

《计算机新技术讲座》第一次报告

71119103 许润

一、报告内容

讲座主题：机器学习在行业中的应用

报告人信息：

王君，IBM 资深大数据分析架构师、数据科学家。作为技术带头人，目前除了带领 IBM SPSS 研发团队为 IBM 大数据分析产品如 DSX, WA, Modeler, Statistics 等开发机器学习算法，同时也为企业做大数据分析方案，多次主持并带领团队为企业量身打造大数据分析平台，涉及的领域有交通、制造业、银行等。

企业信息：

IBM 是国际商业机器股份有限公司的缩写，是美国一家跨国科技公司及咨询公司，总部位于纽约州阿蒙克市。IBM 主要客户是政府和企业。IBM 生产并销售计算机硬件及软件，并且为系统架构和网络托管提供咨询服务。截止 2013 年，IBM 已在全球拥有 12 个研究实验室和大量的软件开发基地。IBM 虽然是一家商业公司，但在材料、化学、物理等科学领域却也有很高的成就，利用这些学术研究为基础，发明很多产品。比较有名的 IBM 发明的产品包括硬盘、自动柜员机、通用产品代码、SQL、关系数据库管理系统、DRAM 及沃森。

讲座内容：

随着 AI 及大数据的发展，人们的日常生活中或多或少都会遇到 AI 类的产品。其背后的机器学习和深度学习技术为何？在行业中有什么应用？如何运用这些技术开发实践？众所周知，在行业内部及行业之间的各种数据，它们存在着相互影响的关系。如何在众多的有纷繁复杂关系的数据中识别出一些关键的信息，挖掘并量化这些影响并将其用于行业量价分析交易中有着很大的价值。然而完成这项工作却有着很大的挑战性。我将详细介绍如何应用 IBM Watson Studio 产品中的时间因果和时空预测机器学习算法来解决行业应用难题。

二、收获和扩展学习

报告收获：

机器学习是一门内容丰富的交叉学科，涉及计算机、概率论、统计学、凸分析、计算复杂性理论等多门学科，属于数据科学的一种。复杂的组成部分也造就了机器学习广阔的使用范围，无论对于图像、视频的处理，还是对于语音、情感的识别等等，机器学习都能完美胜任。

机器学习是怎么获得这么大的能力的？或者说，机器学习到底是怎么工作的。王老师提到，机器学习的首要任务其实是学习决策函数。所谓决策函数，就是一种从数据到标签的映射关系，它表征了实际数据的概率分布，是一种统计意义上的客观规律。

在进行机器学习训练的时候，我们往往需要大量的训练数据。如果我们假设这个决策函数存在的话，我们就可以确定它的形式，使用拟合参数的方法进行学习，所以决策函数的学习就转化成了参数的学习。以 MLP 中的 BP 算法为例，我们认为三层以上的感知机模型级联就可以拟合任意的函数形式，因此决策函数的学习就转化为线型层中的权重的学习。我们使用特定的目标函数进行损失的计算，反向传播更新参数，从而进行学习。

王老师给我们介绍了 Cross industry standard process for data mining。CRISP-DM 的主要过程为：商业理解、数据理解、数据准备、建模、模型评估、部署等六个环节，涵盖了从实际问题到机器学习问题的建模并完成最终的评估和部署的全过程。在了解了大致的流程之后，我们需要知道什么样的模型可以用机器学习模型解决。通常来说，这类问题都有一些共性的特点，比如需要有大量可用的数据资源作为支撑；有明确的代价以构建代价函数；由数量较多的、人类难以思考全面的因素决定等。

最后，王老师为我们展示了很多具体应用等实例，包括企业智能审计助手、企业智能招投标助手、企业智能 HR 平台等，体现出机器学习模型卓越的交互能力、推理发现能力和学习决策能力。相信未来机器学习依然会是行业智能化的中流砥柱。

扩展学习：

如今，机器学习已成为科技公司运作方式的核心。诸如数据可视化、智能工作流和 RPA（机器人流程自动化）之类的工具使组织能够为超级自动化时代做准备，从而允许不同的业务应用程序协同工作以提高效率。

1、自动化任务并解决数据孤岛的复杂性

机器学习将成为三大主流趋势：人才、数据和信任。首先，组织将通过自动化来解决数据科学家的短缺问题，比如新的自动 AI 和一键式机器功能。为数据科学家自动化低级任务将缩短实现价值的时间。第二大趋势将是解决混合云空间中存在的数据孤岛的复杂性。最后，当企业希望部署和扩展 AI 时，信任将成为中心焦点。组织正在意识到，他们必须能够在整个 AI 生命周期中信任他们的模型和业务成果。在 2021 年，他们将继续使用可以帮助提高透明度、管理风险，并建立对 AI 的更大信任的工具。

2、机器学习在银行业务中的重要作用

银行业未来将会更多地使用 AI、机器学习和生物识别技术以及更少的密码。银行将机器学习与生物识别技术相结合，以提供新的体验，例如面部和指纹验证而不是密码。我们已经看到的一个例子是银行利用机器学习来检测和读取物理证件，以进行身份扫描。客户使用智能手机扫描身份证，然后进行自拍照。最后，银行利用生物特征面部比较技术和活体检测来验证身份的真实性和完整性，从而确认个人的身份。

