

# 技术开拓 新领域

百度

迷你书

TECHNOLOGY  
TO EXPLORE  
NEW TERRITORY



# 序：当谈到百度技术沙龙的时候 我谈些什么

作为百度技术沙龙的组织者，过去六年我几乎参加了每场活动，也因此亲历了从互联网到人工智能的技术演进。

与其说亲历，其实是旁观，因为现实又惭愧的是我并没有因为勤勉地组织工程师技术活动就在相关知识领域有什么进展。那么让我来谈技术沙龙我能谈些什么呢？

起初我对技术沙龙最关注的是参与人数，因为数字是最直观的 KPI，座无虚席是活动有人气最直接的印证。后来我又很在意收集每一次沙龙活动的反馈，希望能从投票中取得让自己满意的结果，因为大家说好就是真的好。我还与时俱进策划过 SNS 互动，微博直播、集赞转发抽奖等等，因为这些能传播输出活动的影响力。

这些努力都收获了些效果，然而距离实现技术沙龙愿景仍有相当远的距离，那就是畅想、交流、争鸣、聚会。我观察到其他技术交流活动的现场也常常是这样的情况：时髦话题吸引着听众摩肩接踵，却鲜见与讲者热烈的讨论；嚷着要干

货的人们穿梭在分论坛之间，哪里热闹去哪里；朋友圈里刷图签到的越来越多，用技术博客写下所见所想的人越来越少。

努力会有收获，但每日精进才能成为很厉害的人，成就很厉害的事。所以在此就对希望通过参与百度技术沙龙、QCon、ArchSummit 这些工程师技术交流活动提升自我的朋友们提些旁观者的建议吧——比如在参加活动之前先对内容做充分的了解，由此做出抉择和日程；比如尽可能早到一点，占一个让自己可以看得更清楚，听得更明白的座位；比如及时离开不合口味的 Session，而不是在昏昏欲睡和吐槽中消耗时间。希望你习惯用百度脑图这样的思维工具及时总结思路，热烈地参与讨论；还可以先定一个能达到的小目标，比方说现场向讲师提出一个好问题。

因为从事技术品牌和开发者关系的市场工作，我赶过许多夜航，喝过许多创业咖啡，见识了许多正当最好年龄，用代码改变世界的人。我为自己感到庆幸。若说对百度技术沙龙更上层楼的期许，希望能做到很感谢你来，不遗憾你离开。

**百度 程鸿**

# 目录



- 06** 从有到优：百度前端接入技术的升级之路
- 12** 从海量数据、精准定位、语音识别等技术看百度车联网的实践应用
- 19** 百度室内定位技术揭秘
- 26** IT 科技公司和传统车企眼中的车联网
- 33** 百度开放云助力移动游戏快速发展
- 40** 百度语音识别和唤醒技术解析

## Mission

### 使命

整合全球优质学习资源  
帮助技术人 and 企业成长

## Vision

### 愿景

全球领先的技术人  
学习和交流平台



极客邦科技是一家 IT 技术学习服务综合提供商，旗下运营 InfoQ 技术媒体、EGO 社交网络、StuQ 职业教育三大业务品牌，致力于通过整合全球优质学习资源，帮助技术人 and 企业成长。

了解更多，请访问官网：<http://www.geekbang.org>

从有到优

# 百度前端接入技术的升级之路

作者 薛梁



由百度开发者中心和 InfoQ 联合主办的“纵谈前端接入技术、SEO 和安全运维”主题沙龙活动中，来自百度开发者中心的资深运维工程师们热情洋溢的分享了百度在前端技术、搜索速度优化和全站使用 HTTPS 技术的进展及成果，以及百度在这些方面有哪些宝贵经验可供参考的。演讲嘉宾分别为百度 Golang 委员会成员陶春华、专注于网页搜索无线访问速度的工程师许霞，和处理网页搜索可达性、安全搜索等方向事务的主要技术负责人陈曦洋。

## Go语言在Baidu Front-End方面的应用实践

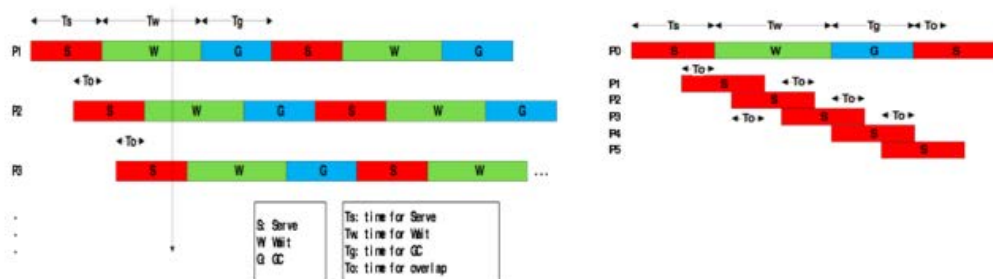
Go 语言的广泛流行取决于部署简单、并发性好、良好的语言设计，以及执行性能好。这也是在重写百度前端这一项目上为什么考虑选用 Go 语言的原因所在。陶春华老师介绍说，促使重写 Baidu Front-End 的诱因主要基于以下三点：一是修改成本高。事件驱动的编程模型本身的编码和调试难度都很大；C 语言本身的难度和开发效率有很多限制。二是配置管理方式落后。为单产品线设计，无法支持平台化要求；配置变更（修改、重载、验证）能力差。三是变更和稳定性的矛盾。例如程序出 core 也是比较头疼的事情。

在此前提之下，团队决定使用 Go 语言来重写前端，但是这里也遇到了一些问题，那就是 GC（Garbage Collection）本身难以避免的时间延迟。BFE 的需求是要在 1ms 以内，最大不超过 10ms，一旦超过这个平均值，那么用户体验将大打折扣。而 Go-BFE 实测 100 万连接，400ms GC 延迟。这就需要不断的对 GC 进行优化。

在这里陶老师也介绍了两种优化思路，第一个常见的方法就是将扫描的小对象合并成大对象，利用 Array 来合并一组对象。第二种方法精算性更高一点，可以把消耗内存较多的内容放到 C 里面，因为 Go 语言有一个 CGO 接口，直接通过 Go 调到 C 可以解决这个问题，只不过代价比较大。但是，问题和方案永远是相生相伴的。用 Array 技术重写网络库，所有的 BFE 将永远用 Array 来写，理论上可行。这里的问题又来了，第一风险太大，第二如果 Go 语言升级了，还能不能继续使用下去。



## GC优化 – 多进程配合



进程数的计算: 
$$N = 1 + \left\lceil \frac{Tw + Tg + To}{Ts - To} \right\rceil$$

图 1

陶老师在这里介绍的解决方法叫做车轮大战。即，在一组工作进程中，进程和服务是等价的，某一个进程跟服务运作到一定时间之后关闭 GC，让它休息，第二个进程代替它服务，以此轮换，构成一个车轮大战的局面。如果在不能直接解决 GC 问题的时候直接关掉服务，然后绕过它。这基本的方案思路也就是关闭继承多进程的轮转战。（如图 1 所示）

## 搜索速度优化的前进之路

在整个百度接入服务里，百度搜索一直秉承提供最基础的三个保障，那就是安全、快速、可靠。许霞首先介绍说，在对速度进行度量之前，先要对数据检测、收集。对客户端数据监测的特点是：可以检测任何对象，成本高，并且监测的指标很固定。JS 埋点检测数据的特点是：可以检测任何指标，甚至可以检测每一条结果的速度。第三方数据检测的特点是：可以定制，并且有一定的海外监测能力，但成本高。

收集数据的意义在于可以很清晰的了解掌握用户的搜索习惯，这对 PV、UV 以及变现收入有很大影响。那么如何贴切搜索引擎的特点做搜索速度的优化？通过三个方面：接入质量提升、后端处理优化和前端渲染优化。接入质量提升主要有两个考察因素：延迟和带宽，对应的也就是优化 RTT 和传输效率。

后端优化其实就是整个搜索引擎的优化了，分为缓存优化和检索优化。缓存



优化最基础的方式就是进入、淘汰机制等等，保证淘汰机制是最合理的。检索优化，则需要对硬件以及硬件方案的选择做一些深入考虑。在前端渲染优化方面，除了考虑节省时间之外还要考虑怎样让它定性化。

对优化做决定性决策只是其中的一种方法，还有更聪明的创新方法，那就是关于无线技术。这里面所涉及的内容包括手机终端、机站以及 IP 网络，传输速度当然是跟这三者有密不可分关系的。机站会根据自己能获得多少收益来处理用户的请求，尽量会缩小头部信息，进行一定程度的数据压缩。手机跟机站之间建立连接以维持这种连接关系。但电耗大是很关键的问题。百度搜索做了维持连接的一些机制，当用户页面空闲很长时间或者放在后台，就可以减少电量的消耗。

