附件一：申报主题目录

申报主题

（以下主题均不限于给定的建议研究方向，可基于研究者背景及兴趣确定。）

1. 机器学习

1.1 受限条件下的机器学习

受限条件下机器学习是一个备受瞩目的研究方向。机器学习算法在实际应用过程中，经常遇到标注数据、训练资源有限，部分任务无显性负样本等问题。如何在这些受限条件下确保机器学习算法的效果，是一项实用价值凸显，挑战性极强的工作。

建议聚焦图像分类问题，基于上述研究方向，解决目前存在的挑战。

建议研究方向：

1) 单样本学习、半监督学习、小样本学习；
2) 单类分类问题（PU learning）、开放集迁移学习（Open set learning）；
3) 有限 GPU 资源下的高效学习。

返回目录

1.2 图神经网络算法研究与应用

近年来，图神经网络（GNN）可以自然地整合节点信息和拓扑结构，已经证明在图数据学习方面很强大的能力。GNN 的这些优势说明其在社交推荐的应用方面可能存在巨大潜力。在社交推荐中的数据可以表示为用户-用户社交关系图和用户-item 的行为交互图;学习用户和物品的潜在关系是关键。然而，基于 GNN 算法构建社交推荐系统还面临挑战。比如：多源异质图的 GNN 算法的设计、图采样算法、大规模图的 GNN 算法训练和推理等。

建议研究方向：

1) GNN 算法优化设计：考虑社交影响力、好友采样方式等条件；
2) 异构图的 GNN 算法研究；
3) 大规模 GNN 算法的训练和推理。

返回目录

1.3 机器学习在材料信息学和新材料设计研发中的关键技术研究

人工智能技术已经在图像、自然语言处理、安全等领域取得了显著成果，并被应用于化学，生物学，制药等跟自然科学更加密切相关的领域。在材料科学中材料的模拟、功能材料的设计、以及材料性质的预测等问题一直都是材料研究领域的基本问题。

近年来，科学家们已经开始尝试用人工智能技术中一些方法，比如深度学习等，来帮助降低材料研发的周期和成本。人工智能中的深度学习，神经网络等方法都具有帮助解决这些材料科学基本问题的巨大潜力。

建议研究方向：

1) 通过机器学习的方法，利用已有的材料模拟或者实验数据（网上开源数据或者运用传统模拟软件进行模拟，合作团队可以提供协助），探索新的材料性质与性能的预测方法，或提升材料模拟效率的有效途径；
2) 通过机器学习的方法，比如深度学习，结合传统的模拟工具产生的数据（数据需要运用传统模拟软件生成，合作团队可以提供协助），对材料在微观尺度建模，加速和改进传统的材料模拟方法；
3) 在腾讯云端建设材料信息学数据库，并接入高通量并行计算的自动化作业流。

返回目录

1.4 基于机器学习技术的能源行业研究

目前，能源行业不仅面临产能的问题，也面临着节能减排的问题。鉴于能源行业积累了大量的数据，机器学习是一个解决这些问题的潜在可行的技术。所以本项目旨在利用公开数据源（如 UCI 机器学习数据库等），推动机器学习技术在能源行业的创新应用，例如能源转化和利用，以期促进能源行业的可持续发展。

建议研究方向：

1) 基于机器学习技术的电力负荷预测；
2) 基于机器学习技术的锅炉燃烧控制；
3) 基于机器学习技术的最大产能预测；
4) 其他能源行业创新方向。

返回目录

1.5 基于强化学习和博弈论的多智能体协作与对抗研究

多智能体协作与对抗问题作为人工智能领域的核心关键问题，广泛影响诸如量化投资、分布式集群控制、无人驾驶等多个领域。基于战术竞技、棋牌等游戏虚拟环境进行多智能体研究的前沿探索，已在国内外人工智能研究领域形成广泛共识。本课题以虚拟世界的多智能体协作和对抗为切入点，重点研究强化学习和博弈论在游戏领域的应用性问题。在研究过程中，由腾讯自主研发的人工智能开放平台将提供相应的程序框架和计算资源。