但是，人工智能和机器学习解决方案的未来是透明性和监视潜在的偏见，这是组织不能再忽视的问题。例如，一家美国的国际投资银行高盛集团（Goldman Sachs）在 2019 年成为调查的对象，当时消费者抱怨称，其苹果信用卡服务（Apple Card）与男性顾客相比，向女性申请人提供的信用额度较低。那些将 ML 工具用作其运营的一部分的公司将需要采取措施，以确保该技术严格遵守道德标准。”

3、更加关注 AI/ML 技术的安全性

虽然今天的大部分关注点都集中在将 AI 技术应用于网络安全上，但是公司很快将在保护他们的 ML 模型和确保算法健壮性方面投入更多的精力。到 2024 年，50% 的大型组织将部署隐私增强技术来支持其 ML 应用程序。客户数据对于洞察力和个性化的价值正在增加，但隐私方面的挑战也在增加，尤其是在全球监管力度加大之后。

4、机器学习在网络安全任务自动化中的作用

机器学习和自动化具有潜力，可以通过处理耗时的任务（例如对安全警报进行优先级排序，减少误报，以及将设备映射到 IP）来释放网络安全分析员的精力，使他们可以从事更重要的战略活动。他们还可以增强安全团队快速检测攻击者行为的能力，这些行为原本需要大量时间才能进行手动调查。ML 还可以建立员工档案，包括他们的同龄人小组和个人电子邮件地址，使分析师能够比以往更快地识别内部威胁。

《计算机新技术讲座》第二次报告

71119103 许润

一、报告内容

讲座主题：微服务：从概念到实现

报告人信息：

李琪，IBM Consulting(原 GBS) 应用架构师。

企业信息：

IBM 是国际商业机器股份有限公司的缩写，是美国一家跨国科技公司及咨询公司，总部位于纽约州阿蒙克市。IBM 主要客户是政府和企业。IBM 生产并销售计算机硬件及软件，并且为系统架构和网络托管提供咨询服务。截止 2013 年，IBM 已在全球拥有 12 个研究实验室和大量的 软件开发基地。IBM 虽然是一家商业公司，但在材料、化学、物理等科学领域却也有很高的成就，利用这些学术研究为基础，发明很多产品。比较有名的 IBM 发明的产品包括硬盘、自动柜员机、通用产品代码、SQL、关系数据库管理系统、DRAM 及沃森。

讲座内容：

本讲座介绍微服务架构风格这种开发方法。我们将介绍如何以开发一组小型服务的方式来开发一个独立的应用系统,并展示如何围绕业务功能进行服务构建,如何通过全自动的部署机制来进行独立部署,实现以不同开发语言和不同的数据存储技术做最低限度的集中管理。

二、收获和扩展学习

报告收获：

讲座结构清晰明了，主要分为 What、Why、How 三个部分、即什么是微服务，为什么要使用微服务，怎样实现微服务。

微服务架构风格是指通过开发一组小服务来组成一个单一应用的方法。这每一个小的服务都自成一体，运行在自己的进程里，彼此之间通过轻量级的机制，通常是 HTTP 资源 API 的方式来进行通信。这些小服务一般基于既定的业务能力范围来进行构建，通过完全自动化部署的机制来进行独立部署。而微服务可以从两个方面理解：什么是微？什么是服务？“微”要遵循以下两个基本前提：①业务独立性首先,应保证微服务是具有业务独立性的单元,并不能只是为了微而微；②团队自主性其次,考虑到团队的沟通及协作成本，一般不建议超过 10 人。当团队超过 10 人，在沟通,协作上锁耗费的成本会显著增加,这也是大部分敏捷实践里提倡的。

微服务架构的特点是组件服务化、按业务能力组织服务、服务即产品、智能终端与哑管道、去中心化、基础设施自动化、兼容设计。因此每个服务足够内聚，足够小，代码容易理解，开发效率高；服务直接可以独立部署，让持续部署成为可能；每个服务可以各自进行水平和垂直扩展，而且每个服务可以根据需要部署到合适的硬件和软件上；提高容错性，一个服务的问题不会让整个系统瘫痪；系统不会长期限制在某个技术上。但微服务同时也存在一些非常明显的缺点：增加了故障排除的难题，运维的技术复杂、相应的运维成本（测试、变更、部署）高，必须建立开发运维一体化机制 Devops，必须有完备的监控手段和自动化恢复手段，分区数据库可能带来的业务数据同步与一致性问题，微服务的接口将成为变更的敏感位置。

SOA 非常庞大，有的时候要在 SOA 上做一些开发与运维，而微服务和 SOA 是互利共存的。SOA 专注于对“大型服务”进行编撰操作，但这些大型服务也可以通过对一系列微服务进行组合而实现，所以说服务的大小并不是定义微服务的好方法。

李老师分别从开发、测试人员视角，服务管理人员视角，业务应用拥有者视角这三个不同的角度，阐述了微服务的价值，即我们为什么要用微服务。但是在使用微服务前也要考虑很多前提，如是否具有快速交付能力、基础监控能力、快速部署能力，代码复杂度等等。