（如图 2 所示）

## 全站HTTPS能否确保网站彻底安全？

2015 年 3 月，百度搜索成为国内首家完成全站 HTTPS 改造的大型站点；且目前来看，全站 HTTPS 已经成为百度产品的首要标准；同时，统一接入平台也大

## 关于无线 | 连接的维持

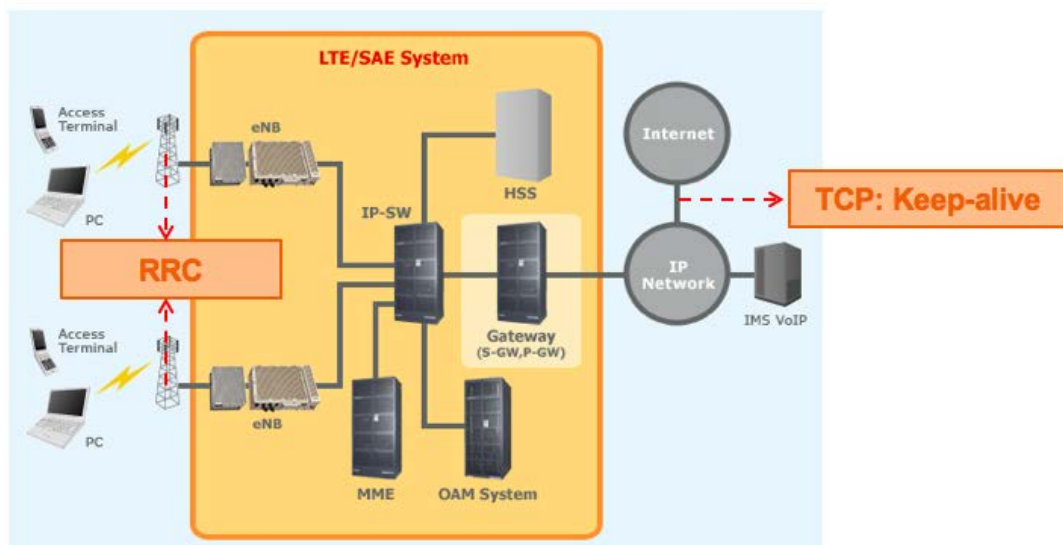


图 2

幅提升了 HTTPS 的接入效率和性能。陈曦洋老师在开讲前是这样介绍大背景的。全站 HTTPS 的原因是为了让用户保持良好的使用体验，解决反馈较多的劫持和隐私泄露等问题。这些问题的具体 case，包括页面被加上 URL 参数，不停刷新；页面被 DNS 劫持到其他网站；用户手机号码遭泄漏；白页，搜索功能异常等等。正是出于对用户数据的安全保密，维护网站正常运作的考量，百度专门成立了由百度搜索和运维部组成的 HTTPS-SUPPORT 团队，对 HTTPS 进行深入研究，提供完整的服务，保障用户正常访问百度原始产品。

陈曦洋老师在这里详细介绍了全站 HTTPS 改造的成本，这也是很多人都比较关心的焦点问题，这不仅涉及到证书的部署，还会涉及到激增的计算量，需要多次协商和握手，而用户端搜索的延迟将会给 HTTPS 改造需要解决的问题。除此以外，对于一个大型网站而言，架构如何解决多业务部署 HTTPS 的问题，巨大的页面和模板数量，以及如何解决实际部署中的各种问题，让用户无损 / 平滑的完成切换，其实是更具有挑战性的工作。

计算性能涉及到密钥（证书）的长度，1024 和 2048 位在性能有什么差别呢？原来使用 HTTP 协议的时候，假设 cps 可以达到 2w+，而转换成 HTTPS 之后，cps 只能达到 2-3 千；在访问速度方面，使用了 HTTPS 之后，不做任何优化，访问百度的速度可能会恶化 250-500ms，一些设计比较差的页面可能会恶化 500-1200ms；在架构和产品成本方面，对于百度这样的综合性网站，牵一发而动全身，仅仅是在页面形式上就要改大量的模板，成本相当大。

那么有没有可选的优化方案呢？陈老师认为，性能优化上优先使用 ECC。这里使用 ECC 密钥长度大小要比 RSA 和 DH 密钥长度短。在硬件的优化上则可以使用硬件加速卡，可以做 TLS 的远程卸载（小型站点在不面对大量的恶意请求时完全可以通过纯软件卸载，只需要保证连接复用率）。在访问速度上的优化上，通过复用连接和协议优化可以尽量减少握手次数，就可以让它接近于原始 HTTP 的性能。怎么去减少握手次数？比如 Session cache 和 Session ticket 可以极大的减少用户在一定时间内再次访问时的计算开销，而 HSTS 能在浏览器内部完成 HTTP 到 HTTPS 的跳转，不再经过一次网络传输和浏览器开销。另外还可以用

## 优化方案 | 访问速度

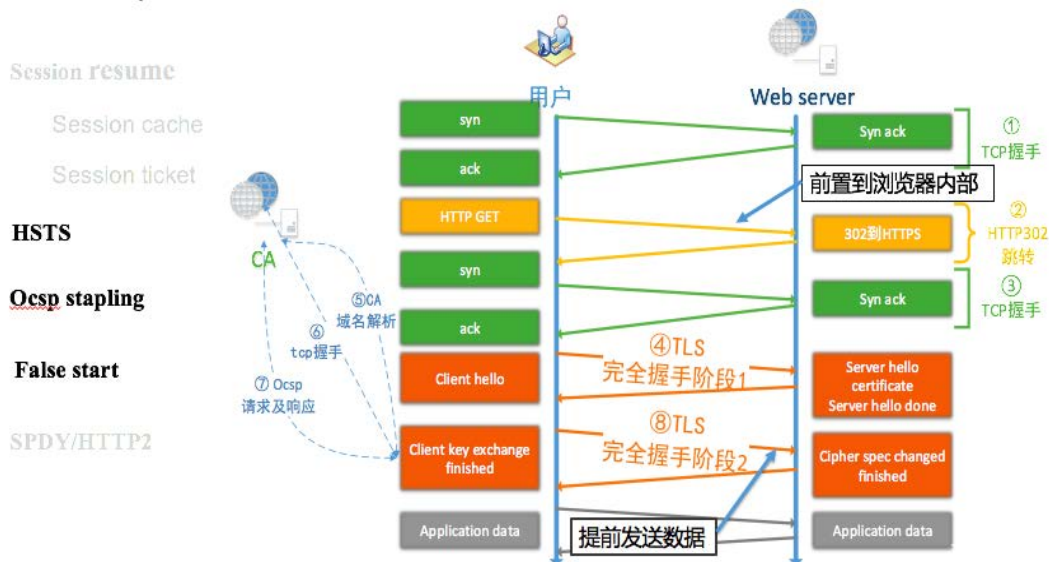


图 3

SPDY-HTTP2 方案，优点是基于单连接，能进一步提升连接复用比例，协议支持 header 压缩，在无线网络下有重要意义，这些都可以提高访问速度（如图 3 所示）。

除了对协议层进行优化之外，也可以在应用层做些优化，预连接就是一个很好的优化方案。比如在网页端或者客户端，用户发起访问请求之前提前把这个握手过程完成，减少延迟，这一点也很重要。另外陈老师建议站点在发展到一定规模时一定要做整体的接入规划，控制域名数量，一些服务需要变成公共的，比如图片，静态资源的存储和访问。

在最后，陈老师也回答了大家普遍比较关心的问题，那就是使用 HTTPS 就代表着绝对安全吗？事实上并没有绝对地安全，代码是人写的，很多问题都是实际的实现上或者依赖的其他环境上出现了漏洞，OpenSSL HeartBlood 就是最典型的案例，甚至连随机数的生成和一些加密算法上也可能有人为埋下的漏洞，CDN 回源这样的路径很多情况下也是使用的 HTTP。百度使用 HTTPS 只能保护用户在浏览百度产品的时候的安全，但是很多手机号的泄露是第三方站点导致的（它们会通过非法渠道购买识别用户手机号的服务），这个问题并不能通过百度的 HTTPS 解决。但是相对于 HTTP，HTTPS 的安全防范性能更高，增加了坏人的做恶难度。

百度张鑫

# 基于 LBS 云的海量位置数据存储和检索实现

作者 薛梁

由百度开发者中心、百度地图开放平台和百度车联网联合举办的第 62 期《基于地图开放技术的车联网实践》百度沙龙上，来自百度开放平台 LBS 云方向的技术负责人张鑫，分享了题为《基于 LBS 云的海量位置数据存储和检索实现》的演讲，介绍了 LBS 云平台的整体架构，以及在整个发展过程中的优化历程。