建议研究方向：

1) 完美信息条件下，基于强化学习和博弈论的多智能体协作与对抗研究；
2) 非完美信息条件下，基于强化学习和博弈论的多智能体协作与对抗研究。

返回目录

1.6 广告转化率预估场景中的若干挑战

转化率预估是在线广告投放技术链条中的核心环节之一。转化目标的多样性，以及普遍存在的转化数据延迟等，给预估任务带来了诸多挑战。本课题旨在寻找高效的应对与解决方案，提升转化率预估准确性。

建议研究方向：

1) 设计实用的多目标/多任务学习算法与模型结构；
2) 对转化延迟建模，或者针对性改进预估模型，减少转化延迟带来的不利影响；
3) 探索统一的用户转化兴趣构建方式，融合多样的用户转化行为；
4) 小样本学习：在训练数据较稀疏的场景下，仍能保证相当的预估准确性。

返回目录

1.7 基于机器学习方法的代码智能辅助技术研究

以深度学习为代表的机器学习理论在软件研发领域已有诸多成功实践，如自动代码补全和智能提示可以提高软件研发效率，软件缺陷检测和自动修复可以提高软件质量，克隆检测和和代码水印可以用来保障代码的合法高效复用。目前代码智能辅助技术仍然是一个热点研究领域，特别在海量代码库等大尺度数据规模条件下，如何更好辅助程序员的代码开发工作，具有非常重要的产业实践意义。

建议研究方向：

1) 基于深度学习等机器学习理论的软件研发质量研究，如代码质量评估、缺陷预测及自动修复方法；
2) 基于机器学习、知识推理及 NLP 方法的软件研发辅助研究，如代码自动补全、智能提示、注释自动生成机制；
3) 软件产权保护和可追溯性研究，如代码克隆检测和传播跟踪。

返回目录

1.8 深度学习在软件安全领域的应用研究

随着软件复杂度的不断提升，大规模源代码和二进制软件的漏洞挖掘工作面临新的机遇和挑战。本研究项目希望把深度学习相关技术（例如自然语言处理、图神经网络、深度强化学习等）应用于软件安全研究中，其成果可以对传统的逆向工程、模糊测试、漏洞挖掘等有较大促进。

建议研究方向：

1) 计算机语言的表征和分类研究，例如识别二进制软件对应的编译器、编译优化选项、第三方库、开发作者等信息；
2) 计算机语言的自动生成和翻译技术研究，例如自动生成用于编译器(解释器)模糊测试的符合语法结构的程序代码；利用机器翻译技术实现二进制和源代码之间的相互翻译工作；

3) 面向复杂交互程序的智能分析方法研究，例如研究代码相似性分析、演化 API 的误用检测、基于程序内状态机的符号执行、用户界面和运行时事件等复杂输入驱动的软件测试方法。

返回目录

1.9 给定模型和数据集下超大 BS 评估与收敛性研究

在机器学习训练场景中，经常通过使用多机多卡来加速训练从而提升迭代效率，但这随之产生了 BS（batchsize）收敛的问题，导致收敛精度下降或不收敛。本课题将研究在给定数据集和模型的情况下，如何科学评估 batchsize 的合理范围，以及评估后，如何在单卡到多卡的扩展过程中，有效保持线性收敛。腾讯将为合作者提供加速机器学习的平台来验证实验效果，并有机会在现场环境中落地。

建议研究方向：

在常见的主流开源模型上，实现一套完整的大 BS 收敛性量化评估手段，通过可以适用于业务实践的科学评估手段，评估最为合理的 batchsize 值，并在多卡扩展中保持线性收敛。

返回目录

2.数字图像处理与计算机视觉

2.1 对比无监督表示学习

无监督表示学习是一个重要的学术研究热点。近期基于对比学习的无监督方法大幅度拉近了无监督与有监督方法之间的性能差距，展现了其潜在的应用价值。但如何进一步提高该类方法的性能，并将其应用于各类视觉任务仍是一项非常具有挑战性的工作。

建议聚焦对比无监督表示学习（Contrastive Unsupervised Representation Learning）算法在人脸识别中的应用,并结合其他前沿研究领域进展开展研究工作，包括但不限于图神经网络模型、元学习等。