关于怎样构建和管理微服务架构的应用，李老师给出了自己的见解，分为服务识别、业务建模、协作开发、服务测试、服务部署、服务运维六个步骤。在构建微服务软件时使用的分布式服务框架也拥有很多特性，这些特性提供了很多复杂问题的解决方案，比如异步调用、超时失败策略、故障隔离、健康检查、流量控制以及自定义路由等，这些特性大大减轻了开发者处理逻辑的负担，还有一些高性能、高可用的保障都增加了我们构建微服务的信心。而且经过多年的微服务实践者的摸索，也总结出了很多辅助运维系统，比如统一日志收集（ELK）、服务管理、服务监控和服务治理等，大大减轻了运维的工作，而且能让我们对复杂的微服务系统做到心中有数。

最后是微服务 PassS 云平台介绍，PaaS 经历了四代的发展，从最开始的 Heroku 到 Cloud Foundry 的 DEA 版本、Diego 版本，最后到 Kubernetes 以及其它容器编排引擎，基于容器+Kubernetes 的新型 PaaS 将会成为云计算的主流。

扩展学习：

目前，国内外学者对基于微服务技术的 SOA 架构及其实现机制的研究进展十分重视，微服务具有独立进程、轻量级通信和独立部署环境等显著的特性；将系统整体功能分解到单个微服务中以实现系统的解耦；这种方式不仅可以降低系统的耦合性，提升系统的内聚性，而且减少服务交互的成本，提供更加灵活的服务支持。

现在微服务设计不仅限于大型的复杂的系统，对于中小型的应用也在采用微服务的模式和思想，提高应用系统扩展潜力。这几年的发展让微服务越来越受关注，一些成熟的技术解决方案也被熟知，应用的领域也越来越广泛，随着技术的不断成熟，微服务架构会成为技术开发的常用手段。

与传统开发模型相同，选择合适的未来开发平台对于微服务十分重要。微服务如何与 2 个主要的新兴平台进行集成，即云平台 and 物联网。随着科技新兴技术与数据时代的到来，两个平台很可能在互联网行业中占主导地位。由于微服务本身具有移植性和可伸缩性等特性，在物联网上运行存在部分难题，若在云平台中运行微服务似乎是恰当的选择。但从系统安全性角度出发，存在部分功能具有低计算能力且具有较高风险的缺点。因此，微服务与具体应用平台相结合，解决微服务与平台相集成的特定实施方案以及安全方案需求，变得更加迫切。

《计算机新技术讲座》第三次报告

71119103 许润

一、报告内容

讲座主题：解锁深度学习，开启 TensorFlow 实战之旅

报告人信息：

杜可，IBM 软件工程师，IBM 中国系统实验室，大数据与 AI 平台 Spectrum Conductor 的开发工程师。从 2016 年起研究与实践 IBM Watson，在实验室先后从事大数据和 Watson 人工智能平台研发，目前从事 Data and AI 平台中知识图谱系统研发工作。

企业信息：

IBM 是国际商业机器股份有限公司的缩写，是美国一家跨国科技公司及咨询公司，总部位于纽约州阿蒙克市。IBM 主要客户是政府和企业。IBM 生产并销售计算机硬件及软件，并且为系统架构和网络托管提供咨询服务。截止 2013 年，IBM 已在全球拥有 12 个研究实验室和大量的软件开发基地。IBM 虽然是一家商业公司，但在材料、化学、物理等科学领域却也有很高的成就，利用这些学术研究为基础，发明很多产品。比较有名的 IBM 发明的产品包括硬盘、自动柜员机、通用产品代码、SQL、关系数据库管理系统、DRAM 及沃森。

讲座内容：

深入浅出介绍深度学习，由简单的感知机到复杂的 CNN、LSTM、物体检测，以及时下最热门的生成对抗网络，带你一览 AI 的大脑深度学习的奇妙世界，并且教你如何上手 TensorFlow，编写一个炫酷的神经网络

二、收获和扩展学习

报告收获：

讲座从关于 IBM、人工智能与深度学习，使用 TensorFlow 开发神经网络模型，迁移学习、翻译和生成对抗网络，分布式 TensorFlow，IBM WMLA 介绍以及深度学习应用案例等几个方面展开。

首先，杜老师引入了一个有趣的例子——啤酒厂都在招 AI 工程师。工程师可以用 AI 来解决啤酒生产的完美配比以及啤酒灌装的问题。从此就可以看出，AI 已经和我们的日常生活生产息息相关了，也侧面反应了 AI 的重要性。

那么什么是人工智能呢？人工智能是计算机学科的一个分支，通过了解智能的实质，产生能以人类智能相似方式做出反应的智能机器，包括机器人、自然语言识别、图像识别和专家系统等。而人工智能也分为强人工智能和弱人工智能，目前我们仍处于弱人工智能阶段。

总的来说，机器学习就是“总结过去，预测未来”，即从数据中不断学习、总结经验，形成一种模式。深度学习就是一种学习算法，是人工智能领域的一个重要分支。它源于人工神经网络的研究。机器学习是一种实现人工智能的方法，深度学习是一种实现机器学习的技术。我们可以看下机器学习下的深度学习和监督学习以及非监督学习，那它们之间是什么关系呢，其实就是分类方法不同而已，他们之间可以互相包含。所以在深度学习中我们可以用到监督学习和非监督学习，而监督学习中可以用到很基础的不含神经元的算法（KNN 算法）也可以用到添加了多层神经元的深度学习算法。深度学习与传统算法的区别就在于能否自动的进行特征提取。比较经典的神经网络有一般神经网络、卷积神经网络、递归神经网络以及生成对抗网络。