百度地图开放平台主要是为地图开发者提供一系列的系统解决方案。目前这些解决方案主要是涵盖六个方面：定位、地图、数据、出行、轨迹和分析。每一个解决方案都有一系列的服务 SDK 和 API，并打包成一个完整的产品来对外提供。开发者也可以根据自己的需求使用集中的服务或者 SDK 进行开发。在数据处理方面的服务主要还是 POI 检索，Place 检索是对百度地图使用的 POI 检索服务做了一层封装。

LBS 云主要是让开发者上传个性化数据，并支持对这些数据的存储及检索。GC 是地理编码引擎，接收一个文本的地址串后解析成百度地址库中对应的地址。还有经纬度信息，RGC 和 GC 恰恰相反，当 RGC 接收到一个 XYZ 经纬度坐标之后就会把这个经纬度坐标转换成一个语义上可以解读地址，也就是一个文本串，或

者是一个结构化的地址。

目前 LBS 云所支撑的业务场景可以通过三个相对典型的案例来详述：

- 首先是物流行业。包裹分发到一个网点，不再需要相应的作业人员看包裹上的地址来进行分发，而是把业务员负责的片区数据存储到云平台上，这个包裹的地址用LBS云的GC转换到一个XY坐标，通过检索系统确定这个坐标所对应的服务片区。
- 其次是酒店。通过百度地图酒店数据来定位选择合适的酒店，这些数据目前就存储在LBS云平台上，对酒店周边服务进行检索都是通过LBS云平台实现的。
- 最后是交友。找附近的人估计是很多社交类开发者比较共性的诉求，LBS云平台系统支撑开发者把这样的数据信息上传到云存储，实现查找附近人的功能。

## LBS云系统的特点

这里可以介绍一下百度 LBS 云系统比较突出的 6 个特点：

- 高可用性： LBS云系统存储的可用性可以达到4个9，检索可用性在5个9以下，可用性的标准是根据系统从收到一个请求到给出应答这一过程的耗时不超过两秒，这是可用性的保准。
- 容量大： LBS云系统存储现在为每个开发者都提供了大概是千万级别的存储，空间可以达到数百G，而且容量还可以根据需求来申请扩大。
- 灵活性强： 系统可以存储各种各样的POI数据，可以根据用户想要上传的字段来灵活安排，这些字段是否参与检索、是否参与排序都可以自定义的。关键是这些字段的属性也是可以修改的，所以灵活性非常强。
- 实时： 开发者将数据上传到LBS云平台，其存储发布到检索是在秒级之间，能够达到及时更新。
- 高效的空间索引： 从前面分享的场景里可以看出，LBS云系统所支撑的开发者大部分在做周边检索，空间索引算法采取的是GeoHash算法，耗时不到7个毫秒，非常高效。



- 数据安全：最后一点大家都比较关注，放到LBS云平台上的数据安全吗？从两个纬度来讲：第一，用户之间的数据是互相隔离的，每个用户都有自己的AK，你在注册成为我们开发者的时候有自己的AK，一个用户的数据只能被该用户使用。第二：开发者上传的数据有多个副本存储在平台上，不会出现数据丢失的问题。

## LBS云系统架构

既然这个系统有如此多的特点，那么它的架构又是怎样的，如何支撑这些特点？

图1 左边部分是系统监控，LBS 云系统里的每一个子模块、子服务都有精密严格的系统监控，这个监控包含了对机器资源本身的监控，CPU、磁盘都有监控。另外对 QPS、耗时、SLA 的监控都是非常完备的，一旦出现任何问题，系统都会灵敏及时的报警。（图1）右边主要包含三部分，一个是针对检索结果和云分析结果的云展示，做数据可视化的展示。另一个是数据管理平台，用户可以通过这个平台对上传的数据进行便捷的编辑、修改。（图1）中间区域是整个架构的核心——控制服务层，主要功能是负责并发控制、配额控制以及集群。通过对服务

## 系统架构

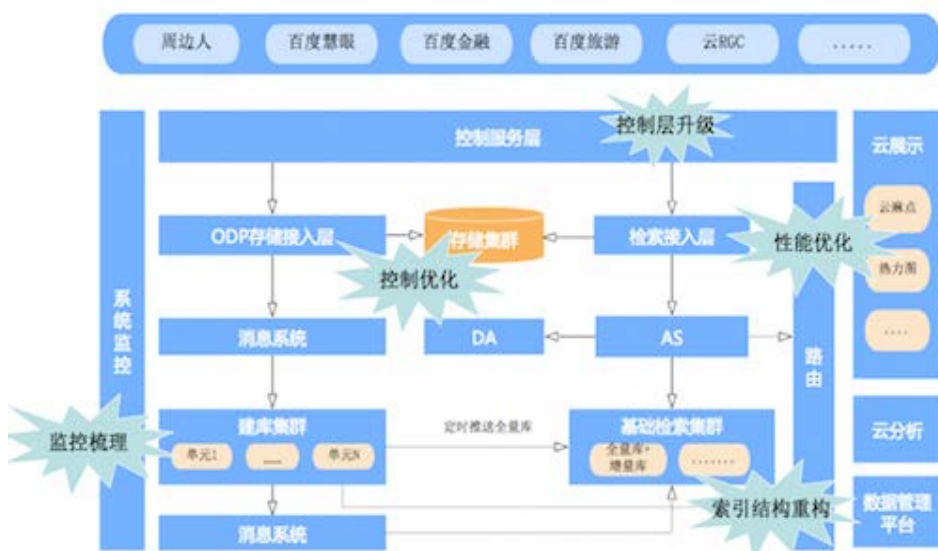


图1

请求的控制，将存储请求分发到检索层，存储请求会通过存储检索层更新到存储集群当中。如果将这些数据发布到检索，是可以通过系统推送到检索端。而如果是从检索端进来的请求，首先会进行索引查询，查询完之后再去找存储集群中查摘要，把索引跟摘要的信息做一个整合反馈给用户，这就是检索接入层。

AS 是高级检索单元，主要负责把查询传给 DA，做一些语义理解，把理解的结果发到基础的检索集群当中。基础检索集群主要负责真正的倒排检索，基础检索集群里的库分为两部分，一是全量库，全量库静态不可修改。另一个是增量库，实时发布的内容会更新到增量库。通过实时查检索全量库和增量库，并将结果反馈给 AS。最后介绍的建库集群有两个功能：一是把增量信息推送到基础检索集群，其次是及时构建全量库，替换基础检索集群库。

LBS 云系统架构在经过全面而细致的调研之后，对各个风险点进行了分析，随后做了如下的一些改进：

- 对基础检索索引进行了一个重构。
- 对检索性能做了全面的优化。
- 对 ODP 存储接入层做了控制优化。
- 全面升级控制服务层。
- 对监控的梳理。
- LBS 云索引结构演进

通过图 2 可以看出，云存储系统可以保证用户数据互相隔离，A 用户不能访问 B 用户的数据，云检索系统中基础检索单元原先的全量库设计是：将所有用户数据的索引混建在一起，取得倒排拉链后，再根据用户的唯一标识 `user_id` 做过滤，这种设计会使得某些用户的检索会被其他用户的数据增长影响而导致性能变差，最严重的是会导致很多用户数据会由于拉链过长截断的问题而一直没有机会返回，当用户量小数据量少的时候问题并不明显，而随着接入 LBS 云的用户越来越多，原先这种索引之间没有彻底隔离的设计对用户体验造成了极大的伤害。