建议研究方向：

1) 无监督的场景自适应人脸识别算法研究；
2) 无监督的大规模人脸识别模型预训练算法研究；
3) 无监督的多模态人脸识别算法研究；
4) 无监督的元学习研究。

返回目录

2.2 由视觉信息生成声音素材及空间音效

本项目旨在充分挖掘视觉及声音的关联性，提出在视觉输入下产生声音的任务，进一步可以根据视频画面动态变化产生相匹配的动态空间音频效果。这类功能可以在影视及虚拟现

实场景中得以应用（自动为虚拟场景生成声音），为有视觉障碍的人提供对图像或视频的参考信息。本项目拟利用已公开的视频-声音数据集（日常多类场景，目标对象包括人和动物等）进行模型评估。目标是产生相当真实的声音，并且与视觉输入具有良好的时间同步性和空间效果。

建议研究方向：

1) 建立输入为视觉信息，输出为单通道声音的端到端模型；
2) 建立输入为视觉信息和单通道音频，输出为高阶空间音频的端到端模型；
3) 视觉场景目标检测、识别与场景辨识，建立目标语义到声音的映射；
4) 对视觉目标场景进行检测，提取空间及深度信息以获得目标移动轨迹，从而生成空间音频效果。

返回目录

2.3 长视频的表征学习

视频的表征学习是视频理解的重要任务。目前基于 3D 卷积结构的视频特征学习的研究主要关注如何获取视频中短小片段（clip）上的信息。一方面，短小片段的信息有一定的局限性，不能提取视频序列在较长时间内随着时序演进产生的有效信息。另一方面，现有的长视频建模的方法大多数从视频中稀疏采样特定的帧，然后在多帧图像特征的上层做融合。这类方法虽然捕捉到了全局信息，但是由于融合方式比较简单，不适合完成细粒度的任务。此外，使用稀疏的图像帧提取特征容易遗漏一些细节信息。而由于 3D 卷积复杂度高，直接将 3D 卷积应用到长视频建模会有很多实际的问题，如内存和计算资源消耗过大等。此题目将探索如何高效并且有效的学习长视频中的特征表示，并将其应用到视频相关的计算机视觉问题中。。

建议研究方向：

1) 设计一种基于长视频上下文关联挖掘的高效率视觉特征提取的模块；；
2) 将此模块应用到大规模视频识别标准公开数据集测试性能，并尝试根据腾讯内容数据建立大规模的(私有)长视频数据集。

返回目录

2.4 基于 AI 技术的视频图像交通分析和数据建模

实时采集的道路交通图像和视频是感知交通动态的最直接数据源。本项目通过研发和应用最新的视频图像 AI 技术来自动识别解析视频图像内容、构建准确的全方位交通动态模型、智能理解复杂场景驾车行为及突发事件。视频和图像来源包括路口路测固定设备采集及移动车辆设备采集。

建议研究方向：

1) 基于机器学习技术的视频图像目标检测分类和追踪，包括车辆行人识别分类统计和交通流分析；
2) 视频图像内物体的空间精准定位、场景关系识别、时序特征建模、突发事件提取、驾驶行为分析建模、道路设施健康检测等；
3) 交通视频图像 AI 模型的压缩、提速、迁移和终端部署等技术研究。

返回目录

2.5 指纹与人脸认证算法安全性分析

目前刷脸支付及指纹支付已经在许多场景普及，由于涉及用户资金安全， 厂商们一直不断努力在硬件层面与机器学习算法层面提升其产品的安全性，但现有生物认证算法仍面临对抗攻击等新的威胁，如何测试与提升人脸与指纹认证算法的安全性是一项非常重要的课题，这将有助于保障腾讯公司金融支付产品的安全。为了提高研究效率，我们团队将提供测试终端，用于预训练的图像数据集以及其它安全方面的技术指导。