接着是深度学习的开发框架。如何开发深度学习应用呢？主要有两种方法，一是使用现有的 API Service,另一种是使用深度学习开发框架搭建模型，如 Theano、Torch、Pytorch、TensorFlow、Caffe、CNTK。本次主要介绍的就是 TensorFlow，TensorFlow 是一个开源的、基于 Python 的机器学习核心框架，它由 Google 开发，可以帮助您开发和训练机器学习模型。并在图形分类、音频处理、推荐系统和自然语言处理等场景下有着丰富的应用，是目前最热门的机器学习框架。Keras 是一个由 Francois Chollet 于

2014-2015 年开始编写的开源高层深度学习 API，目的是提供一个简单。灵活。强大的深度学习开发框架。它把底层运算（例如 `add`, `matmul`, `transpose`, `conv2d`）封装成常用的高层神经网络模块，增强易用性和代码的可读性，提高深度学习开发者编写模型的效率。

开发神经网络模型简单的神经网络的训练过程包括以下几个步骤：①定义一个包含多个可学习参数（权重）的神经网络；②对输入的数据集进行迭代计算；③通过多层网络结构来处理输入数据；④计算损失值（输出值与目标值的差值）；⑤反向传播梯度到神经网络的参数中；⑥根据更新规则来更新网络中的权重值。

TensorFlow 还有一个优点就是迁移学习和微调，就是如何将学习到知识从一个场景迁移到另一个场景。机器翻译自出生到现在经历了四个时期的演变和发展：基于规则的机器翻译、基于实例的机器翻译、基于统计的机器翻译以及神经网络机器翻译。**GAN(Generative Adversarial Nets)**是用对抗方法来生成数据的一种模型。和其他机器学习模型相比，**GAN** 引人注目的地方在于给机器学习引入了对抗这一理念。**GAN** 是生成模型的一种，生成模型就是用机器学习去生成我们想要的的数据，正规的说法是，获取训练样本并训练一个模型，该模型能按照我们定义的目标数据分布去生成数据。主要应用于生成图像数据集，图像转换和文字和图片之间的转化、图片编辑、图片修复等等。

最后就是深度学习应用案例。有信用卡识别、肿瘤扩散评估、太阳能电池板的自动品控分类。人工智能正在潜移默化的影响我们，改变着消费品与时尚、音乐创作、银行甚至是客户服务。

扩展学习：

深度学习除了在已经广泛应用的领域不断深化完善，也在持续拓展新的应用领域，推动各个行业的智能化。得益于硬件和算法的进步，深度学习的强大能力逐渐显现。这种进步体现在连续学习过程中使用模式识别的算法，使它们能够自主训练执行任务，而不需要显式编程代码。

一些较成功的深度学习应用集中在图像识别、文本分析、产品推荐、欺诈预防和内容管理等领域的增量和实际改进上。深度学习可能会带来未来更强大、更具颠覆性的应用发展，如无人驾驶汽车、个性化教育和预防性医疗保健。

深度学习除了在应用领域的快速扩张，一些基础技术研究依然保持着很高的热度。主要可以分成以下几个方向：深度学习的基础理论研究;深度学习和其他方法的结合扩展;深度学习一些优势的深耕和发扬光大;深度学习现阶段一些局限性的解决。当然这几个方面也是相互影响，相互交织的。

深度学习的理论性研究侧重给出对深度学习有效性更好的解释，对深度学习的泛化能力更好的证明，以及如何更快收敛更好地学习等。在和其他方法结合方面，一方面最成功的当属深度强化学习，从静态监督学习扩展到动态交互时序学习，这在 **AlphaGo** 上已经得到了验证。但是对于深度强化学习是否就是最接近人工智能的解决方案还是存疑的。另一方面是在一些非游戏类的实际应用中，深度强化学习发挥好的效果往往需要针对性的设计和实验。除此之外，也有很多关于神经和符号相结合的探索、数学模型和认知心理学方法相结合的探索。

《计算机新技术讲座》第四次报告

71119103 许润

一、报告内容

讲座主题：全真时代，音视频技术在互联网中的应用

报告人信息：

Ensonhu，腾讯高级后台工程师（腾讯云与智慧产业事业群），硕士毕业于东南大学。17 年校招加入腾讯后，一直从事腾讯云音视频产品开发工作，目前主要负责音视频海外以及超低延时直播相关技术和业务工作。

企业信息：

腾讯是一家世界领先的互联网科技公司，用创新的产品和服务提升全球各地人们的生活品质。腾讯成立于 1998 年，总部位于中国深圳。公司一直秉承科技向善的宗旨。我们的通信和社交服务连接全球逾 10 亿人，帮助他们与亲友联系，畅享便捷的出行、支付和娱乐生活。腾讯发行多款风靡全球的电子游戏及其他优质数字内容，为全球用户带来丰富的互动娱乐体验。腾讯还提供云计算、广告、金融科技等一系列企业服务，支持合作伙伴实现数字化转型，促进业务发展。