针对这个问题重新设计了 LBS 全量库的倒排结构，设计目标：

- 全面隔离用户，用户之间互不干扰，有各自独立的索引区间。



## 索引结构演进

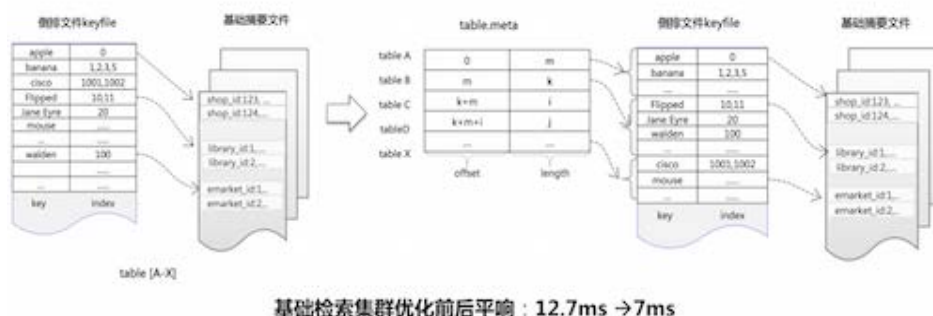


图 2

- 重构后性能至少要比现状好，解决随着用户增长，系统性能不合理地变差。

具体方案：

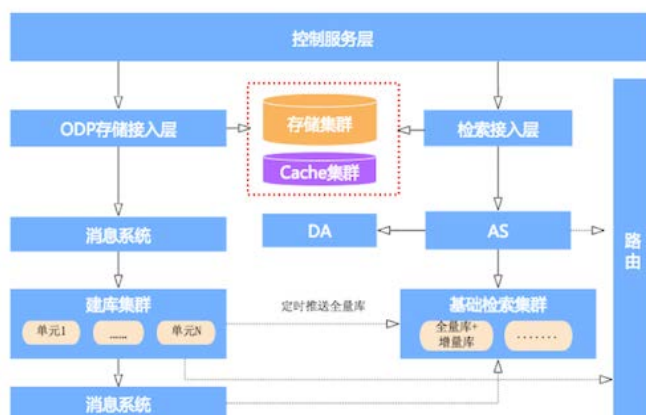
- 重新设计了全量库索引结构，提供了支持区间有序，允许全局有重复key的基础索引文件结构，提供了区间二分检索的查询。
- 采用新的索引结构的同时，设计了Table.meta二级索引结构，用于维护每个用户对应的索引区间，通过记录(起始位置，长度)信息来界定用户自己的索引区间，打破了原来用户索引杂糅无法分离的设计。
- 图2中table ABC...的索引内容完全分离，互不影响。基础检索集群的平响从12.7毫秒降到了7毫秒，两类PV量最大的空间检索平响超过100ms的长尾比例从原来的2.82%，12%减少到1.58%，5.89%，也彻底解决了用户检索效果上的问题。

## 存储演进

对存储以及监护层性能优化主要分为三点，第一点是把检索监护层从 PHP 改成了 C++，因为存储集群搭配的 PHPdriver 本身性能是非常差的。第二点是对 driver 进行升级。经过测试验证，连接池比短连接性能稳定非常多。第三点改进比较重要，引入 cache 集群来提升 table 层的性能。

关于提升 table 层的性能有两点需要说明，第一点性能方面我们如何保障的？（见图 3）

## 存储演进



存储+检索接入层优化前后平响：150ms→10ms

图3

首先 Cache 集群采用了 Redis 做了存储接入层。其次就是将 Cache 存储结构里的用户信息缩减。还有就是把存储集群当中的数据加载到 Cache 集群中，而加载操作不是由检索监护层完成的，检索监护层发一条消息给消息系统，由另外一个模块完成这种存储集群到 Cache 集群的更新。

第二点是如何有序的更新，以及保证最终一致性。假设现在在更新数据，同一时刻不同的 POI 数据也在更新，要解决有序性这个问题。方法就是去掉并发，不同的数据可以并发，相同的数据有序更新。

## 存储控制优化

在存储控制优化方面也做了很多改善，V1.0 版本只支持非批量操作，单套 POI 数据实时增删改。V2.0 版本引入批量操作，支持用户对自己的表进行批量上传、删除，以及发布检索等批量任务。引入批量操作对系统造成了很大压力，因为部分用户会上传超过严格配额的数据。优化方法就是增加配额管理，减少数据过载现象。（见图4）

任务调度系统也是优化的重点，此前，只要用户池里有任务就会被提取出来发到后端消息系统，更新到建户集群当中，这样的做法对后端建户集群造成很大压力，消息系统有很多消息堆积。解决方法就是根据后端建库系统性能对任务做一些调整，使得不会有过量任务冲击到后端建库集群，保证一些实时单条数据更

# 存储控制优化

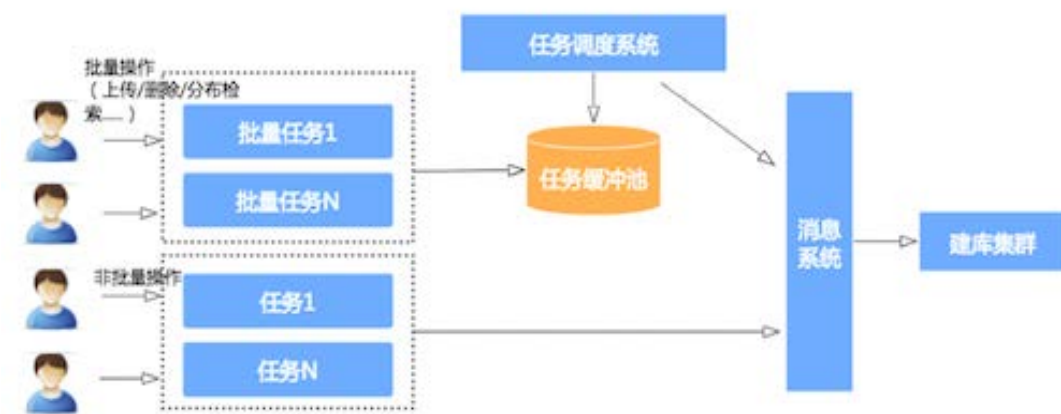


图 4

新不受太大干扰。

控制服务层升级是非常有意义的事情，将之前控制的服务层内部 RPC 组升级到最新最好用的组建，组建升级之后就使得检索的可用性提升了近一个 9。HTTPS 解决。但是相对于 HTTP，HTTPS 的安全防范性能更高，增加了坏人的做恶难度。

张鑫是百度高级研发工程师，目前是百度开放平台 LBS 云方向的技术负责人，主要负责 LBS 云存储和云检索的架构优化和 RGC（逆地理编码）的系统架构优化产品布局，主导了百度 RGC 效果性能优化项目，项目获得百度创新奖。重新设计优化了云存储基础检索集，摘要检索的架构。

# 百度贾海禄为你揭秘 室内定位技术的神秘面纱

作者 薛梁

地图软件已经成为我们日常出行必不可少的工具之一，在导航和搜索上带来了很多便捷的帮助。而定位技术在地图软件里的应用可谓无法替代，尤其是在小范围的室内定位环境下，如何帮助用户高精度搜索导航，是百度定位技术的研究重点。

由百度开发者中心、百度地图开放平台和百度车联网联合举办的第 62 期“基于地图开放技术的车联网实践”技术沙龙上，来自百度定位组的资深软件工程师贾海禄，分享了《室内定位技术揭秘》的主题演讲，介绍了在使用地图导航和搜索过程中都用到了哪些不为人知的定位技术，以及如何克服定位技术本身的技术缺陷。

首先，贾老师举了一个例子，通过百度地图的室内导航技术帮助用户定位导航到位于购物商城五层的商店。由于该商城使用的是地磁定位，效果比 Wi-Fi 好，定位精度在 3 米以下，所以根据地图给出的导航路径很快乘坐电梯到达了目的地。

在这一过程当中用到的室内定位技术在哪些领域还有用武之地？其实在其他场景的应用很多，像商场、机场定位、交通枢纽定位导航、停车寻车、店铺查找、团购活动推送、用户画像挖掘等领域都很有价值。

关于室内定位估计很多人会存在一个误区，那就是认为手机上的 GPS 能实现一切定位技术。答案是否定的。虽然 GPS 精确度比较高，也支持连续定位技术，但是缺点是功耗太高，且初次打开 GPS 时所用的定位时长较长。而对于室内定位来说，GPS 最大的不足就是室内不可用，因为信号覆盖不到室内，地下停车场更不用说了。

## 室内定位常用的这些方法是什么呢？

基于 LED 灯具和手机的搭配可以进行定位。经过编码的特殊 LED 灯具通过电路控制设备，能够被手机前置摄像头感知 LED 灯本身的阴暗变化，从而定位具体位置。（如下图所示）比较常用的定位技术主要还是基于 Wi-Fi 和蓝牙，原因之一就是成本较低，部署简单，便于维护。

## 常用室内定位方法

方法	精度	成本	设备依赖	技术优势	技术缺陷
VLC(LiFi)	1M内	高	LED灯具等+手机	抗干扰性强、成本一般	新技术、国外推广阶段
红外线	5~30M	较高	红外接收+发设备	精度一般	易受墙体等阻隔，传输距离短
超声波	10CM	较高	测距器+标签	功耗小、精度一般、抗干扰性强	多径效应、非视距传播影响大
超宽带	6~10CM	较高	UWB接收器+标签	穿透力强、功耗低、抗多径效果好，安全操作性强	造价高
蓝牙	10CM	中	蓝牙设备+手机	体积小	传输收视距影响、稳定性稍差、受噪声信号干扰大
Wi-Fi	5~20M	低	无线路由器+手机	普及度高，精度较低	易受干扰