建议研究方向：

1) 如何攻击物理世界中黑盒的金融支付级设备的活体检测（指纹/人脸）匹配算法与模型；
2) 如何结合对抗攻击提升生物认证算法与模型的鲁棒性。

返回目录

2.6 安全可靠的新一代人脸识别技术研究

近年来，基于深度神经网络的人脸识别技术取得了突破性的进展，并且成功地应用于很多重要领域，包括金融、安防、社交、电商等。然而，深度神经网络存在着严重的潜在的安全隐患，难以防御对抗样本攻击、后门攻击等各种新兴的攻击手段。因此急需开发安全可靠的新一代人脸识别技术，以确保人脸识别技术能更安全可靠地为社会大众服务。

建议研究方向：

1) 新的攻击手段和方法研究；
2) 有效的防御手段和方法研究；
3) 在真实场景中安全可靠的新一代人脸识别系统研究。

返回目录

2.7 针对 CG 内容的图像质量评价与增强

游戏画质是决定玩家体验的主要因素之一。目前，基于视频流的云游戏方案已成为主流，但由于算力、带宽、显示设备等因素限制，游戏图像在整个编解码-传输-显示流程中必然存在损失。因此，一个可靠的图像主观体验评价模型、及对于受损图像的恢复与增强方法，对测量与优化云游戏的编解码、传输、终端显示设备适配等方案至关重要。

图像与视频的感知质量评价、以及针对主观感知体验而进行的图像增强、超分辨率等工作多年来一直是业界与学界的研究重点。但目前的研究大部分针对自然图像，针对游戏画面的相关研究却非常匮乏。游戏作为 CG 内容，无论在数据分布、人眼感知特性上都与自然图像有极大的不同。这类研究目前对于计算机图形学、计算机视觉、编解码、图像处理等领域都有着重大的意义。

建议研究方向：

1) 针对 CG 内容的图像质量评价模型；
2) 针对 CG 内容的图像去噪、增强与超分辨率技术。

返回目录

2.8 游戏视频自动特效编辑

为了使视频具有更强的表现力，能够吸引更多的用户进行观看，内容创作者往往会对视频进行二次加工，选取其中的精彩片段，通过添加特效、蒙太奇等方式，使得最终呈现的视频具有更强的节奏感和故事性。如何通过算法对视频进行自动的视频编辑和特效加工，是机器创作领域一项有挑战性的重要课题。

本课题主要关注游戏视频场景下的自动视频编辑，希望通过计算机视觉、音频处理算法对游戏视频和背景音乐进行分析和理解，结合计算摄影学，为视频添加蒙太奇（可包括慢放、回放、暂停、倒放、放大局部、添加滤镜、摇镜头等），使得呈现的游戏视频更加生动。

建议研究方向：

1) 基于视频内容理解的自动视频编辑技术：通过对游戏视频内容的分析和理解，自动添加蒙太奇，丰富视频的表现手法；
2) 基于音频节奏分析的自动视频编辑技术：通过对音频节奏和情感的分析，自动为音频段选择合适的视频片段，并添加蒙太奇，使得画面节奏与音频节奏相一致；
3) 如何提取游戏视频内容和音频内容的语义标签信息。

返回目录

2.9 基于学习的游戏动画合成和角色控制

游戏普遍采用动画切片+状态机的方法，对角色动画进行精确控制，以快速响应玩家输入，并获得逼真的动作、确定性的行为。近些年有研发团队尝试采用 Motion Matching 的方法，从动画数据库中检索最合适的动作序列，可达到与动画状态机相当甚至比它更好的效果，但在内存和计算的消耗上都比较大，应用场景受限。

基于本次合作，团队希望共同探索一种基于学习的 locomotion 动画合成方案，合成的动画品质接近 Motion Capture 所获得的动画，达到尽量短的响应时间的同时，避免滑步、抖动、动作细节丢失等问题，并可以覆盖 locomotion 足够多的动作类型（走、跑、跳、爬、后退、蹲伏前进等），能够感知并适应于包括起伏、障碍物及常见建筑在内的虚拟环境。