讲座内容：

主要介绍音视频技术概述，以及在中国互联网近 10 年的一个演变发展历程，以及以腾讯云为代表中国音视频云产品走向国门的一些技术发展历程。

二、收获和扩展学习

报告收获：

Ensonhu 首先分享了他的职业选择历程，他对于自己的定位就是想从事海量服务的开发工作，基于自己的想法和兴趣，因此最终选择了腾讯的音视频团队。

接着开始进入本次讲座的主题，认识 PaaS 音视频。PaaS 音视频已经广泛扎根于各行各业之中，包括教育、电商、游戏、影视、娱乐、体育等等方面，并逐渐成长为行业发展不可或缺的基础建设能力。而腾讯云音视频更是从 IaaS、PaaS、SaaS 等方面都有了很多方面的涉足。腾讯云音视频不仅仅做国内的方案，从 2017 年开始，腾讯的机房建设和服务部署能力已经延伸到了海外，在全球范围内拥有 2000+ 的加速节点，120Tbps+ 的带宽储备，覆盖了 50 多个国家和地区。在国内外很多方面都彰显出了腾讯云音视频的实力，其中就包括 2019 年的 SX 赛事，在千万级用户并发和数亿人次播放量的情况下为电竞赛事保驾护航，提供了大片级超清观赛体验、安全可靠远距离传输、毫秒级超低延时直播等服务。

然后 Ensonhu 工程师以腾讯云为例，开始讲述音视频广阔的技术方向。从方向的选择上，会有音视频、智能 AI、云储存、云网络、云数据库、云大数据、云服务器、云物联这些方面。然后以秀场直播为例，介绍了直播架构是什么样的流程以及使用的技术栈。

首先是主播方，它是产生视频流的源头，由一系列流程组成：第一，通过一定的设备来采集数据；第二，将采集的这些视频进行一系列的处理，比如水印、美颜和特效滤镜等处理；第三，将处理后的结果视频编码压缩成可观看可传输的视频流；第四，分发推流，即将压缩后的视频流通过网络通道传输出去。其次是播放端，播放端功能有两个层面，第一个层面是关键性的需求；另一层面是业务层面的。先看第一个层面，它涉及到一些非常关键的指标，主播方播出的视频到达播放端之前，需要经过一系列的中间环节，也就是我们这里讲的直播服务器端。直播服务器端提供的最核心功能是收集主播端的视频推流，并将其放大后推送给所有观众端。另外，对于一个主播端推出的视频流，中间需要经过一些环节才能到达播放端，因此对中间环节的质量进行监控，以及根据这些监控来进行智能调度，也是非常重要的诉求。实际上无论是主播端还是播放端，他们的诉求都不会仅仅是拍摄视频和播放视频这么简单。

然后又通过经典案例分享介绍了状态击穿问题如何优化。缓存在某个时间点过期的时候，恰好在这个时间点对这个 Key 有大量的并发请求过来，这些请求发现缓存过期一般都会从后端 DB 加载数据并回设到缓存，这个时候大并发的请求可能会瞬间把后端 DB 压垮。那么如何解决呢？业界比较常用的做法，是使用 mutex。简单地来说，就是在缓存失效的时候（判断拿出来的值为空），不是立即去 load db，而是先使用缓存工具的某些带成功操作返回值的操作（比如 Redis 的 SETNX 或者 Memcache 的 ADD）去 set 一个 mutex key，当操作返回成功时，再进行 load db 的操作并回设缓存；否则，就重试整个 get 缓存的方法。

最后又总结了他这几年的工作经验，他告诉我们除了需要开发的基础，还要有优秀的架构能力，不能满足当前的技术，要学会专业进阶，学习一些新的知识，进行创新。同时一个人的力量也是有限的，要懂得跨界合作以实现价值最大化。总的来说就是：面对机遇，要抓住机遇，勇于挑战，资源整合，合作共赢，大胆创新，小心求证，及时复盘，张弛有度；面对挑战，要耐得寂寞，受得委屈，勤于学习，善于思考，控制预期，及时反馈，顶住压力，柳暗花明。

扩展学习：

音视频技术本身将会不断发展，首先是视频编码技术会快速迭代，目前在用的主流的视频编码标准是 H264，H264 第一版标准完成于 2003 年，距今已经 18 年了，H265 因为专利的原因并没有被广泛使用。

我们预期新的编码标准将会比较快的应用起来，尤其是 AV1，作为开源媒体联盟 AOM 制定的第一代标准，除了有非常好的生态支持，还提供了免费的专利政策，相比 H.265 等知识产权政策不明确的标准，有巨大的优势。

其次是 AI 技术，它对整个 RTC 行业带来的影响是非常深远的，客观地说，以深度学习为代表的 AI 技术的发展与成熟，对音视频关键技术的突破提供了另一个有希望的方向。

有些使用传统技术无法解决的问题，可以通过与 AI 技术的融合，大大降低问题解决的难度。无论是 RTC 行业的核心音视频编解码，还是其他一些音视频处理和增强技术，比如现在比较热的音频降噪、视频超分、对象分割识别、语义识别等，在 AI 的加持下，都获得了突破，从而催生了智能客服、智能监控、虚拟形象等新领域。

