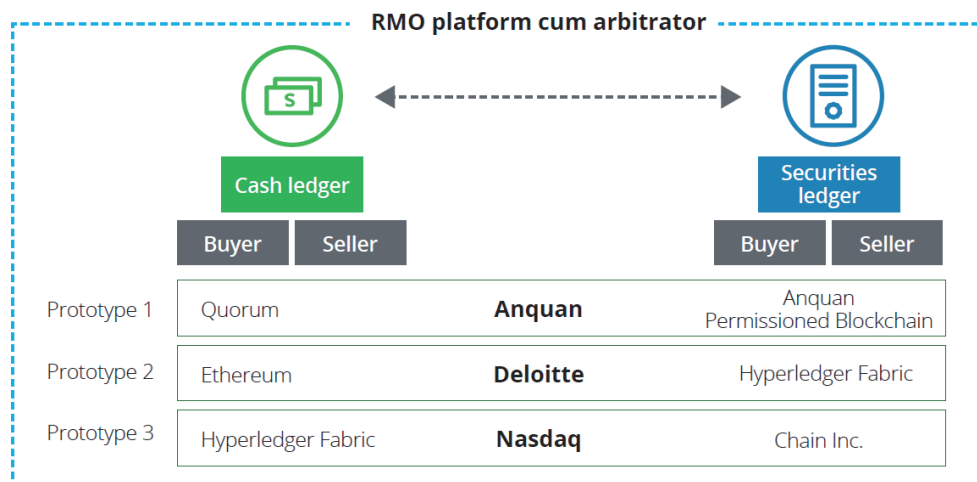


图 8 新交所基于分布式账册的 DVP 系统层次图

该系统与纳斯达克相关系统类似，也采用了分层的体系结构，分别由安泉资本、德勤和纳斯达克等不同的技术供应商实现不同层级的原型交互系统。各层的原型交互系统由供应商的商用系统构建，根据采用的具体技术实现向上层提供不同性能和功能的互联互通。其中纳斯达克和德勤主要负责实现基于集中式订单簿撮合的工作流，安泉资本主要负责实现基于 OTC 的工作流。基于智能合约的 DVP 确保了流程中权力和责任的一致性和完整性，减少了市场整体的信任成本。交易所作为一个 RMO（认证市场运营人）和仲裁人，负责监察和维护市场正常功能。现货账册是由央行维护的数字化新币资产账册，证券账册是由交易所维护的代币化证券资产（如国债、CDR）账册。

2019 年新加坡交易所进一步与亚马逊公司合作，共同基于区块链开发产品、进行金融服务产品推荐。根据新交所的规划，将在 2020 年正式对外推出基于分布式账册的数字资产。

三、创新案例：基于区块链的场外交易结算服务平台

传统的场外交易缺乏公开性、透明性以及有效的风险管理，譬如过去曾屡见不鲜的“萝卜章”事件，充分暴露了传统场外交易模式所蕴含的诸多局限。金融危机之后，场外交易场内化已经成为一种趋势，包括产品的场内化、交易的场内化和风险管理的场内化。场外交易的核心问题在于信用风险管理机制不足，对交易对手方的信用情况缺乏有效的跟踪与评价手段。将区块链应用于场外交易系统可以具有以下优势：一是可以保证场外交易的真实性、不被篡改，便于看穿和追踪，并降低传统 CCP 登记、信用担保、强制执行的成本，有效控制风险；二是通过智能合约实现实时的自动信任建立、交易、清算和结算，简化场外发行和交易流程，提升交易效率；三是利用区块链构建监管节点，可以针对特定主体、产品、

交易者和规模等，设定定制的监管方式和手段，使得监管更加有效和智能。

“百川”平台是将区块链技术应用于场外交易及结算方面的一个概念验证原型系统，采用私有分布式账册方案。主要提供智能合约、托管清算和智能监管功能，由上交所资本市场研究所和技术公司合作完成。相较于“产品标准化”的场外产品场内化思路，“百川”利用智能合约实现的“清算智能化”覆盖大多数场外交易场景，降低合约执行成本，简化交易后业务环节。当场外交易智能合约在平台中成功执行了权益的转移操作之后，平台账册的记录随之改变。随后，平台可依据相关登记结算机构的业务规则，代理市场参与者完成基础资产权益的转移。

“百川”平台使监管的实施上具有以下优势。一是“看得到”的穿透式监管。若所有的场外交易均在平台中执行，则监管者可以轻易获取合约内容、合约关联方等信息，实现穿透式监管。二是“查得完”的智能分析功能。相较于文字合约，智能合约的程序代码更容易被人工智能等技术获取和分析。平台可以提供智能化的监管工具，实现合约内容理解、合约方关系判定、关联交易追踪和市场风险压力测试。三是“管得住”的侵入式监管。场外交易在特定平台内运行，可以给监管层提供深入业务运作一线的监管机会。当发生风险事件或发现违规行为时，监管机构甚至可以直接中止相关交易的执行。

为验证智能合约的灵活性和可用性，“百川”平台中试验了五组不同复杂程度的示例智能合约。它们分别是普通欧式期权、多边参与期权、关联交易合约组、产品嵌套合约组和辅助交易合约组。经验证，“百川”平台可以提供自动和半自动的合约运行方式，灵活地支持场外合约设计。

“百川”平台的技术实现主要基于 HyperLedger 面向商用的区块链智能合约开发框架 fabric。此外，“百川”平台还引入了 Echarts、Nodejs 等数据可视化和 web 工具。

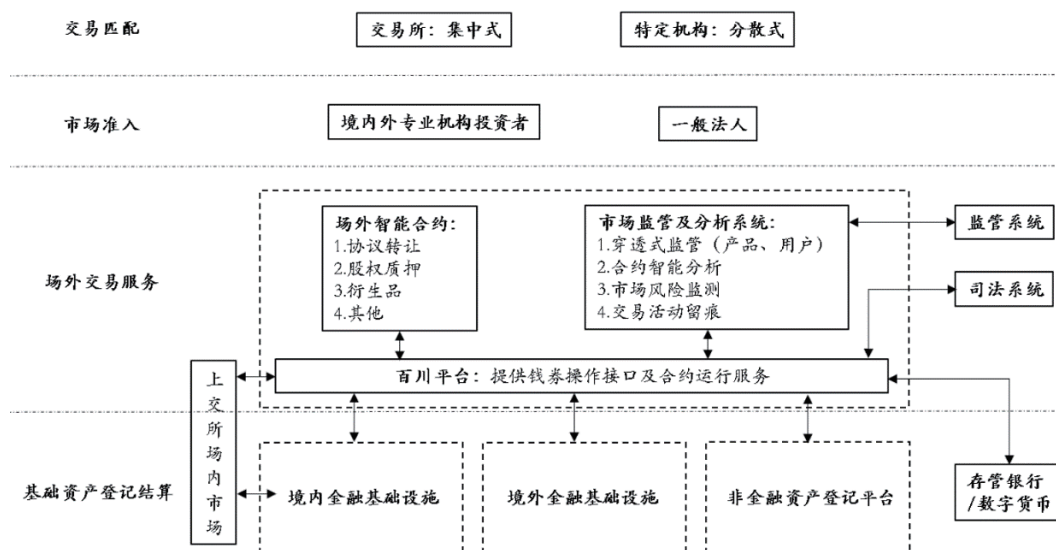


图9 “百川”平台系统架构图

四、结论及建议

一是充分重视区块链技术价值，明确将其作为行业金融科技重要发展方向

从行业境外经验上看，区块链技术目前主要应用于清算结算环节，并逐渐向场外市场和交易环节延伸。总体上，区块链的实际落地还较不广泛、影响还较小。但随着相关技术成熟度的快速提高，以及与传统证券业务的快速融合，不排除未来5到10年会通过对基础设施的变革，从而对传统证券业务产生巨大影响，而且在互联网时代可能由此改变全球证券市场格局。建议提高区块链在行业战略发展中的定位，将其作为金融科技的一个重要发展方向，纳入行业整体发展战略。同时，应加大投入，积极开展区块链技术的规划、研究和开发，对外开展积极合作，对内增强技术储备。