建议研究方向：

采用监督学习方法，从 MoCap 数据集中学习，来合成 locomotion 动画，合成过程中既可以响应游戏玩家的控制指令，也可以适配环境。

返回目录

1. 知识图谱与自然语言处理

3.1 深度文本理解技术探索

文本理解技术被广泛应用于搜索、个性化推荐、广告匹配、智能对话等场景，用来对自然语言文本进行结构化分析与处理。随着近年来深度学习方法的兴起，文本理解技术取得了很大的进步，但是在深度理解文本语义方面的能力和水平，跟人类相比还有较大差距。本命题研究和探索基于语义分析和知识推理的深度文本理解技术。

建议研究方向：

1) 深层语义分析模型和技术；
2) 文本理解的新模型和架构；
3) 引入常识及外部背景知识的文本理解模型；
4) 细粒度命名实体识别及其语义分析；
5) 知识图谱的表示、构建和推理；
6) 以及文本理解技术在相关场景中的应用。

返回目录

3.2 信息安全领域知识图谱构建与应用

通用领域深度学习模型的可解释性差，且缺乏先验领域知识。知识图谱是一种可人工编辑的结构化知识载体，如果将这些知识融合到深度学习模型中，则能够有效提升模型的可解释性和可编辑性，使得模型能够根据具体业务场景进行人工定制。目前学术界和工业界均在知识图谱方面进行布局，并取得了显著成效。然而，各大公司将知识图谱视为内部资源而不愿意开源，导致可获得的高质量知识图谱较少。另一方面，针对信息安全这一特定领域，领域知识图谱将有助于改善各项下游业务的效果。因此，构建一个信息安全领域知识图谱是一项十分有必要的工作。

建议研究方向：

1) 构建信息安全领域百万级的知识图谱；
2) 研究知识图谱与深度学习模型融合的方法，解决模型解释性差的问题；
3) 针对信息安全领域下游任务，提供基于知识图谱的解决方法。

返回目录

3.3 医疗自然语言理解

医疗自然语言处理面临患者口述口语化、标注难度大、临床电子病历结构化等难题，我们希望通过医疗领域的自然语言理解技术，如：医疗知识图谱和医疗语言模型，来提升深度学习模型在健康助手、在线问诊、智能诊断等拥有大规模用户的产品中的表现。

建议研究方向：

1) 医疗语言模型，知识蒸馏；
2) 长文本分类，阅读理解，摘要生成，文本匹配；
3) 问答系统、对话系统；
4) 低资源的信息抽取和知识图谱扩增方法、医疗知识常识推理问答等。

返回目录

3.4 多模态医学知识图谱

医学数据的信息化产生了大量的多模态数据，包括文本数据，图片数据，影像数据，时序数据等等。这些数据中蕴含着大量的知识，而目前没有被很好的挖掘利用。知识图谱是一种表达能力强、扩展性好的知识表示方式，能够兼顾人类认知与机器自动处理。如果能将医学多模态数据中蕴含的知识抽取出来，并且以知识图谱的形式表示，可以更好的支撑行业应用，例如基于知识图谱编码和路径的智能推荐、基于知识图谱实体关系的多轮对话系统等。

建议研究方向：

1) 从海量的医学多模态数据中挖掘出知识，并且以知识图谱的方式进行表示抽象；
2) 使用多模态知识图谱落地应用到实际的医学场景，包括基于知识图谱的医学文章推荐和患者多轮对话系统等。

返回目录

3.5 面向 K12 阶段的教育领域 AI 技术研究

伴随着在线教育的持续发展，教育领域内相关计算机技术的应用进入“深水区”——从早期的“教育+互联网”逐步演进到“教育+AI”。教育领域内的应用，比以往任何时候更注重通过 AI 技术的运用来达成教育效率与教育效果的提升，进而推动教育公平的实现。本课题主要关注与 K12 阶段教育领域相结合的 AI 技术研究。

建议研究方向：

1) 教育领域的基础 NLP 问题，包含且不限于词法分析、句法分析、篇章分析以及学科相关的语义理解（如数学公式理解）；
2) 教育领域的学科图谱构建，以及学科图谱与教育资源（习题、教案、教学视频等）的自动关联技术；
3) 自适应学习技术，包含且不限于知识追踪、题目及学习资料推荐；
4) 自动批改技术，包含且不限于作文批改、数学应用题批改以及问答题批改。