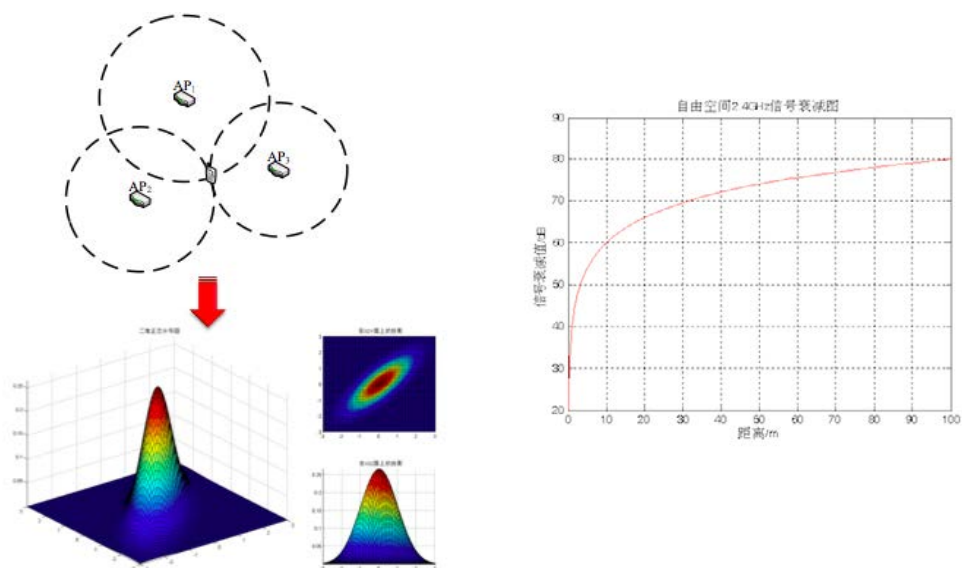
那么在没有 GPS 同时要减少部署的情况下，如何实现室内定位呢？当然是靠手机获取基站信号，能够正常通话的手机都能检测到基站信号。其次就是通过 Wi-Fi、蓝牙和地磁，基本上每个建筑都会有地磁信号，并不需要特殊的发射源。

信号与位置之间的相关性：

- 离无线设备越远，信号越弱。
- 不同位置扫描到不同的Wi-Fi、蓝牙等信号。
- 设备固定、各处信号强度相对稳定。

通过以上三个因素，就可以制定出可行的技术方向，那就是基于室内无线电信号构建出一套由无线电信号到 XYZ 映射的技术，即室内定位技术。基于这样的思路制定出最简单的一种定位方法——三角定位法。当空间有三个真实的位置 AP ( $X_1Y_1$ ,  $X_2Y_2$ ,  $X_3Y_3$ )，处于其中的手机就能检测出 AP 信号强度，最终得到一个几何位置。这种方法的优点是模型简单，容易实现；缺点是 AP 信号辐射分布刻画不足，精度不高。

但即使是在这样一个简单的定位模型下，还是会存在很多需要解决的技术问题。比如，随着和 AP 位置的距离远近导致信号强度逐渐递减。当然，因为手机不同的制造工艺，不同的芯片，不同的机壳材质，都会导致无线电信号 RSSI 感知的差异，呈现出非线性的规律。（见下图）



如果说三角定位法的模型效果不好，那如何去解决这些问题呢？下面来介绍几种使用场景较多的定位方法。

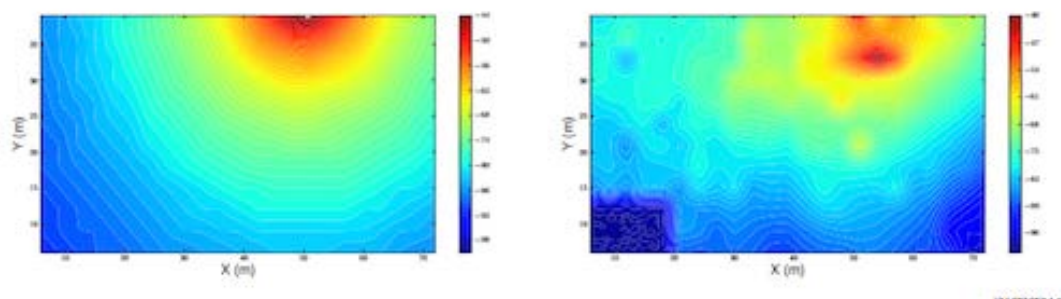




## 指纹法

指纹即对空间中各个点的各种信号采集数据，数据代表了复杂模型。通过合适的指纹相似度计算公式，即各种距离计算公式来计算实质是 KNN，求得 TopN 相似指纹的加权中心点。

对真实的 AP 空间每个点的信号强度进行采集，检测每个点上 AP 辐射的强度，在服务器当中对噪音做一些校正处理，最终把这些数据输入指纹库，而这个指纹库就是根据真实空间分布的一些数据组成的。AP 室内辐射分布如下图所示。

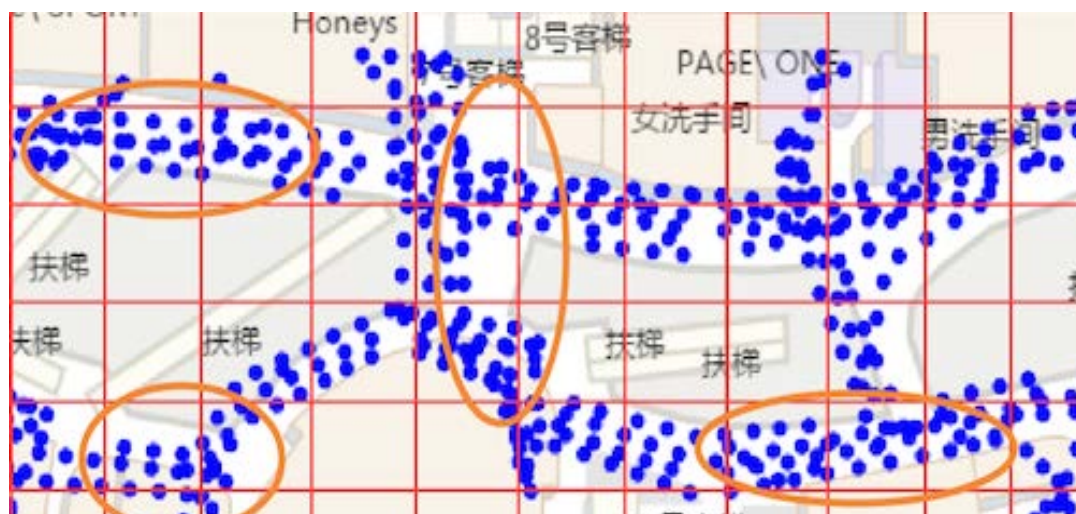


指纹法的特点是实现简单，精度尚可，但是要将所有数据和样本进行比对计算，消耗存储空间，计算量非常大。

## 局部模型法

局部模型法的原理就是对室内空间进行小网格划分。针对每个 AP 在网格空间进行最小数量 + 最小误差的模型训练，得到 AP 多个覆盖区域的简单模型。计



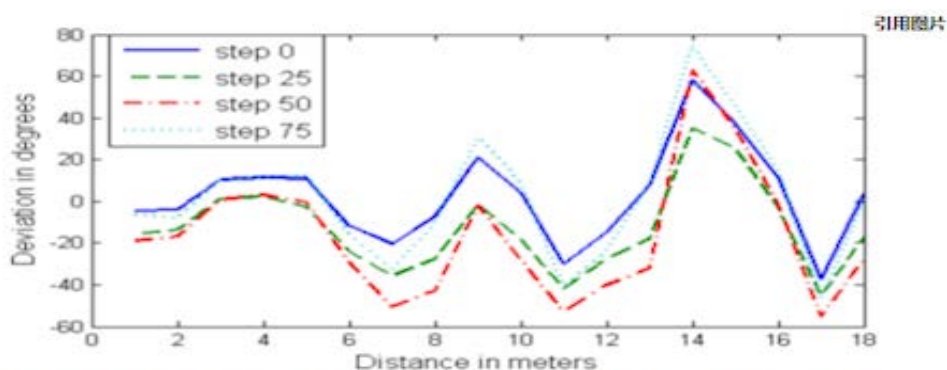


算时先识别定位网格区域，再利用该区域 AP 的精细模型实行进一步位置推算。

这种模型的典型特点就是，训练过程相对复杂，定位精度逼近于指纹法，具备一定的鲁棒性，而且存储需求较小。

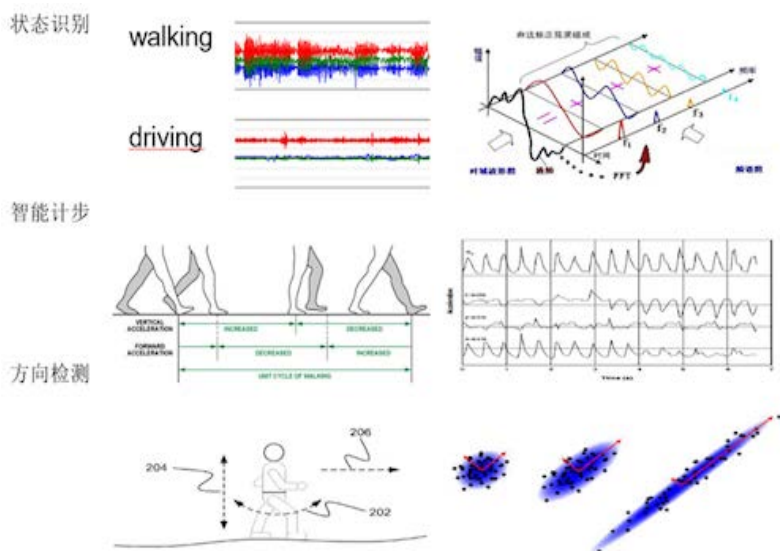
## 地磁定位法

地磁信号来源于地球本身，因为地心引力，地磁场本身存在。钢筋结构和地板等物体都会对地磁产生影响。地磁强度由于建筑结构不发生变化，基本上还是比较稳定的。在一段时间内磁场强度变化本身就是一种指纹。通过计算定位地磁序列指纹与指纹库中的子序列相似度来得到匹配位置。相当于一个模糊子序列查找，实际上因为搜索空间巨大，会结合粒子滤波提供匹配性能，最终定位效果达到 3 米以下，这是专业测试达到的效果。