二是积极探索区块链在各业务领域的应用，鼓励可行性较高者以概念验证项目方式先行先试

如前所述，区块链可以对证券行业的多个业务领域和环节产生影响，但在重要性、紧迫性和可行性方面又各有不同。从境外交易所经验来看，大多都是以可行性较高的方面出发、痛点问题出

发，从概念验证项目的阶段入手，渐进但又快速地推进，这种类似敏捷开发的推进做法，大大降低了项目的试错成本和时间成本，同时为后续进入正式开发打下基础。从另一个角度，区块链技术本身也具有一定的优劣性，不是万能灵药。通过概念验证项目可深入理解区块链在各类应用中的优势和局限，有利于深入、科学和系统地评估其在应用于具体问题的可行性和有效性，以及运营成本效益等重要问题。

建议搭建行业区块链技术研发测试平台，对重要性、紧迫性较高且具可行性的方案，优先启动概念验证项目，鼓励各方积极承接课题、开展研究，共享交流成果，引领行业在区块链上共同进步。

三是积极开展对外交流与合作，提高国际话语权

建议行业机构积极参与区块链技术的业界交流，尤其是以各种方式积极参与全球交易所分布式账册变革进程，及时了解动态，并在区块链相关监管政策、业务发展和行业规范等国际事务中主动积极发声。加强与相关境内外技术组织、联盟、标准化组织和监管机构的交流互动，在区块链应用与发展、标准和相关知识产权事宜中掌控更多的话语权，为未来规模化应用以及跨境跨领域发展铺路架桥。

券商证券交易系统质量评估 框架

何志东 / 华锐分布式技术实验室

朱立 / 上交所技术公司

邓博 / 华锐分布式技术实验室

随着中国证券交易市场的快速发展，自然人投资者的增长给券商交易系统带来了系统容量上的压力，机构投资者的增长给券商交易系统提出了低时延等个性化的需求。交易系统的性能、容量、扩展性、对新业务的支持、可维护性等在当前业务场景下都面临着巨大的挑战。为了应对以上市场需求，目前中国证券行业正步入新一代券商交易系统选型建设的阶段，如何科学、客观地评估度量交易系统的质量显得尤为重要。本文的旨在探讨如何建立券商交易系统质量评估框架，以业界常用的五种质量模型 McCall、Boehm、FURPS+、Dromey、ISO 为基础，结合交易系统自身的特殊要求，根据行业的场景定义提出了细致可执行的度量方案。



1 引言

中国交易市场一直在快速发展,根据《上海证券交易所统计年鉴 2019 卷》和《深圳证券

交易所市场统计年鉴 2018》的统计,近十年每年新开户的投资者数量稳定在每年 250 万 -1000 万之间,2016 年 -2018 年新开户自然人投资者分别是 3237.6 万、2499.3 万及 1943.2 万,机构投资

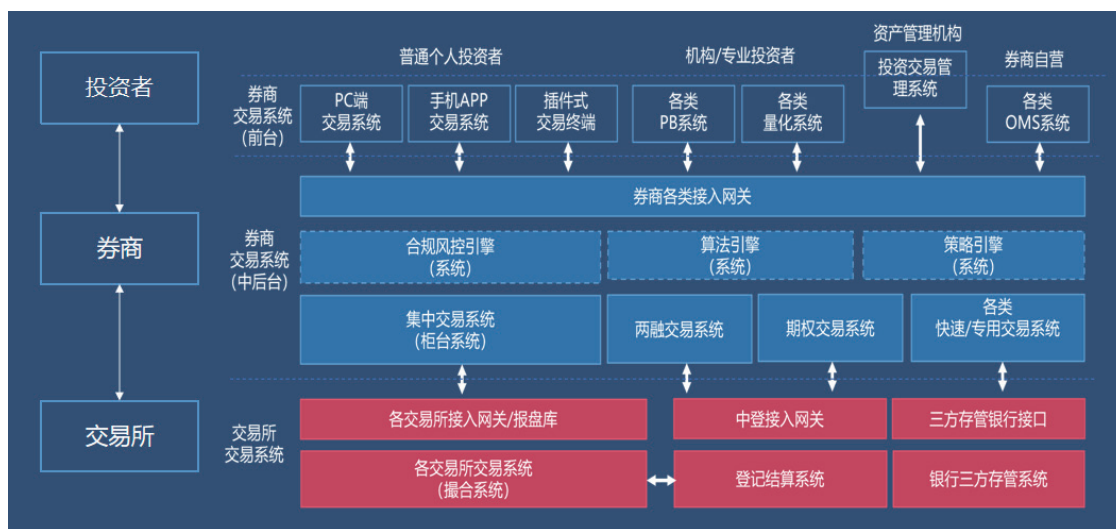


图 1 中国证券市场交易系统概览图

者在 2016 年 -2018 年分别新增 6 万、6 万及 4.5 万。截至 2019 年，沪市共有 2.14 亿投资者，其中机构投资者 64 万；深市共有 2.76 亿投资者，其中机构投资者 64.06 万。自然人投资者在某些年份井喷式地增长，直接给证券集中交易系统带来的是系统容量上的压力；机构投资者数量上的增长，则给交易系统提出更多个性化的需求，这两年最重要的趋势是部分机构投资者向交易系统提出了低时延的要求。

中国证券市场交易系统从 1992 年至今，经历多个阶段的发展，目前证券经纪业务交易核心系统习惯地叫做柜台系统或者集中交易系统，全行业目前正在使用的集中交易系统是从 2004 年开始替换的，到目前已经运行十五年有余，其系统容量、扩展性、对新业务的支持、可维护性等在当前业务场景下都面临着巨大的挑战；尤其对于机构投资者提出的性能需求，应对措施更是捉襟见肘。为了应对以上市场需求，目前中国证券行业正步入新一代集中交易系统选型建设的阶段，而为了应对机构投资者的需求，全市场更是掀起了证券极速交易系统测试选型的热潮。

在这个中国证券行业发展正在革新的历史阶段，如何科学地、客观地评价证券交易系统就显得尤为重要，本文的目标则是探讨如何建立证券

交易系统质量评估框架。

2 交易系统

2.1 交易系统范围

本文的评估框架面向的对象是证券交易系统，不同视角的证券从业人员对证券交易系统的定义是有差异的，广义的证券交易系统包含了所有围绕着证券交易构建的系统的集合，狭义的证券交易系统则特指证券交易关键路径上的各个系统的集合。