返回目录

4. 语音信号处理与语音合成

4.1 面向实时语音通信处理算法的客观质量评估

实时语音通信应用需要部署去混响、噪声消除、丢包补偿等各类算法应对由于环境影响或网络损伤带来的质量降低，而在精细化衡量各类算法效果和性能的过程中，作为主观评估方法的替代和补充，需要使用客观化的方法或方案，来贴近用户主观感觉，并降低评估的难度和提高评估结果的重复性。本课题将开展相关的前沿研究和工程技术创新，腾讯也将为项目提供丰富的落地实践场景。

建议研究方向：

1) 研究衡量单通道/多通道去混响或抑噪算法效果的评价手段，以单一或者复合的指标以及可复现的评估方案来确定去混响/抑噪处理在提高声音质量和可懂度方面的性能；
2) 研究衡量回声消除算法效果的客观评价手段，以单一或者复合的指标以及可复现的评估方案来确定算法处理的性能；
3) 研究以无参考的方式来建模和评估被算法处理或丢包补偿之后的单向话音质量，提供新的方案来预测主观用户评估的结果。

返回目录

4.2 多说话人多风格的高可控通用语音合成系统研究

现今大部分语音合成系统所涵盖的说话人及风格比较有限，当业务应用对特定嗓音及风格产生需求时，往往需要从头进行数据录取，标注及模型训练。此研究旨在扩大用于语音合成系统训练的数据规模并探索能够有效接受各类控制信号的模型结构及训练准则，构建出涵盖各类说话人音色及风格的通用语音合成系统。此系统在合成时能通过控制信号灵活合成出如朗读，对话，傲娇，嗲声嗲气等各类语音。

建议研究方向：

1) 扩大现今用于合成系统的训练数据规模，利用尽可能可以利用的数据，如语音识别数据，在线有声读物数据等；
2) 研究能够有效接受各类控制信号的神经网络结构；
3) 研究能够充分利用各类风格控制监督信号的训练准则。

返回目录

5. 密码学

5.1 基于国密算法体系的密钥保护机制

对于密码学来说，在客户端如何保证密钥的安全存储是一个基础的应用场景，一般的密钥嵌入代码、文件加密存储方案都有泄漏的风险存在。更极致的做法是采用硬件 KEY 的方案，但是在移动设备场景或者现在流行的小程序场景，不太可能让用户随身携带硬件 KEY。

同时，2020 年是国密密码法实施的第一年，国家以法律要求使用国密算法，但业界在这方面的成熟应用还比较欠缺，特别在类似 web/小程序这种高风险的运行环境下，如何保障密钥安全，是一个很基础但又没有很好安全解决方案的需求，目前市场需求很大，具备很高的实用性。

本课题希望结合基于国密算法体系的机制，在保证实现正常数据加解密，签名验签功能的前提下，在协同签名的基础上更进一步研究密钥保护机制，并符合 GM/T 0028-2014 《密码模块安全技术要求》安全二级要求。

建议研究方向：

1) 通过软件沙箱机制来隔离密钥数据；
2) 基于同态加密、零知识证明等技术，在不泄漏密钥数据到内存的情况下实现加解密，签名验签；
3) 协同签名的基础上进一步加强保护本地部分切割密钥的机制。

返回目录

5.2 基于数据隐私保护的多方联合建模

随着产业数字化的推进，许多行业开始基于多方联合平台进行建模协作，特别是业界热门技术区块链平台。以区块链平台来说，在这个过程中，多方联合建模协作需要通过区块链的分布式账本技术实现多方联合建模，通过区块链访问参与方相关数据，但所访问的共享数据涉及到参与方的商业机密，且参与方又不希望敏感商业数据透露给对方。采用密码学做隐私保护是一个被广泛认可的研究方向，业界提出了零知识证明、安全多方计算、同态加密、密文等值测试等不同技术解决方案。本命题希望结合有明确隐私保护需求的特定业务场景，如多方联合建模风控、多方广告精准投放，通过采用密码学等技术，提出安全性和性能都符合商用标准的解决方案。