**b. Four heading directions measured at four steps.**

地磁定位法的特点是实现相对复杂，需要做地磁校正，但是精度效果优秀。



## 连续定位法

室内定位应用场景还有连续使用的情况，地图 App 在打开的时候是持续收集信号的，持续的感知 Wi-Fi 和电磁信号的变化，以及各种传感器的变化。这就需要利用多种传感器在较短时间段内的感知变化，并识别出模式的改变，提高连续定位的体验。

这里不得不谈到传感器技术，而这些传感技术主要有两个途径：

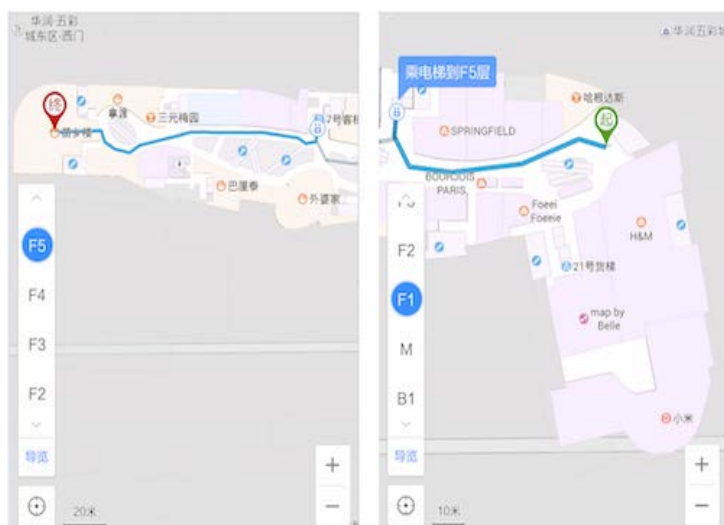
- 状态识别。智能计步是通过手机对加速度传感器的处理，根据人行走的快慢来感知加速度数据，判断用户是在移步，还是在快速前行。
- 楼层检测。根据加速度传感器收集数据，来判断用户是在一层里面走，是不是上下梯，扶梯或者升降梯。

## 实践操作

在百度地图的移动端 App 上，通过室内定位技术来实现精准定位的几个简单操作如下。

## 写在最后

在介绍了众多室内定位技术之后，技术交流又回到了车联网领域应用这个问题上。对于现实中驾车用户来说，尤其是在陌生城市或者商场车库里，定位导航



的功能尤为重要，未来的车联网导航应用中，会直接展示地下车库车位分布图，然后给出高精度的定位点，用户下车后随身携带的百度地图将会自动检测记录停车点，等用户逛完商场后可直接发起到停车点的导航，帮助用户快捷的实现停车找车目的。

**贾海禄，百度资深软件工程师，目前负责室内高精度定位技术方向的研发工作，专注于利用机器学习技术持续优化室内定位服务，对惯性传感器、WIFI 定位、蓝牙定位、地磁定位等相关技术有深刻研究，快速推动了室内高精度定位技术在百度地图产品上的应用。**

# 从 IT 科技公司和传统车企看 车联网发展现状

作者 薛梁

由百度开发者中心和百度车联网联合举办的第 62 期“智能语音和大数据护航车联网解决方案”技术沙龙上，来自百度车联网 OEM 解决方案 CoDriver 和 MyCar 的项目负责人刘凯和栾宁给大家详细讲解了百度是如何帮助车主提供基于智能语音识别技术和大数据分析技术的行车解决方案的。同时，来自传统车企长安汽车的车联网平台负责人蔡春茂站在传统车企的角度，分析了他们的切入点以及在车联网方面的发展成果。

## 解放双手的语音副驾 CoDriver

刘凯说，在目前的车载场景下，其实有很多痛点需要解决的。因为生活场景和信息服务方式在不断的发生变化，驾驶者获取信息和获取服务的需求不能很好地被满足。这时候就需要通过智能语音助手来满足驾驶需求，带来更好的驾驶体验。在保证驾驶者的安全问题前提下，CoDriver 真的能像副驾驶一样，帮助驾



图 1

驶者分担查询信息、获取服务、导航、打电话、听音乐等需求，真正地帮助驾驶者解放双手。

下面我们就一起来看看 CoDriver 到底具备哪些核心技术（见图 1）

CoDriver 目前需要通过像“小杜你好”这样的词来语音唤醒它，通过语音技术的识别把用户说出的话识别成文字文本，下一步要做的就是解析成一个 CoDriver 可以理解的驾驶者意图的结果，这个过程其实就是一个分词的过程。

语音合成，目前百度在语音合成部分积累了包含多音字的合成以及多种有感情语音的合成，根据用户需求也推出了不同的可定制化的明星语音 GPS 播报，主要用于导航过程中给用户比较好的体验。

多轮对话技术，这也是归属到语音解析的自然语言处理范围之内，刘凯老师举例说：“明天北京飞上海的航班”，CoDriver 可以结合上下文的理解，选的是国航的北京飞上海的航班，结合上下文的理解作出判断。目前多轮对话技术已经是在 CoDriver 的 POI 检索以及多轮打电话的联系人确认的过程中应用了。

其次就是降噪算法，在车载场景下的一些噪声处理，有专门的软件降噪算法以及硬件的麦克风阵列的降噪算法（见图 2）。

麦克风阵列是将两个麦克风的信号耦合为一个信号。亦即在两个麦克风的正前方形成一个接收区域，进而削减两个麦克风侧向的收音效果。语音增强是指当



# 麦克风阵列

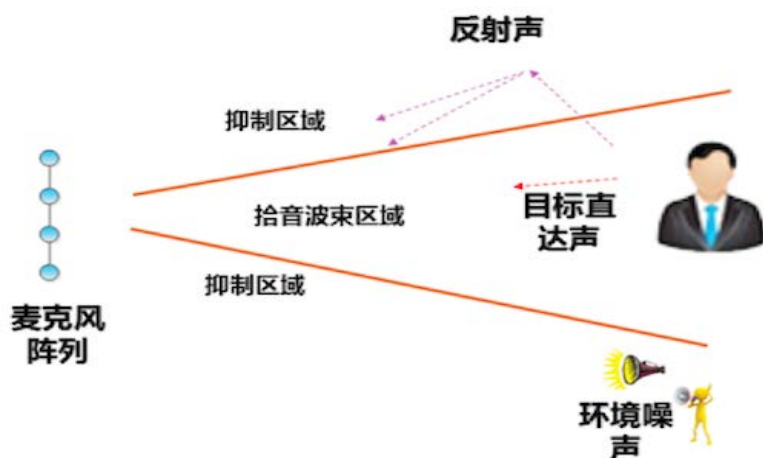


图 2

语音信号被各种各样的噪声（包括语音）干扰甚至淹没后，从含噪声的语音信号中提取出纯净语音的过程。所以在车载嘈杂环境下，也能准确识别语音指令。

所以麦克风阵列技术不限制说话人的运动，不需要移动位置以改变其接收方向，具有灵活的波束控制、较高的空间分辨率、高的信号增益与较强的抗干扰能力等特点，因而成为智能语音处理系统中捕捉说话人语音的重要手段。

功能方面，有一些最核心的几个需求，例如地图导航、电话、短信的收发需求，生活信息的查询等等，这些都可以通过百度强大的搜索能力来完成。对于车主服务，比如洗车、加油、美容等服务，结合车后服务和车主服务来实现这些功能是非常容易的事情。所以通过语音、语义、智能反馈的技术叠加，再加上百度的云服务以及视觉技术和百度大数据技术，百度 CoDriver 的能力以后会越来越丰富。