图 1 是中国证券市场交易系统概览图，图中基本囊括了订单从投资者到交易所交易系统的全路径的各种交易系统的类型。

从系统的功能以及归属关系来分，交易所的交易系统归属于交易所，为市场提供交易撮合服务；券商的交易系统归属于券商，为市场提供投资管理以及交易通道服务；进一步地，如上图所示，券商的交易系统又分为前台和中后台系统。在确定评估对象之时，笔者发现范围定义得太大会导致整个框架泛泛而无可着力点，最终导致需要进一步定义具体的执行方案时无可行性，因此本文评估的交易系统特别圈定的范围为：券商交易系统（中后台）中的集中交易系统、两融、期权以及各类快速交易系统和专用交易系统。后



图 2 交易系统组件架构缩略图

续章节中除非有特别说明，否则“证券交易系统”专门指代以上圈定范围内的系统。

2.2 交易系统架构

无论是上一代的集中交易系统，还是目前行业正在使用的机构交易系统，亦或是市场上正在推出的新一代分布式交易系统，这些交易系统从功能模块的划分上基本上还是保持一致的，图 2 为华锐分布式核心交易平台（ATP，Archforce Trading Platform）的组件架构缩略图。

蓝色框的部分是交易系统的核心组件，网关（GW）负责订单的接收及路由、交易引擎（TE）负责订单的检查以及订单生成，报盘服务（ORS）则对接交易所的交易网关，将订单报送给交易所。

不同开发商的交易系统系统架构可能会有差别，可能是集中数据库架构的，或者是总线型架构的，又或者是微服务架构的，最终从功能上，都会有以上功能的实现；从进程上的划分上，以上几个功能模块可以分开从而实现架构的松耦合，也可以为了追求极致的性能将模块合并为一个进程，减少进程间的交互，但是以上模块的功能实现，都会在代码实现上得到体现。

本质量评估框架也将基于以上交易系统架构展开讨论。

3 质量度量模型与测试标准描述

3.1 软件质量模型

质量是产品的生命，它直接影响着产品的使用和维护。我们不能直接来测量软件的质量，我们只能通过软件的一些具体的质量特性来测量。给质量特性规定标准，于是就建立起来了具体的质量度量标准。质量度量的标准也可以叫做质量模型，它是质量管理过程中的核心部分。

如今业界常用的质量模型有五种，分别是：McCall、Boehm、FURPS+、Dromey 和 ISO。

我们稍微介绍应用相对广泛的 ISO 模型：

ISO 是目前比较完善和权威的软件质量模型，1993 年制定第一版，目前使用的是 2011 年重新修订版。当前版本的模型由 8 大基本特征和 5 个软件使用质量特征组成。每一个特征又可以分解成为可以度量的多个子特征。

8 大基本特征如图 3，基本可以涵盖软件的内外部的质量考察点，也包含了软件的功能性和非功能性质量检查点。

ISO 定义了 5 个软件使用质量特征，根据用户使用软件的结果而不是软件自身属性来进行度量，使用质量是面向用户的内部和外部质量的组合效果：



图 3 ISO 25010 软件内、外部质量

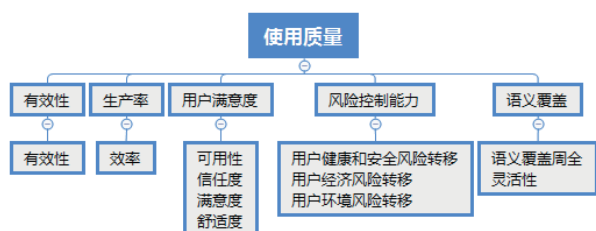


图 4 ISO 25010 使用质量

3.2 证券期货行业软件测试规范^[8]

2019年9月30日，中国证券监督管理委员会正式发布了《证券期货行业软件测试规范》，规范针对证券行业的特征提出了行业软件产品的测试内容和测试类型。其中，测试类型按照功能测试和非功能测试进行区分，包含了功能测试、性能测试、可靠性测试、安全测试、可移植性测试、可维护性测试等，并针对每个测试类型做了

详细的测试内容和关注点定义，例如，对性能测试类型，明确要求关注负载测试、压力测试、容量测试与业务响应时间测试几个方面。其关注点与质量模型内容互相对应，形成一套对本行业软件的质量评估测试方法。

众多的质量模型都有自己的优势和劣势，测试规范也定义了本行业的适应的测试方法和关注内容，通过对模型与规范的结合，我们定义了一个针对券商证券交易系统的质量评估框架。

4 交易系统质量评估框架

证券交易系统是软件系统中的一个极小的子集，它与其他软件系统既有共性，亦有自身的特殊要求，给交易系统建立质量评估框架，本质上在已

表 1 证券交易系统非功能性质量特征与描述

证券行业交易系统非功能指标		
性能	吞吐量	交易系统特定场景下能达到的业务并发能力
	时延	交易系统特定场景下的业务时延性能
	容量	交易系统能承载的最大业务数量
	资源占用	交易系统在特定场景条件下的资源使用情况
可靠性	成熟性、容错性	交易系统有对系统内外部错误的防御能力和错误消除能力
	可恢复性	系统在失效状态下，恢复至正常状态的能力
	稳定性	交易系统长时间运行的能力
可伸缩性	系统可伸缩性	交易系统支撑的业务容量的水平伸缩能力
可拓展性	功能可拓展性	交易系统易于变更或增添业务功能的能力
可移植性	周边系统对接	与周边系统对接的接口定义及对接能力
	可移植能力	硬件平台、操作系统类型及操作系统版本的兼容能力
	可安装性	系统有明确安装指引
可维护性	可监控	组件运行状态、事件、指标等的收集、分析、展示、处理
	集中运维	集中管理的能力
	时延分析	对系统时延性能的分析能力
	部署复杂度	易部署、可升级、可回退、有版本管理
	维护可操作性	日常运维流程、应急处理手段
易用性	易使用	易理解性、易学习性、易操作性、符合习惯
	易开发	二次开发难度低
安全性	安全可靠	保密性、完整性、不可抵赖性
	无高危漏洞	减少系统被利用和被攻击风险

表 2 吞吐量指标度量标准

吞吐量指标	度量单位	指标定义
系统吞吐量	笔订单/秒	全系统进行压力测试，系统每秒能处理的最大的订单数量，测试时间应该持续 30 分钟，系统时延稳定正常无波动，队列无消息积压；
交易引擎(TE)吞吐量	笔订单/秒	TE 每秒能处理的最大的订单数量，测试时间应该持续 30 分钟，交易引擎内部时延稳定正常无波动，队列无消息积压；
交易网关(GW)吞吐量	笔订单/秒	GW 每秒能处理的最大的订单数量，测试时间应该持续 30 分钟，交易网关内部时延稳定正常无波动，队列无消息积压；
	笔查询/秒	GW 每秒能处理的最大的查询数量，测试时间应该持续 30 分钟，交易网关内部时延稳定正常无波动，队列无消息积压；
报盘服务(ORS)吞吐量	笔订单/秒	ORS 每秒能处理的最大的订单数量，测试时间应该持续 30 分钟，报盘服务内部时延稳定正常无波动，队列无消息积压；
系统其他瓶颈环节的吞吐量	笔消息/秒	测试一些瓶颈环节，系统的吞吐能力，测试其每秒能处理的消息个数的能力