建议研究方向：

1) 基于密文等值测试方案,实现云数据存储检索共享隐私保护；
2) 基于区块链的分布式账本安全多方计算，在不透露参与方交易数据的情况下，实现两方以及多方协作建模风控体系；
3) 广告精准投放协作过程中,如何在不暴露多方商业原始数据情况下进行可信计算和深度学习，以达到精准投放。

返回目录

6. 数据库

6.1 基于数据状态实现数据一致性的并发访问控制技术

数据库的事务处理领域，有多少种数据异常？数据异常和并发访问控制技术、基于依赖图的可串行调度技术，之间的关系是什么？是否存在一个模型：能够描述已知数据异常的本质，能够发现更多数据异常，能够用一个模型统一各种数据异常的描述并揭示数据异常和并发事务之间的关系（之前的技术都是单独 case by case 式地描述一个个数据异常，散乱没有逻辑）？

建议研究方向：

1) 通过深入研究数据异常的本质，找出更多的数据异常，并进行形式化证明；
2) 通过深入研究现有的技术如依赖图、冲突可串行化等技术，挖掘这些技术的优缺点，确认这些技术“为什么能解决数据异常？”；
3) 构造新的算法和体系，实现“可串行化”目标；
4) 研究分布式系统中，数据异常和并发访问控制技术以及隔离级别等。

返回目录

7. 智慧交通

7.1 基于多源信息融合的交通信息处理方法

多源信息融合就是充分利用不同时间与空间的多传感器信息资源，在一定准则下加以综合分析，得到被测对象的一致性解释与描述。随着国家新基建项目开始在全国范围内推动智慧道路建设及智能网联汽车在国内的推广，可以预见，数量众多且种类繁杂的传感器将部署在主干道路上或者被智能网联车辆搭载。通过这些多源信息数据，可以获取包括车速、车流量、道路占有率，湿滑雨雪等天气信息在内的丰富的实时道路交通信息，并有助于对道路交通状况的变化做出准确预测，从而实现更加有效的交通控制和管理。这是一项有挑战性的重要课题，其成果将有助于腾讯智慧出行及智慧交通产品的战略决策和及时改进。

建议研究方向：

1) 研究基于车联网环境的多源信息融合系统架构；
2) 利用多源信息融合，实现面向百公里以上骨干路网的交通状况趋势预测；
3) 对接腾讯车路协同平台，搭建基于多源信息融合的交通信息分析系统，并具备对外展示的能力。

返回目录

8. 微众银行区块链与大数据专题

8.1 跨链协议的事务机制研究

当前区块链平台种类众多，各种平台的接口、协议、架构等多方面都存在异构的情况，要实现跨区块链之间的交互就需要有一套可靠的跨链协议。跨链协议除了要打通异构区块链平台的交互，还需要实现在异构平台之间的原子性操作。现有的跨链协议包括哈希时间锁定、中继、侧链、分布式密钥交换等，但是大部分仅考虑了资产交换这种单一场景。在面对更复杂的联盟链场景下的数据交换、合约接口调用等场景，如何保证跨链操作的原子性，保证跨链访问的事务性，还需要有更多理论和工程层面的探索和突破。

建议研究方向：

本课题希望就跨链协议事务机制展开研究，包括链间协同、链上链下协同机制的设计，结合分布式一致性技术原理进行研究，也可以结合密码学、博弈论等机制进行设计等。可以选择类似以上内容的一点或几个点展开，可以结合 FISCO BCOS 和 WeCross 平台进行研究。

返回目录

8.2 区块链系统和信息安全研究

区块链网络需要运行在包括公网、局域网、云间网络等多种网络上，连接多个机构，承载高价值数据和交易，在用户机构的私钥管理、节点的网络层、存储层、计算层（合约），共识算法、以及密码学算法均有可能遇到安全风险，本课题研究针对联盟链的攻击模型，包括且不限于网络渗透、DDOS 攻击、化身攻击、访问劫持攻击、注入攻击、智能合约漏洞攻击、恶意记账者等等。可以选以上内容的一点或几个点展开，可以结合 FISCO BCOS 开源联盟链平台进行研究。