音视频技术也会随着 5G、AR/VR 等技术的发展而发展，将会给用户带来更加沉浸式的体验，也将催生更多行业和场景来应用音视频技术。5G 浪潮的爆发正在为实时音视频创造低时延和高带宽的网络条件，这些基础设施的改进，为超高清视频、全景视频技术的落地创造了可能。AR/VR 为音视频应用的场景和生态拓宽了道路，加速 3D 视频、虚拟现实等技术的落地，共同为下一个互联网形态“元宇宙”提供强有力的技术基建，实现科技向善的美好未来。

中国是个充满机遇的国度，短短几年成就一批行业大亨。中国的市场环境很好，全世界你找不到第二个综合素质如此高的环境（包括人口基数、消费能力、营商环境）。那么这个时代能有多辉煌，还是由行业里众人来创造。未来可能并不太远。音视频技术其实就是解决人们感官上的数字技术问题。而对于中间件厂商来说所追求的依然是，如何使其高效、多元化地融入我们这个以互联网商业模式为主导的大环境中。这是一个综合性的考量，这也是对于如何生存的考量。先要活下来，再要活得好些，然后活得长久。

《计算机新技术讲座》第五次报告

71119103 许润

一、报告内容

讲座主题：ToB 技术挑战与架构实践

报告人信息：

彭嘉，从业互联网十多年，先后投身过安全、视频、人工智能、企业应用等业务领域，深耕服务端开发；2019 年加入字节跳动，现任财务营收平台研发 Leader，负责架构设计、架构优化、技术团队管理工作。

企业信息：

北京字节跳动科技有限公司，简称字节跳动，是一家位于中国北京的跨国互联网技术公司，成立于 2012 年 3 月，旗下有产品今日头条和抖音（及其海外版本 TikTok）、西瓜视频等。至 2018 年，字节跳动的移动应用月度用户超过十亿人，估值 750 亿美元，超越 Uber 成为全球最有价值的创业公司。截至 2019 年 7 月，字节跳动的产品和服务已覆盖全球 150 个国家和地区、75 个语种，曾在 40 多个国家和地区位居应用商店总榜前列。

讲座内容：

ToB 即 To Business，面向的是企业服务市场；ToC 即 To Customer，面向的是普通用户市场。随着互联网的高速发展，渗透率增高，C 端流量红利日渐消失，B 端市场已成为互联网下半场的新增长点。本次讲座将从 ToB 的业务特点与发展趋势、ToB 的技术特点与挑战、ToB 的架构实践来进行介绍，让同学们对 ToB 的业务、技术、架构有大概的了解。

二、收获和扩展学习

报告收获：

ToB 即 To Business，通过为企业提供平台、产品、服务来获取利润。ToB 的业务模式比较多样，主要有 B2C 平台（比如京东、美团、天猫等平台）、B2B 平台（比如阿里巴巴平台）、提供咨询和解决方案产品（比如广告公司、管理咨询服务、企业培训等服务）、云基础设施服务（比如 AWS、阿里云、金山云、火山引擎等）、SaaS（比如钉钉、企业微信、飞书等）五大类。在国外，SaaS 模式已经在全世界得到了广泛认可，并得到大范围的普及和全面发展。提供 SaaS 服务产品的供应商主要有 SAP、Oracle、Microsoft、Salesforce.com 和 NetSuite 等公司，获得较大的成功。回首看 40 年前，美国刚开始发展 To B 业务，SAP（1972）、微软（1975）、甲骨文（1977）等国际巨头值此时诞生。步入 21 世纪，Salesforce、Workday、Servicenow 等百亿级公司应运而生，背后的本质是美国企业遭遇市场红利枯竭，各种降低成本提升效率的 To B 服务工具适逢其时。国内 SaaS 的起步比较晚，目前还处于早期发展阶段，每年的 ARR 还是少于 10 亿的规模，并且市场份额只占全球的 5% 左右。但是 ToB 的领域还有很多机会，市场前景方面：中国每年的 SaaS 市场份额增长率都在逐步上升，国内市场规模逐步扩大。大浪淘沙之下，中国的 SaaS 路线正在愈发清晰。相信在未来的十年甚至更久，SaaS 企业都以一种缓慢但不容阻挡的速度持续前进，成为中国产业互联网的发展的代名词。

接下来是 ToB 技术特点与挑战。以美国 To B 业务规模和投入作为对比，如果其指数为 100%，那么中国 To B 业务的规模和投入低于 5%。因此，中国的 To B 业务有着巨大的机会，但同时也面临着很多的技术挑战。第一个特点是解决方案比较复杂，有六个主要的点：复杂、非线性、涌现、反馈、套叠以及连锁。由于不同的企业业务存在着较大的差异，比如做支付服务、航空服务等，对支付的要求完全不一样，因此就造成了解决方案的复杂度。此外，To B 企业对于服务水平、可用性、扩展性要求特别高，这也造成了解决方案的复杂度。二是实施周期长。由于企业里面的分工特别明确，一环扣一环，比如评测、摸底、访谈、了解、分析、实施等，全是由不同的小组和不同的人员来完成，之间的