备注：以上性能指标应基于特定部署前提，单实例部署可建立度量基准，整系统测试则度量系统的能力。

有的质量度量模型下，将证券行业需求放大，并进一步根据行业的场景定义细致可执行的度量方案。

交易系统业务功能质量的保证是系统可用的基础，是目前各个开发商以及券商重点关注对象，大家也已探索出行之有效的质量保证体系，故本质量评估框架不再对证券交易系统的功能质量指标做定义，本框架着重关注证券交易系统非功能性指标。

根据上述标准质量模型，本文提出证券交易系统相关子特征和特征描述如表 1，本文尽可能描述所有相关指标，后续具体实施过程可以根据不同情况进行适当裁剪和指标判断。

4.1 性能

正常运行的情况下，系统的功能实现决定了交易系统能提供何种服务，系统的性能则决定了系统提供服务的质量、效果。在时间特性和资源

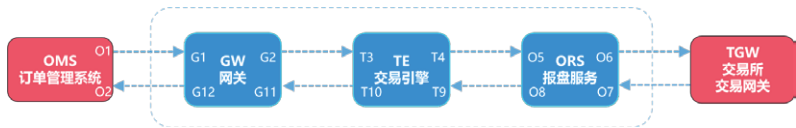
利用性上两个方面，证券交易系统着重关注以下几个子性能指标：吞吐量、时延、系统容量、资源占用。

4.1.1 吞吐量

吞吐量体现的是系统的负载，是系统的最大服务能力。交易系统是一个业务系统，处理的核心业务是订单生成递交给交易所，指标的定义应围绕订单业务场景进行。对此，应该定义以下几项指标（表 2）。

4.1.2 时延

系统速度快与慢是一种相对的且主观的形容方法，系统处理时间是则衡量速度的最好标准，它体现了业务响应时间指标。对于请求应答类型的服务，我们通常用响应时间（Respond Time）来度量系统的快慢；对于交易系统这种类通道型的服



备注：图中的 G1 表示订单进入 GW 打的时间点，G2 表示订单出 GW 打的时间点，其他时间点同理。

图 5 交易系统流程缩略打点图

表 3 时延指标度量标准

时延指标	度量单位	指标定义
系统上行穿透时延	微秒	订单从发出订单管理系统到离开报盘服务（对于沪市则是上海报盘库写库返回）之间的时间差：O6-O1
系统下行穿透时延	微秒	成交确认从进入报盘服务到到达订单管理系统之间的时间差：O2-O7
组件内部时延	微秒	订单进入组件到订单处理结束离开组件之间的时间差：G2-G1，T4-T3 等
系统回路时延	微秒	订单管理系统发出订单到收到对应订单的委托确认时间的的时间差：O2-O1
系统内部时延	微秒	(G12-G1)-(O7-O6)

上表时延指标虽然只定义了 5 个，但是在不同的速率下、不同的账户数量下、不同脉冲吞吐下、不同业务组合模式下等各种场景组合下，会衍生出很多有度量价值的指标，建议单一时延指标明确输出均值、中位数、90%、95% 等统计指标。

4.1.3 容量

系统能承载的最大业务数量，业务维度包括账户容量、委托容量、成交容量、证券数量等，技术维度则包括网关和 ORS 能对外提供的服务能力。

4.1.4 资源占用

交易系统在特定场景条件下的资源使用情

况，包括 CPU、内存、硬盘、带宽等，资源占用是一个动态变化的过程，随着交易时间向后增长，内存、硬盘占用都会相应地增长；不同的业务并发量下资源占用的情况也会有较大变化。

4.2 可靠性

4.2.1 成熟性、容错性

产品成熟性与容错性用于表示在面对内外部错误和故障时的错误消除能力和防御能力，它们都可以通过业务结果来展现。证券交易系统要求报送给交易所的订单不重不丢不乱，不重指的是同一笔交易订单不能重复递交给交易所，不丢指

表 4 容量指标度量标准

容量指标	度量单位	指标定义
系统账户容量	个	交易系统支持的最大账户数量
系统订单容量	笔	交易系统单个交易日能支持的最大订单数量
系统成交容量	笔	交易系统单个交易日能支持的最大成交数量
系统证券数量	个	交易系统能支持的最大证券数量
网关连接容量	个	单个交易网关能对外提供的最大连接数量
报盘服务交易单元容量	个	单个报盘服务能支持的最大交易单元数量
报盘服务连接容量	个	单个报盘服务能支持的与交易所报盘网关（报盘库）建立的最大连接数量
数据库容量	笔	数据库能支持的最大订单数量及成交数量

表 5 资源占用指标度量标准

资源占用指标	度量单位	指标定义
CPU 占用	百分比	交易系统运行时各个组件的 CPU 使用率
内存占用	字节	交易系统运行时各个组件的内存占用数量,尤其要关注内存占用与订单数量及运行时间的关系
硬盘占用	字节	交易系统运行时硬盘空间的占用增长与与订单数量及运行时间的关系
组件带宽占用	比特	交易系统运行时各个组件的输出与输入带宽占用,尤其要关注带宽占用与订单数量及运行时间的关系,以及带宽占用和直接通信对象个数的关系
系统带宽占用	比特	交易系统运行时交换机的带宽占用,尤其要关注带宽占用与订单数量及运行时间的关系

的是不能不将投资者的订单提交给交易所同时没有任何知会；不乱是为了保证交易的公平性，投资者先递交的订单应按序先递交给交易所。在面对系统内外部异常的时候，系统依然能有效运行并保证业务正确性。

通常发生重单、丢单或者乱序是系统中某些组件发生故障或者故障恢复时发生的，因此应重点检查故障或故障恢复时以上指标的情况。

4.2.2 可恢复性

中国证券市场对核心交易系统故障零容忍的监管规定，确定了容错能力是证券交易系统的核心能力指标，容错能力有两个行业公认的核心指标：RTO (Recovery Time Objective, 用于度量系统恢复提供服务所需要的时间) 和 RPO

(Recovery Point Objective, 用于度量发生故障到系统恢复数据丢失时长)。具体化到不同的交易场景，对交易系统的容错能力也是有阶梯式需求的，应根据不同的需求，也衍生出不同的部署方案，针对不同的部署方案，我们可以列举出相应的容错指标。

4.2.3 稳定性

4.3 可伸缩性

证券行业的发展都是周期性的，因此交易系统容量的需求也是周期性的，在功能保持不变的情况下，交易系统仅靠增加硬件设备就能支撑更大业务容量的能力非常重要。

4.3.1 系统可伸缩性 (Scalability)

表 6 可靠性可用性指标度量标准

可靠性指标	度量单位	指标定义
不重单	百分比	系统发生故障的情况下，交易系统重复递交订单的数量占比
不丢单	百分比	系统发生故障的情况下，交易系统丢失订单的数量占比
不乱序	百分比	系统发生故障的情况下，交易系统订单乱序的数量占比
可用性	百分比	交易系统在运行时间内的可用程度=正常提供服务时间/系统运行时间
容错性	个	通过冗余组件提供容错能力的情况下，能够最多容纳多少个冗余组件失效