建议研究方向：

可从网络渗透、DDOS 攻击、化身攻击、访问劫持攻击、注入攻击、智能合约漏洞攻击、恶意记账者等方向或其他相关方向选择开展研究，结合 FISCO BCOS 开源联盟链平台进行研究。

返回目录

8.3 高效可用的场景化隐私保护机制研究

隐私保护关乎个人、机构的敏感信息，其重要性日益显著。隐私保护与区块链结合，业界已有一些研究和探索，包括可搜索加密、零知识证明、同态加密、安全多方计算、可信硬件计算环境等多种技术被运用于区块链架构中。然而这些通用的密码学技术存在较大的性能瓶颈。隐私保护的范畴是非常广泛的，包括了个人或机构身份数据的隐私、行为的隐私、状态的隐私等，不同场景的隐私保护需求也大不相同，需要针对不同场景问题设计高效可用的场景化隐私保护方案。

建议研究方向：

本课题希望研究一些特定场景下的高效可用的隐私保护方案，例如在机器学习场景下的隐私保护、在数据挖掘场景下的隐私保护、在边缘计算场景下的隐私保护等。可以选择类似以上场景的一点或几个点展开，可以结合 FISCO BCOS 和 WeDPR 平台进行研究。

返回目录

8.4 分布式数字身份的应用研究

分布式数字身份体系的技术和行业应用已成为近年来区块链研究的重点方向之一，基于区块链的分布式数字身份体系，目前以 DID 规范为范本已经有了工业化的实现（微众银行开源方案 WeIdentity），并有多个应用落地，并实现了选择性披露、零知识证明等特性。本课题期望就分布式数字身份体系的技术和行业应用方向，对行业应用场景、相关技术挑战（性能、存储、安全、隐私等）、技术和合规风险等方面进行分析。提出有创新性的场景、或致力解决传统业务领域向分布式数字化身份转换接入的痛点问题，或重点解决其中的一些技术挑战问题。本课题建议对 WeIDentity 有一定的基础性了解。

建议研究方向：

本课题期望就分布式数字身份体系的技术和行业应用方向，对行业应用场景、相关技术挑战（性能、存储、安全、隐私等）、技术和合规风险等方面进行分析。提出有创新性的场景、或致力解决传统业务领域向分布式数字化身份转换接入的痛点问题，或重点解决其中的一些技术挑战问题。

返回目录

8.5 区块链和物联网

区块链和物联网是近年来业界最关注的方向之一，如何利用区块链的优势解决物联网的痛点已成为行业关注的焦点。该问题域非常广泛，包括从广域网到边缘计算再到区块链整个范围，牵涉设备和设备之间、设备商和设备商之间的、基于区块链的协作模式。

建议研究方向：

可选择物联网设备和区块链的接入协议、设备标识和管理、路由寻址、隐私保护、消息响应模型等，重点解决一到几个问题，或者提出一个场景（工业、家居、交通、能源、社会治理等等），设计完整闭环的物联网和区块链联合工作场景，并重点识别其中的商业痛点、技术痛点、提出解决方案。

返回目录

8.6 小微企业信用与风险评估

目前被使用的小微企业信息有限，现有的信用模型主要依赖企业纳税记录等信息建立风险模型。在除税务信息之外，还有其它有关的企业信息，无法快速评估一个企业的信用和风

险程度，因此需要探索研究是否能用户风险评估，尽可能挖掘出信用可靠的企业，响应国家大力支持小微企业融资贷款的政策。

建议研究方向：

1) 法律文书，比如法庭公告，判决书等包含大量风险信息，研究这些并提炼出相应机制进行企业风险评估；
2) 某企业用户在使用 A 产品，研究分析预测该企业用户可能会使用 B 产品的概率；
3) 从公开信息我们可以获取到企业法人、高管等个人姓名，但是这些姓名存在大量的同名现象；
4) 研究通过外部信息来区分同名的不同自然人，使得我们在不需要个人身份敏感信息的情况下，建立起自然人跟企业的关系；
5) 研究企业关系图谱，丰富图谱中的数据纬度，用于优化企业营销和风险模型。

返回目录