沟通协作需要大量的时间，过程特别漫长，真正解决技术本身问题所耗费的时间其实并不多。三是对服务水平要求比较苛刻。如果 To B 方案出了问题，会造成比较大的影响，因此就要求在实施、验收的时候做到一丝不苟，因此对服务水平要求特别苛刻。四是客户对成本的敏感度相对较低。对于企业来讲，要的是长期的、优质的、稳定的服务，宁可多花钱，也绝不希望某个应用存在一点风险，所以成本不是企业主要的考虑因素。对于 To B 业务面临的主要技术挑战，主要来自于六个方面：安全、可配置、数据隔离、可扩展、高可用、热部署。对于 To B 业务技术挑战的解决之道，每个企业应根据以上六个方面，制定适合自身公司的解决方案。企业要的是安全的系统、稳定的运维、可快速扩展的容量和快速扩展的架构，一般的产品经理是无法搞清楚这些事情的，只有架构师这种具备高素质技术且同时有一定业务经验的人，才能真正成为技术专家。

最后就是 ToB 的架构实践。彭讲师以财务营收平台系统为例，讲述架构发展主要有四个阶段 1.0、2.0、3.0 和 4.0。第一阶段是单体服务+单数据库，是最早期建立的架构，适合小团队快速搭建 MVP；第二阶段是微服务+多数据库，分为四层：业务系统、业务中台、财务中台以及财务后台，优点是逻辑清晰、易于扩展，可独立维护和迭代，生产效率大幅度提高；第三阶段是分库分表+并行计算，分库分表目前已经是成熟的技术，并以流水路由表和认款路由表为例展示了分布式系统。并行计算有任务队列、计算节点和数据节点三层。因此可以支持大规模数据，性能提升了多个量级；第四阶段是混合 OLTP+OLAP，相比 3.0，它可以支撑海量数据，性能再次提升了多个量级。

扩展学习：

中国互联网目前出现了一个很明显的趋势，即人口红利终结、流量红利终结、超常规的高速增长终结，即便如阿里、腾讯这样的互联网巨头，也开始出现用户增速和业务增速下滑的情况。就国内市场而言，ToC 模式的路变得越来越难走了，ToB 就势必会成为下一个重要的增长点。

产品制胜。前阿里巴巴的 CEO 卫哲在一次分享中说：“所谓的 B2B，实质是 Business Person To Business Person”，可谓是一语中的。毕竟所有服务的最终都是 C，都是每个活生生的人。如果 ToB 产品能在体验方面超出用户期望可能会获得意外收获。

服务制胜。ToB 和 ToC 最大的本质区别就是决策流程和决策成本不同。企业的购买决策是由流程来决定的，不会像 ToC 一样存在消费冲动。企业选择使用你的产品，有时候相当于人体更换一套血液循环系统，这是非常复杂的工程，甚至会出现很大的危险。因此，ToB 不仅是产品，本质上是产品加上服务，需要一个强大的服务团队，管理 ToB 复杂的决策流程，打消用户心中的疑虑，确保每一个环节都能够非常顺利地执行。

下沉制胜。ToC 可以“一个产品打天下”，所以容易在“马太效应”中形成赢家通吃的局面。但 ToB 不一样，很难做到一个标准产品适合所有客户，这是社会分工细化的底层逻辑决定的。ToB 业务要耐得住性子，不仅是下沉到垂直行业，更要下沉到企业的生产销售营运全流程。

数据制胜。B 端企业大部分的交易过程还是集中在线下，将线下的业务导流到线上，用数据指导业务的优化迭代，会事半功倍。充分发挥平台的品牌优势，利用集采的思路，为客户提供价美质优的、企业营运必备的财税、快递、法务、视觉设计等企服产品，在掌握基础数据的情况下，配合支付产品，结合收单数据，科学地评估企业营运现状，为客户提供相应的金融服务，助力企业的高速发展，是一个不错的思路。

ToC 改变世界，ToB 帮人改变世界，也许这就是 ToB 业务的魅力。

《计算机新技术讲座》第六次报告

71119103 许润

一、报告内容

讲座主题：Zero Trust（零信任）和 SASE（安全访问服务平台），XDR（扩展检测和响应平台）

报告人信息：

Janet Yao（姚淑依），东南大学（生物医学工程），目前趋势科技资深产品经理，曾任趋势测试经理，项目经理。

Sarah Ye（叶青），河海大学（计算机科学与技术），目前趋势科技资深产品经理，曾任趋势科技项目经理。

企业信息：

趋势科技，网络安全软件及服务领域的全球领导者，以卓越的前瞻和技术革新能力引领了从桌面防毒到网络服务器和网关防毒的潮流，以独特的服务理念向业界证明了趋势科技的前瞻性和领导地位。总部位于日本东京和美国硅谷，在 38 个国家和地区设有分公司，拥有 7 个全球研发中心，员工总数超过 6000 人。是一家高成长性的跨国信息安全软件公司。趋势科技分别在日本东京证券交易所和美国 NASDAQ 上市，并分别在 2002 年 10 月和 2004 年 9 月入选日经指数成分股（4704）和道琼斯可持续性指数。

讲座内容：

介绍了最近几年技术领域与市场营销层面的热门概念。企业在云时代面临的问题，SASE 的优势和挑战，如何构建“威胁感知”的网络，XDR 的诞生和发展。

二、收获和扩展学习

报告收获：

零信任英文是 Zero Trust。Forrester 于 2010 年提出零信任的概念，谷歌在 2014 年通过 BeyondCorp（beyond the corporate network）浇了一把油之后，成功地火遍安全界。