## MyCar在大数据处理方面的核心技术

首先，百度 MyCar 的项目负责人栾宁给大家分享了 MyCar 究竟为何物：百度 MyCar 是集车主服务、车后服务和云计算数据的私有云服务平台，为车提供个性化定制、安全管理的解决方案，为每一辆联网车辆装上“百度大脑”，提高智能用车体验。

那么百度为什么要做车联网这件事？首先是基于百度地图积累了很多的路况

数据，其次，百度已经为超过六千万的车主提供了服务，在车主数据以及 B2B 上有着得天独厚的优势。MyCar 希望利用百度大数据处理优势，进一步分析数据，给车主提供优质的服务。

下面我们来看一下 MyCar 都有哪些功能，以及它的核心技术在哪方面！

首先，车上的数据可以传输到云端，用户可以通过手机 APP 远程查看操控车辆，以及获取车辆的相关服务，包括位置服务、智能体系、车辆管家、数据分析或者是托管服务。

## MyCar提供的实时获取车辆位置有什么好处？

可以远程地查找车辆，包括室外停车场以及地下停车场。

车况的展现，包括油耗分析、续航里程、百公里油耗以及车门车窗的开关情况以及车况的自检。

远程控制，包括远程开关车门车窗、后备厢、天窗等等。

除了数据的展现之外，MyCar 还做了进一步的简单化数据分析，包括油耗分析，行驶轨迹分析，这些数据对于建立车智能模型和大数据分析起到了很好的铺垫作用。MyCar 更多的是基于车辆数据的采集，数据的传输，以及服务打通和安全性方面提供优质服务的。

下面介绍一下 MyCar 所使用的两项核心技术。

第一个是鹰眼。它是一个位置数据采集、上传、存储的技术。首先它包含了室内定位，室外定位，轨迹和回传。

第二个是数据处理器的存储和检索服务。百度服务器的稳定性非常高，还可以实现实时的轨迹检索和历史服务。

第三个是慧眼。百度慧眼是基于百度地图 GPS 数据分析的工具。它可以基于人的属性、兴趣爱好、消费场所等因素制作用户画像（见图 3）。

最后一点，栾宁也谈到了安全方面的技术问题，如果数据不安全，那么一切都是空谈（见图 4）。

栾宁从系统层、应用层和云端技术三个角度介绍了 MyCar 的安全性。

- DDoS防护：自动清晰攻击流量防御SYN Flood，UDP Flood，CC等DDoS攻



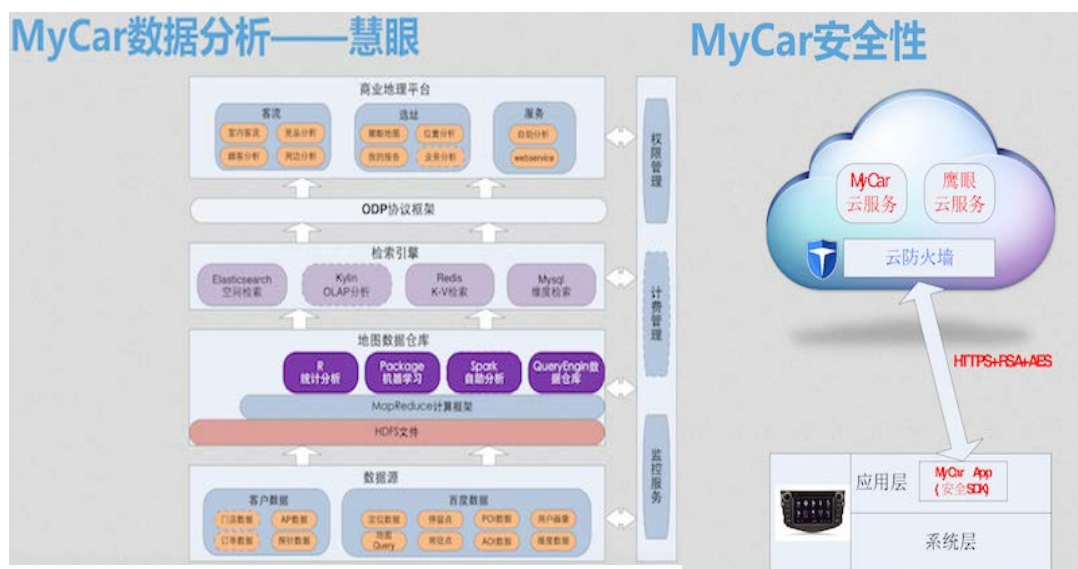


图3 图4

击。

- 云服务器防护：自动拦截密码暴力破解，实时检测通知异常登录。
- Web漏洞检测：智能扫描云服务器上网站可能存在的SQL注入，CSRF、XSS等漏洞。
- 端口安全检测：定期扫描服务器上的开放端口，通告风险端口，降低服务器被入侵风险。

## 长安汽车车联网平台在建心得

蔡春茂是长安汽车车联网平台的项目经理兼技术负责人，他主要是从整车厂商的角度阐述传统车企在面对车联网大浪潮时的困惑、反思、平台建设思路、设计理念、技术架构、以及和互联网公司合作的案例心得。

首先，蔡春茂站在车主的角度分析了现在汽车后端服务行业里的几个困惑：用户续费意愿低、服务体验差、产品没有粘性、功能定位方向错误、厂商无数据、创新慢。车联网仍在探索，参与环节和厂商越发繁多、庞杂，各利益主体之间纠葛随之增多。但即使是这样，不管是前装车厂还是后装车厂，都在搭建自己的平台。而传统车企在这一方面所遇到的困难相对较多，没有标准，数据少，整车是比较封闭的，包括车主也不愿意把数据提交给车厂。其次是车上功能很复杂，操



图 5

作麻烦，体验不一致。

那么长安汽车是怎么做的呢？探索商业模式创新，构建安全、绿色、便捷的出行服务整体解决方案，即长安自己的车联网平台，以此不断提升产品价值。具体措施是自建+合作。整体思路是走“轻平台”思路：核心自建、数据自有、服务合作的原则。详细流程见图 5。

这里，我们重点来看一下云端的建设。云端建设更多的是保持开放性、扩展性，终端、平台、服务商解耦合，强化管控；支持多服务商，多设备的原则。长安汽车现在正在建的 TITAN 平台，大数据分析平台、服务集成平台、容器管理平台等等，都处于探索阶段。

根据长安汽车的车辆数量，设计了如下图这个高可用设计的云端架构，所有的数据都可以通过这个平台发给第三方合作伙伴，内部也是在数据存储上做服务，通过数据异步写入 HBase、HDFS，利用 Spark 进行数据处理，最后会把数据分析放到不同的 MySQL 里进行预处理（见图 6）。

在大数据分析这一块，长安汽车也属于起步阶段，通过采取场景建模、ETL 标准规则建立和预测建模等方式，制定出相似度计算、推荐算法、文本挖掘算法、机器学习等规范化成果。通过对第三方数据交易、共享数据交换、数据合作模式研究等数据整合处理手段，获得驾驶行为、群体驾驶特征、车辆工况进行分析，

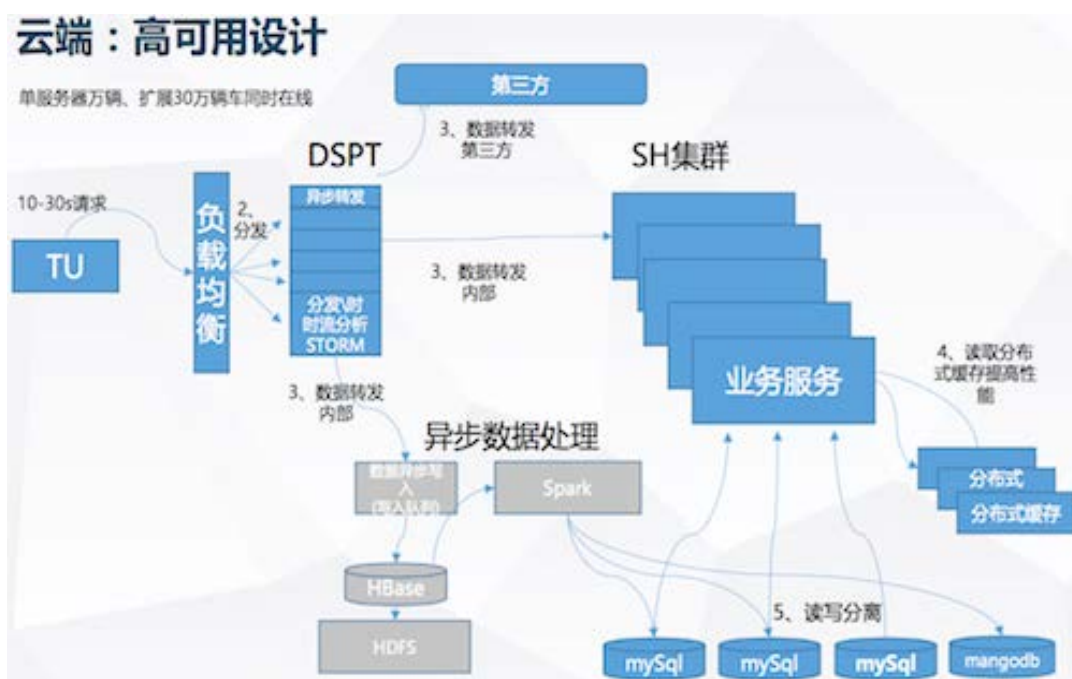


图 5

车辆零部件跟踪和故障隐患分析、车辆能耗消耗等有价值的结果。

当然，对于一个传统的正在尝试新领域的车企来说，自建是一个好的办法，但是向互联网科技公司借力也是一个不错的方法。在跟百度的合作中，百度提供了包括 MyCar、CoDriver、远程控制等方面的支持。