表 7 可恢复性指标度量标准

部署方案	容错性指标	度量单位	指标定义
单数据中心 部署	RTO	微秒	不同故障场景下，故障恢复时间
	RPO	微秒	不同故障场景下，故障恢复丢失数据
	切换方式	无	不同故障场景下，自动切换还是人工切换
同城双中心 部署	RTO	微秒	不同故障场景下，故障恢复时间
	RPO	微秒	不同故障场景下，故障恢复丢失数据
	切换方式	无	不同故障场景下，自动切换还是人工切换
两地三中心 容灾部署	RTO	微秒	不同故障场景下，故障恢复时间
	RPO	微秒	不同故障场景下，故障恢复丢失数据
	切换方式	无	不同故障场景下，自动切换还是人工切换

表 8 稳定性指标度量标准

健壮性指标	度量单位	指标定义
长时间空转运行 指标	分钟	系统长时间空转不出异常的时间长度
长时间正常业务 量运行指标	分钟	系统长时间正常业务量运行不出异常的时间长度（未达到硬件容量）
异常数据处理指 标	无	系统正常或系统异常的情况下，交易系统丢失订单的数量

可伸缩性（Scalability）可以反应软件系统计算处理能力。通过很少的改动或者硬件设备的增加，实现整个系统处理能力近似线性的增长。

4.4 可延展性

证券行业的发展都是周期性的，新业务新功能的推出常常是周期性的，评估交易系统快速响

应业务需求完成所需业务变更的能力非常重要。

4.4.1 功能延展性（Extensibility）

4.5 可移植性

广义的交易系统是围绕证券交易构建的各种应用系统的集合，而交易的核心在于交易系统，这就涉及到交易系统的快速部署以及与周边系统

表 9 系统可扩展性度量标准

容量可扩展指标	度量单位	指标定义
吞吐能力可扩展	无	吞吐量能否横向水平扩展，对性能有何影响，扩容成本如何
账户容量可扩展	无	账户容量能否横向水平扩展，对性能有何影响，扩容成本如何
订单容量可扩展	无	订单容量能否横向水平扩展，对性能有何影响，扩容成本如何
成交容量可扩展	无	成交容量能否横向水平扩展，对性能有何影响，扩容成本如何
证券数量可扩展	无	证券数量能否横向水平扩展，对性能有何影响，扩容成本如何

表 10 功能可延展性度量标准

功能可扩展指标	度量单位	指标定义
新增业务功能	无	符合新增业务功能，对性能有何影响，新增功能成本如何
支持二次开发	无	系统是否提供二次开发框架以及接口

表 11 周边系统对接能力度量标准

周边系统对接指标	细化对接指标	度量单位	指标定义
网关对接能力	支持第三方交易客户端接口	无	已经支持的第三方交易客户端数量以及目前支持的第三方交易客户端
	自身的接入接口	无	系统自定义的接入接口，是否符合行业习惯
数据上下场对接	无		已经支持的周边系统的数据上下场接口
实时指令接口	无	无	已经实现的实时指令接口

的对接兼容问题，而这也是评估一个交易系统能否快速上线的重要考虑因素。

4.5.1 周边系统对接

与周边系统的对接能力，体现本系统的兼容性，也体现了系统的可安装和可替代特性。

4.5.2 可移植能力

可移植能力涵盖了应用软件与硬件平台、应用软件与操作系统、操作系统与硬件平台间的兼容性，以及软件系统能在所有特定的目标环境（硬件、软件、中间件、操作系统等）中正确运行的能力。

4.5.3 可安装能力

4.6 可维护性

交易系统的核心功能是后台的业务处理系统，包括上述提到的交易网关、交易引擎和报盘服务，但是系统是需要管理运维的，保证系统运行不发生故障，系统的可维护性非常重要。

4.7 易用性

4.8 安全性

交易系统涉及证券交易的核心业务，后台系统直连交易所，从技术上看系统运维及运营人

表 12 可移植能力度量标准

可移植指标	度量单位	指标定义
服务器兼容性	无	服务器的兼容列表（X86/ARM/小型机等），并在相应设备上能正常运行
网卡兼容性	无	网卡兼容列表，系统能正常运行并能兼容对应网卡
操作系统及版本兼容性	无	操作系统、版本兼容列表，系统能在相关操作系统正常运行

表 13 可安装能力度量标准

可安装指标	度量单位	指标定义
安装指引	无	系统发布有明确的安装指引说明书或指导手册

表 14 可维护性能力度量标准

可维护性指标	度量单位	指标定义
可监控	无	组件运行状态、事件、指标等的收集、分析、展示、处理能力； 提供监控接口；
集中运维	无	对交易系统集中运维、运营管理的能力； 提供运维管理接口；
时延分析	无	实时展示交易系统处理时延的能力； 提供时延数据接口；
部署复杂度	无	易部署、可升级、可回退、有版本管理
维护可操作性	无	有明确的日常运维流程、应急处理流程等
问题可跟踪	无	日志信息精确可跟踪、有问题排查手册
完备的工具集	无	能够提供的处理突发问题的工具集合
可测试	无	配套性能和功能测试工具

表 15 易用性能力度量标准

易用性指标	度量单位	指标定义
易使用	无	易理解性、易学习性、易操作性、符合习惯
易开发	无	二次开发门槛低，提供完整的高性能开发库及业务开发库

表 16 安全性度量标准

安全性指标	度量单位	指标定义
保密性	无	有完整的账户权限管理体系
不可抵赖性	无	交易数据全路径落地，原始数据以及状态信息完整
问责性	无	操作留痕、可审计
无高危漏洞	个	全系统经过专业扫描工具如 Nessus 扫描，无高危漏洞存在

员跨过了券商的各路安全控制，错误不合规的操作能够直接影响市场，因此交易系统的安全性指标非常重要，它是守护我们交易安全的最后一道关卡。

5 结语

证券交易系统是证券行业最重要的最核心的

系统，交易系统的能力决定了券商对投资者提供服务的能力，最终决定了券商的核心竞争力，如何科学、客观地评估度量交易系统的质量，在过去、现在及未来都是一个非常重要的课题。本文基于目前中国证券市场的发展阶段，提出了一个证券交易系统质量的评估框架，更多地是希望在行业内与大家探讨，形成标准规范的评估方案，推动行业系统质量往前发展。

作者简介：

何志东，hezhidong@archforce.com.cn，华锐分布式技术实验室主任，具有丰富的交易系统、消息总线和高可靠架构设计与开发经验，目前主管华锐分布式技术实验室，对分布式系统架构、高性能计算基础设施、低时延技术、系统性能调优等领域有深入研究。



I 信息资讯采撷 nformation

监管科技全球追踪

监管科技全球追踪

国际组织动态

七国集团发布全球稳定币影响力报告

10月18日，七国集团（G7）稳定币工作小组发布《全球稳定币影响力调研报告》。该报告指出，第一波加密货币发展并未给社会提供可靠和有吸引力的支付与价值存储方式。而如今，稳定币的出现可能会解决这些问题，提供更便捷、更廉价和更具包容性的全球支付体系。当然，稳定币产业发展可能也会给公共政策、行业监管、反洗钱和反恐怖主义融资等方面带来重重挑战和风险。报告认为，在没有充分解决这些风险和挑战之前，不应该发行任何全球化的稳定币。