如它名字所言，零信任从不信任任何请求，无论请求源起何处，也无论它的目的归于哪里。请求者无法看到整座城，相反他只能看到城里的一个酒楼或者驿馆。这其实类似所谓“Dark Cloud”的概念：用户只能看到该看到的些许服务，其它的不该看的，对他而言一概屏蔽。零信任将关注的焦点从网络转移到服务本身。每一个服务都需要单独进行安全控制，这有效地阻止了横向移动，因为每个连接都是需要重新认证和授权，且是暂时性的连接。

零信任将网络安全的控制粒度细化到了每个服务层级，当然给 IT 的日常管理带来了更大的难度和更高的挑战。比如：

- （1）如何去基于每个服务进行认证和授权呢？
- （2）如何控制面向服务的暂行性连接呢？
- （3）如何在整体上进行 audit 以及事后分析呢？

那么业界给出的解决方案是什么样子呢？下面轮到 SASE 登场。

零信任和 SASE 搭配组成了软件定义的网络边界（Software-Defined perimeter，SDP），这个虚拟的边界围住了用户和企业使用的所有服务，包括 on-premise、IaaS、PaaS、SaaS。它作为一个集中式的网络控制平面以 SaaS 形式提供服务，极具扩展性和安全性。连接控制平面的一端是访问者，与他的访问位置无关，连接平面的另一端是企业所有的服务，与这些服务运行在哪里无关。

零信任用于控制哪些用户可以访问企业使用的服务，而 SASE 则用于控制对这些服务的连接。简

而言之：

- (1) 无论请求是从何处发起，发往何处，首先需要进行认证和授权
- (2) 如果请求是从企业所用服务往外，SWG 是一个很好的控制节点
- (3) 如果请求是从外部通向企业所用服务，则需要用到 ZTNA
- (4) CASB 常被用来管理企业使用的 SaaS 服务

XDR 是一种基于 SaaS 的、特定于供应商的安全威胁检测和事件响应工具，它原生地将多个安全产品集成到一个统一的安全运营系统中，该系统统一了所有许可的安全组件。XDR 系统的四个主要功能是：

- (1) 成为开箱即用的通用安全产品集合
- (2) 集中和规范化中央存储库中的数据以进行分析和查询
- (3) 多种安全产品协同工作，提高检测灵敏度
- (4) 作为恢复过程的一部分，相关的事件响应能力可以改变单个安全产品的状态

XDR 是特定于供应商的威胁检测和事件响应工具，可将多个安全产品统一到一个安全运营系统中。主要功能包括安全分析、警报关联、事件响应和事件响应手册自动化。XDR 工具需要不断更新有关攻击者工具策略和技术的情报。他们还需要数据规范化和其他形式的预处理来支持分析和关联。它们通常还需要大量基于 SaaS 的数据存储，最好是在能够连接未预定义事件的图形数据库中。XDR 工具将面向威胁的安全组件（例如 EPP/EDR、防火墙、NIPS、SEG、CASB 和 SWG）结合到一个有凝聚力的安全运营系统中。

扩展学习：

中国市场 XDR 发展前景亦十分广阔，据 CIC 最新发布的《网络安全威胁情报行业发展报告（2021 年）》显示，当前我国威胁情报市场规模约为 10.69 亿元。另外，在 11 月，全球网络安全领导企业 Palo Alto Networks 和普华永道中国宣布扩大合作，以 XDR 等产品为依托，在中国市场提供安全运营服务。同时，以腾讯安全为首的威胁情报服务提供商纷纷入局 XDR 赛道。凡此种种，都标志着 XDR 在研发投入与创新动力方面将迎来一波高潮。

聚焦到中国市场，中小企业数量占据了中国企业总数的半壁江山，2020 年底中小企业数量便已突破 4200 万家。研究机构 Gartner 发布的报告显示，2021 年全球信息安全和风险管理技术与服务支出预计将达到 1504 亿美元。为了抵御攻击，企业机构将扩展和规范威胁检测和响应活动，由于没有足够的网络安全人才或技能来推出自己的集成架构，中型企业和小型企业对 XDR 有着强烈的需求。

XDR 作为融合了“EDR+NDR+SOAR+威胁情报+安全中台+X”的一站式安全检测与响应平台，不论是传统 IT 环境，还是公有云、私有云、单云、多云、混合云架构，XDR 都能够更快、更准确地检测网络攻击活动，并以开箱即用的自动化操作快速应对各类繁琐枯燥的安全任务，有效应对告警过载难题。

当前，各类产业主体都在积极围绕以 XDR 为代表的威胁情报技术及商业模式展开探索。在有科技产业界风向标之称的 Hype Cycle（技术成熟度曲线）中，端点安全 and 安全运维都提及了 XDR 技术。此外，据 ESG 对 339 位企业安全专业人员的调查数据显示，58% 的人认为 XDR 可以通过增强、改进、聚合当前的安全分析功能来实现 SOC 的现代化，55% 的人认为 XDR 可以通过与 SOAR 集成来实现安全流程自动化。XDR 已获得市场发展的绝佳时机，安全技术厂商正抓住机会，聚焦企业需求痛点，探索 XDR 热门解决方案的发展方向。