国际清算银行发布了《监管科技时代》专题报告

10月17日，国际清算银行发布了《监管科技时代》（The supotech generations）的专题报告，对全球39个金融监管机构的监管科技举措进行了分析。报告指出，监管机构采用的辅助监管技术目前共经历了四个阶段：1）大量手动输入的数据管理工作流程，并且主要提供描述性分析；2）数字化并自动执行数据管道中的某些手动处理环节；3）涵盖大数据架构；4）涉及添加AI作为定义特征。而所谓“监管科技”（Supotech）主要涉及后两个阶段，尤其是第三阶段的数据收集解决方案（如数据报告、数据管理和虚拟助手），以及第四阶段的数据分析解决方案（如市场监督、不当行为分析、微观审慎监管和宏观审慎监管）。受访的39个金融监管机构中，几乎有一半都已

经设立了明确的“监管科技”（Supotech）战略或已经开始了相关研发工作。其中，有些是以大数据和AI为核心的路线图计划，有些则是将这两项技术作为大范围数字创新计划的一部分。而对于那些想要在投入大量资源之前先探索特定“监管科技”（Supotech）工具的机构而言，可以先通过创新实验室、加速器等制度进行尝试。

金融稳定委员会发布专题报告，探讨大型科技企业、云计算对金融稳定性的影响

金融稳定委员会（FSB）的职责之一是在国际层面协调国家金融机构与国际标准制定机构的工作，为促进金融稳定而制定有效监管政策。12月9日，FSB发布《金融领域的BigTech：市场发展和潜在的金融稳定性影响》和《云服务中的第三方依赖关系：对金融稳定性的影响》两份报告，重点关注大型科技企业（BigTech）和云计算、数据服务对金融稳定性的影响。在这两份报告中，FSB指出，大型科技企业也可能对金融稳定构成风险。某些风险与金融公司的风险相似，例如杠杆、期限转换和流动性不匹配以及其他运营风险。大型科技企业拥有大量的资源和用户数据访问渠道，能够通过网络效应迅速扩张。但是，未来少数企业可能因此在金融服务中占据垄断地位，影响金融科技的多样化发展。此外，云服务平台可能会带来全新的监管难题，尤其是在跨境服务的情况下，不同司法辖区的监管机构有时会难以评估金融服务是否符合所在辖区的法律要求。

欧美动态

欧盟：金融科技监管、创新与融资的 30 条建议

去年，欧盟委员会成立了一个专家小组，专门负责审查现有欧盟法律法规对金融科技行业的适用性。12 月 13 日，专家小组完成了审查调研工作，并发表了一篇名为《金融科技监管、创新与融资的 30 条建议》的专题报告。专家小组指出，金融科技是一个不断发展变化的产业，欧盟地区监管机构应该秉持多维、协作监管思路，努力保持高标准的消费者保护、市场完整性和金融体系稳定性。此外，专家小组还建议在加密资产、金融数据方面采取更加大胆的政策行动，调整现有法规的适用性，紧跟市场发展变化的速度。

英国金融行为监管局发布《气候变化与绿色金融：意见回复及未来行动》专题报告

2019 年 10 月 16 日，FCA 整理并发布了《气候变化与绿色金融：意见回复及未来行动》专题报告。本报告重点关注以下 5 个主题：1) 证券发行人与气候风险披露；2) 受监管公司与气候风险披露；3) 可持续性的通用指标和标准；4) 利益相关者的担忧、商业重点和增长障碍；5) 行业参与度。FCA 制定了一些近期的绿色金融发展行动计划，其中包括：1) 就全新监管规则展开咨询，改革证券发行人的气候风险要求，阐明现有信息披露义务；2) 最终确定规则变更，要求独立治理委员会（IGC）监督和报告公司的环境、社会与治理（ESG）和管理政策以及单

独的规则变更，增加长期资本投资机会；3) 针对财务报告委员会关于管理权的联合讨论文件发布反馈声明，并罗列解决有效管理权最主要障碍的措施；4) 鉴别出那些以环保和可持续发展为幌子的企业，阐明绿色发展理念，采取适当措施保证消费者免受误导；5) 为几个重要的合作计划做出贡献，包括气候金融风险论坛、公平有效市场审查工作组、政府主导的信息披露跨监管工作组以及欧盟委员会的可持续金融行动计划。

美国商品期货交易委员会将金融科技研究实验室升级成为独立办公室

2019 年 10 月 24 日，美国商品期货交易委员会（CFTC）宣布将金融科技研究机构 LabCFTC 升级为 CFTC 下属的独立办公室部门，并向机构主席 Heath Tarbert 直接汇报。2017 年，CFTC 成立了金融科技研究机构 LabCFTC，希望借此加强与金融科技行业交流。目前，该机构由 CIO 兼董事 Melissa Netram 领导，她此前曾在硅谷金融服务公司 Intuit 任职。同一天，LabCFTC 还发布了首份金融科技与人工智能应用的入门手册。这是一份面向市场参与者和消费者的公共教育工具，也是 LabCFTC 致力于向公众提供有关金融科技创新信息的一次尝试。

纽约金融科技创新实验室开放 2020 年项目申请

10 月 21 日，纽约金融科技创新实验室宣布开放 2020 年项目申请。该项目由埃森哲集团与

纽约 Partnership Fund 共同创立，为期 12 周，主要面向金融服务领域的创企。2020 年，项目将主要关注来自以下 5 个领域的创企，包括 1) 云端（云管理、迁移和安全性）；2) 数据（替代数据分析、数据管理、隐私与可视化工具）；3) 数字化；4) 参与度（客户获取、产品订制、预测分析等工具）；5) 安全性（网络、风险与身份管理）。参与项目申请的机构必须已经完成相关技术的测试版，并保证可以进行市场化测试。实验室将挑选业内资深的银行、财富管理专家与创企面对面交流，为其提供技术、市场等战略帮助。

三家美国监管机构加入全球监管沙盒计划

10 月 24 日，美国证券交易委员会（SEC）、货币监理署（OCC）和联邦存款保险公司（FDIC）联合宣布加入全球监管沙盒 GFIN。此前，美国商品期货交易委员会（CFTC）已经在 GFIN 创始初期加入了这一组织。GFIN 由英国金融行为监管局（FCA）牵头于 2018 年创建，其目标是成为全球性、跨境金融创新业务的监管沙盒。

亚马逊与美国证券交易监管机构达成一项重要的云服务合同

12 月 4 日，亚马逊宣布亚马逊云服务部门（AWS）正与美国金融业监管局（FINRA）的子公司 FINRA CAT 合作，监管美国证券市场的交易，而美国证券市场每天将吸收 22 家证券交易所和 1,500 家券商的 1,000 多亿次市场活动。FINRA CAT 首席技术官斯科特·唐纳森表示，这笔交易的意义不仅限于数据分析和监管。它最终将通过提高监管机构监管美国证券市场交易的能力来保护投资者。

美国五大金融监管机构发表联合声明，支持使用替代数据

12 月 4 日，美联储委员会、美国消费者金融保护局、美国联邦存款保险公司、美国国家信用社管理局和美国货币监理署就“替代数据”（Alternative Data，也称“另类数据”）使用发表联合声明。声明指出，上述机构了解到，银行和非银金融公司在信贷承销、欺诈检测、营销、定价、服务和账户管理中正在使用或正在考虑使用替代数据。本次声明重点关注在信贷承销过程中使用替代数据对消费者保护的影响，潜在的好处和风险。这五家机构表示，使用替代数据的确可以提高信用决策的速度和准确性，帮助企业评估目前在主流信用体系中无法获得信用的消费者的信用度，可以使消费者通过基于对还款能力的增强评估来获得更多产品和 / 或更优惠的定价 / 条款。这些创新反映了自动承保和信用评分模型的不断发展，具有降低信用成本和增加获得信用的潜力。与信用担保发展的先前发展（包括信用评分的出现）一样，使用替代数据和分析方法也引发了有关如何有效保证新技术发展与适用的消费者保护法相一致的问题。因此，上述机构鼓励企业在使用 / 规划使用替代数据的时候秉持负责任的态度。

国际律所 Ashurst 成立金融科技法律实验室

近日，国际知名律师事务所 Ashurst 宣布成立金融科技法律实验室（FinTech Legal Labs）计划，为创企和投资人了解金融科技相关法律问题提供更多空间和机会。据悉，首次活动将于 11 月 20 至 22 日举行，主要讨论知识产权、数据与隐私、监管、企业结构、技术协议与合作、招聘与税收等多个问题。

亚太动态

澳大利亚参议院委员会发布金融科技与监管科技问题报告

2019年9月11日，澳大利亚参议院宣布成立金融科技与监管科技专责委员会。该专委会根据早期和非正式地与金融科技和监管科技初创企业的交流回复整理并撰写了一份问题文件（Issues Paper）并于10月23日发布。报告明确了提高澳大利亚金融科技与监管科技竞争力的5个主题方向，其中包括：1）资金与成本：报告文件将其称为“关乎创新发展生死的核心问题”，并在意见征询过程中主要关注吸引私募股权和风险投资的机制、采购与IT设备成本等内容；2）税收政策：专委会将讨论针对初创企业和投资工具的税收减免等优惠政策；3）技术与人才：专委会将探索各种新途径，解决高速发展企业面临的人才短缺问题，比如改革高技术人才的签证政策；3）文化氛围：专委会致力于打造澳大利亚本土的创新文化，将澳大利亚打造成世界金融科技行业的领军力量；3）监管法规：专委会将调查《消费者数据权利法案》计划的应用和成本。此外，该机构还将探讨与共享KYC支票和IFTI（国际资金转移指令）报告有关的创新，以及将政府财产数据提供给金融科技公司的可能性。

日本金融厅发布《金融行业网络安全报告》

2019年10月28日，日本金融厅（简称FSA）发布了一份金融部门网络安全专题报告。报告依据日本金融厅发布的《加强金融部门网络安全政策方法》，总结了2018年度整个日本

金融领域网络安全方面的现状和常见问题。该报告指出，大型金融机构正在越来越多地使用新技术，尤其是云服务和RPA，并且正在积极学习新知识并聘用专家以确保适当的技术风险管理。鉴于数字化的行进导致对外部供应商的依赖日益增加，至关重要的是对技术外包也采取相应的措施。

新加坡金管局与中国银保监会达成跨境监管合作协议

10月21日，新加坡金融管理局（MAS）和中国银行保险监督管理委员会（CBIRC）签署了监管合作谅解备忘录。双方表示，将强化双方跨境监管合作，促进双方在银行保险监管及危机管理方面的信息交流和沟通。当日，双方还举行专门的主题圆桌会议，讨论了中新两国在数字银行、基于风险的资本框架、保险中介市场等方面的监管发展。

新加坡金管局联合德勤和标普开发金融科技研究平台

11月12日，新加坡金融管理局、德勤和标准普尔全球市场财智宣布合作开发了一个面向整个金融科技行业的研究平台原型，帮助投资者和金融机构寻找适合投资或合作的金融科技初创公司。这个数字化金融科技研究平台旨在提供更透明的金融科技公司信息，包括商业模式、应用案例、融资和科技堆栈。平台也通过综合来自公司本身和第三方的信息，增强投资者对金融科技公司的信心。此外，用户也能在平台上对比不同公司，

根据他们重视的因素选出更符合要求的业者，从而提高决策速度。

第四届新加坡金融科技节顺利举办，重点关注可持续发展与绿色金融

11月11日至15日，第四届新加坡金融科技节(Singapore Fintech Festival)在新加坡顺利举办。新加坡金融科技节由新加坡金融管理局、新加坡银行业协会合作举办，已经成为全球规模顶尖的金融科技领域的盛会，同时为全球新晋金融科技发展成果提供了展示平台。新加坡政府积极发挥监管作用，不仅将可持续发展与气候变化列为本次科技节大会的首要主题，同时承诺将在未来推出一系列绿色金融发展激励措施，覆盖绿色基金、贷款、债券、保险和风险转移解决方案等多个领域。此外，2020年，新加坡金融管理局可能会颁布一份全新的金融科技投资计划，总额约为2.25亿新元，重点关注网络安全与人工智能领域，希望吸引更多创新企业入驻新加坡，推动当地技术进步、增加创新行业就业岗位。

韩国金融服务委员会提出金融科技发展的24项关键任务

12月4日，韩国金融服务委员会(FSC)宣布将大力推动金融科技产业规模化发展，并推出8项不同领域措施，涉及24项关键任务。其中，关键措施包括完善现行监管沙盒制度、进行监管改革以促进金融科技发展、降低金融业入门限制、建立数字时代的监管基础、开发金融创新的新增长引擎、促进金融科技投资，并建立以私人行业投资为核心的风险投资生态系统、协助金融科技企业进行海外扩张以及扩大公共领域对金融科技企业的支持。

中国人民银行下发《个人金融信息(数据)保护试行办法(初稿)》征求意见

10月10日，中国人民银行下发《个人金融信息(数据)保护试行办法(初稿)》发至各家银行，目前正在征求意见中。据了解，《办法》中第十二条规定：(金融机构)不得从非法从事个人征信业务活动的第三方获取个人金融信息；第十八条规定，金融机构不得以“概括授权”的方式取得信息主体对收集、处理、使用和对外提供个人金融信息的同意。待到《办法》正式出台后，银行将根据该办法的要求，对提供业务数据的第三方机构进行摸排，对不能保证数据来源合法性的数据供应商，要停止合作。

央行等部委将金融科技产品纳入国家统一推行的认证体系

10月29日，中国人民银行、国家市场监督管理总局联合发布公告，将金融科技产品纳入国家统一推行的认证体系(以下简称“国推认证”)，以标准落地实施为手段，持续强化金融科技安全与质量管理，切实防范因技术产品质量缺陷引发的风险向金融领域传导，着力提升金融科技守正创新能力和综合治理水平。其中包含《金融科技产品认证目录(第一批)》以及《金融科技产品认证规则》。《金融科技产品认证目录(第一批)》包含客户端软件、安全芯片、安全载体、嵌入式应用软件、银行卡自动柜员机(ATM)终端、支付销售点(POS)终端、移动终端可信执行环境(TEE)、可信应用程序(TA)、条码支付受理终端(含显码设备、扫码设备)、声纹识别系统、云计算平台共计11种产品种类。《金融科技产品认证规则》中明确了金融科技产品认证将“型式

试验+获证后监督”作为基本认证模式。同时，金融科技产品认证实行统一的认证标志管理。

中国银保信牵头打造的“中国银保信行业信息联盟链”已实现应用落地

中国银保信联手部分险企打造的首个保险领域行业级区块链——“中国银保信行业信息联盟链”已实现在电子保单托管平台等业务场景中应用落地，并取得了积极成效。该“联盟链”主要依托以 HyperLedger 的 Fabric（HyperLedger 是 Linux 基金会下的一个项目，Fabric 是其中的一个子项目）为基础的区块链技术，在联盟成员中采用共享账本的分布式存储方式。利用基于“中

国银保信行业信息联盟链”的区块链技术难以篡改、信息分布存储等特点，打造行业互信联盟，防止保单信息单方面篡改，增强了保单的互认公信力。

中国人民银行启动金融科技创新监管试点工作

12月5日，中国人民银行宣布支持在北京市率先开展金融科技创新监管试点，探索构建符合我国国情、与国际接轨的金融科技创新监管工具，引导持牌金融机构在依法合规、保护消费者权益的前提下，运用现代信息技术赋能金融提质增效，营造守正、安全、普惠、开放的金融科技创新发展环境。

2020年第一季度征稿主题如下

《交易技术前沿》由上海证券交易所主管，上交所技术公司主办，以季度为单位发资料，主要面向全国证券、期货等相关金融行业的信息技术管理、开发、运维以及科研人员。

2020 年一季度征稿主题如下：

一、云计算

(一) 云计算架构

主要包含但不限于：云架构剖析探索，云平台建设经验分享，云计算性能优化研究。

(二) 云计算应用

主要包含但不限于：云行业格局与市场发展趋势分析，国内外云应用热点探析，金融行业云应用场景与实践案例。

(三) 云计算安全

主要包含但不限于：云系统下的用户隐私、数据安全探索，云安全防护规划、云安全实践，云标准的建设、思考与研究。

二、人工智能

(一) 应用技术研究

主要包含但不限于：语音识别与自然语言处理，计算机视觉与生物特征识别，机器学习与神经网络，知识图谱，服务机器人技术。

(二) 应用场景研究

主要包含但不限于：智能客服、语音数据挖掘、柜员业务辅助等。

主要包含但不限于：监控预警、员工违规监控、交易安全等。

主要包含但不限于：金融预测、反欺诈、授信、辅助决策、金融产品定价、智能投资顾问等。

主要包含但不限于：金融知识库、风险控制等。

主要包含但不限于：机房巡检机器人、金融网点服务机器人等。

三、数据中心

(一) 数据中心的迁移

主要包含但不限于：展示数据中心的接入模式和网络规划方案；评估数据中心技术合规性认证的必要性；分析数据中心迁移过程中的影响和业务连续性；探讨数据中心迁移的实施策略和步骤。

(二) 数据中心的运营

主要包含但不限于：注重服务，实行垂直拓展模式；注重客户流量，实行水平整合模式；探寻数据中心运营过程中降低成本和提高服务质量的途径。

四、分布式账本技术（DLT）

(一) 主流分布式账本技术的对比

主要包含但不限于：技术架构、数据架构、应用架构和业务架构等。

(二) 技术实现方式

主要包含但不限于：云计算 + 分布式账本技术、大数据 + 分布式账本技术、人工智能 + 分布式账本技术、物联网 + 分布式账本技术等。

(三) 应用场景和案例

主要包含但不限于：结算区块链、信用证区块链、票据区块链等。

(四) 安全要求和性能提升

主要探索国密算法在分布式账本中的应用，以及定制化的硬件对分布式账本技术性能提升的作用等。

五、信息安全与 IT 治理

(一) 网络安全

主要包括但不限于：网络边界安全的防护、APT 攻击的检测防护、云安全生态的构建、云平台的架构及网络安全管理等。

(二) 移动安全

主要包括但不限于：移动安全管理、移动互联网接入的安全风险、防护措施等。

(三) 数据安全

主要包括但不限于：数据的分类分级建议、敏感数据的管控、数据共享的风险把控、数据访问授权的思考等。

(四) IT 治理与风险管理

主要包括但不限于：安全技术联动机制、自主的风险管理体系、贯穿开发全生命周期的安全管控、安全审计的流程优化等。

投稿说明

1、本资料采用电子投稿方式，投稿采用 word 文件格式（格式详见附件），请通过投稿邮箱 ftt.editor@sse.com.cn 进行投稿，收到稿件后我们将邮箱回复确认函。

2、稿件字数以 4000-6000 字左右为宜，务求论点明确、数据可靠、图表标注清晰。

3、本期投稿截止日期：2020 年 3 月 30 日。

4、投稿联系方式 021-68813289, 021-68800293 欢迎金融行业的监管人员、科研人员及技术工作者投稿。稿件一经录用发表，将酌致稿酬。

《交易技术前沿》编辑部
证券信息技术研究发展中心（上海）

附件：投稿格式

标题：(黑体 二号 加粗)

作者信息：(姓名、工作单位、邮箱) (仿宋 GB2312 小四)

摘要：(仿宋 GB2312 小三 加粗)

关键字：(仿宋 GB2312 小三 加粗)

一、概述 (仿宋 GB2312 小三 加粗)

二、一级标题 (仿宋 GB2312 小三 加粗)

(一) 二级标题 (仿宋 GB2312 四号 加粗)

1、三级标题 (仿宋 GB2312 小四 加粗)

(1) 四级标题 (仿宋 GB2312 小四)

正文内容 (仿宋 GB2312 小四)

图：(标注图 X. 仿宋 GB2312 小四)

正文内容 (仿宋 GB2312 小四)

表：(标注表 X. 仿宋 GB2312 小四)

正文内容 (仿宋 GB2312 小四)

三、结论 / 总结 (仿宋 GB2312 小三 加粗)

四、参考文献 (仿宋 GB2312 小四)

杂志订阅与反馈

各位读者，如您想订阅《交易技术前沿》纸质版，欢迎扫描右侧二维码填写问卷进行订阅，同时可以向我们提出关于《交易技术前沿》的建议与意见反馈。如您希望赏阅电子版，欢迎访问我们的电子平台 <http://www.sse.com.cn/services/tradingservice/tradingtech/sh/transaction/> (或扫描封面尾页二维码)。我们的电子平台不仅同步更新当期的文章，同时还提供往期所有历史发表文章的浏览与查阅，欢迎关注！





扫描在线浏览

联系电话：021-68813289

021-68800293

投稿邮箱：ftt.editor@sse.com.cn

ITRDC

证券信息技术研究发展中心（上海）



中国上海浦东南路528号

邮编：200120

公众咨询服务热线：4008888400

网址：<http://www.sse.com.cn>

内部资料 免费交流

本资料仅为内部交流使用，本季度印750册，编印单位为上交所技术有限责任公司，面向证券期货行业发送，印刷日期为2020年1月，印刷单位主人印刷厂。

部分图片或文字来源于互联网等公开渠道，其版权归属原作者所有。如有版权相关事宜，请发送邮件至ftt.editor@sse.com.cn