## 总结

最后，引用吉利集团董事长李书福的一句话作为总结：传统的汽车公司是不会那么容易被颠覆的，但汽车公司必须拥抱互联网，与互联网相依相生、携手共进才能在信息革命时代长久发展下去。

百度鲁玮

# 百度开放云助力移动游戏快速发展

作者 孟夕

随着移动互联网的快速发展与广泛普及，移动游戏已成为最火爆的移动互联网垂直领域之一，而移动游戏本身对高性能、稳定的服务器、高质量的外网带宽、快速的游戏速部署、极速弹性扩容、安全防攻击等方面有极高的要求，但是传统 IDC 中各方面的不足，很难满足移动游戏的快速发展，从而严重影响游戏的推广与营收。

由百度开发者中心、百度开放云联合举办的第 64 期“百度开放云移动游戏和直播技术解读”沙龙上，来自百度的高级产品经理鲁玮，介绍了百度开放云在移动游戏方面的整体解决方案，并就相关实际案例给出具体讲解。

## 一、移动游戏行业的发展现状

首先，鲁玮老师介绍了近年来中国移动游戏发展的现状，从 2008 年到 2015 年，大概 8 年的时间里，游戏收入从 185 亿增长到 1400 亿，从各个细分领域里的收入增速和发展趋势来看，排名前三的分别是手游、端游和页游（见图 1）。

## 近年来中国游戏行业发展规模分析

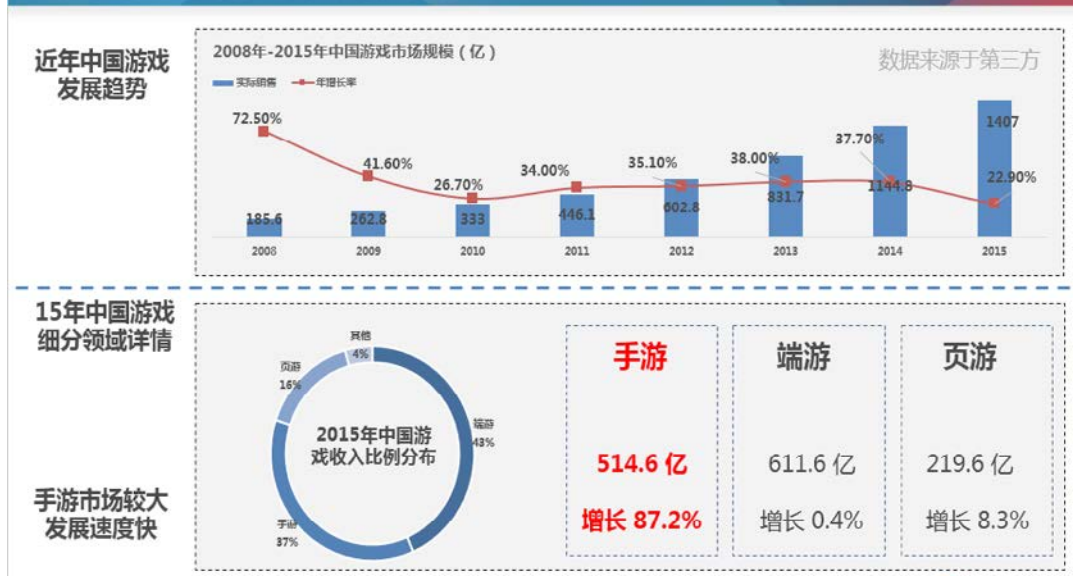


图 1

近几年来，手游经过了爆发式的增长，2015年的时候，手游的增长率达到87.2%，手游已经变成了游戏行业里面最主要的而且是发展速度最快的方向。对于移动发行和移动应用来讲，游戏一直都是最主要的垂直模块，移动游戏也已变成了移动互联网里面变现最快的领域。

## 二、手游发展趋势 – 向中重度方向发展

手游发展到今天，从刚开始的休闲游戏到卡牌、MMO、MOBA等，手游也在向中重度方向发展。归结原因，第一点是技术层面，移动处理器和GPU的快速发展会保证游戏跑得更加顺畅，重度游戏不会像原来那么卡顿；精细化的3D和情景式的游戏，对用户黏性很大，而且开发难度越来越低，这是手游往中重度发展的原因。另一个是运营层面，不能光靠游戏的长期下载量，要让用户沉浸在游戏里的时间更长，而最早出的游戏是偏休闲类的，比如斗地主、连连看，这种相对黏性较低，因为缺少互动模块，它会很容易流失掉用户，所以需要提高互动（帮会、聊天等模块）吸引用户，来产生更多的黏性。另外一个重要原因，这些玩家会付费购买道具来提升自己的级别，支付意愿较明显，所以中重度手游变现能力非常强。



## 移动游戏部署/运行时遇到的痛点



图 2

### 三、移动游戏在部署和运行过程中遇到的痛点

游戏作为一个移动行业里面变现最快的行业，部署和运行过程中肯定会遇到很多问题（见图 2）。

首先手游生命周期短，购买物理资源浪费严重。手游对于其它 APP 来讲生命周期较短，很多游戏的厂商为了支撑游戏峰值业务，购买大量的硬件，而 90% 的情况下不能利用，导致成本和收益不成正比。

第二，流量峰值无法预测，已有 IT 资源无法支撑。很多时候要进行活动大推，或者由于社会热点（影视剧、网络小说）等因素导致游戏受到关注，从而引入大量玩家。这种情况下，原来的服务集群的计算性能无法支撑新涌入的玩家，所以搞活动、大推的时候，流量无法很好的预估。原来游戏部署在 IDC 时会选用物理机，但是物理机宕机恢复时间非常长，至少需要 30 分钟，这种游戏体验对玩家来讲很难接受，它会导致用户大量流失，而且这种传统的架构系统也是基于烟囱式的单点上面部署 OS、搭建应用服务，扩展性很差。

另外，中小游戏厂商使用 IDC，本身其防攻击能力很弱，所以会经常受到攻击，严重影响游戏收入。

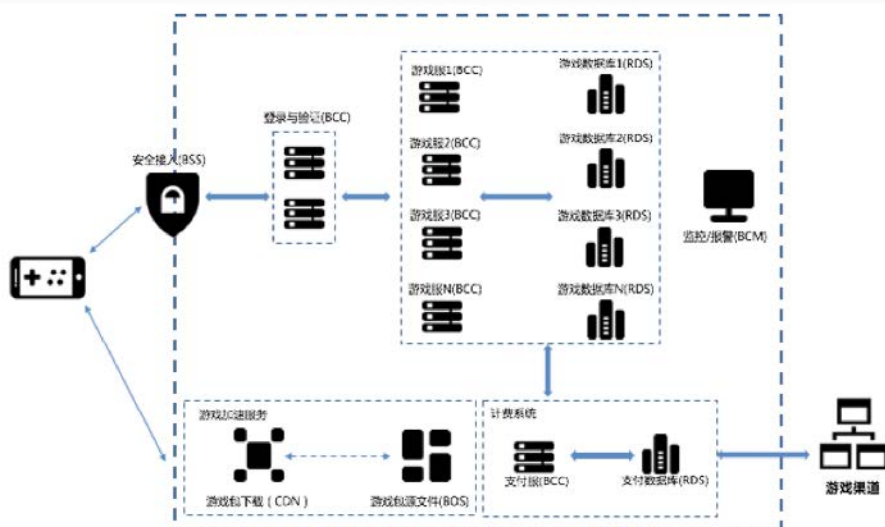


图 3

## 四、百度开放云移动游戏解决方案介绍

下面我们从百度开放云游戏客户的实际案例中，分享不同游戏类型的架构解决方案。

### 卡牌类型手游

图 3 是卡牌类游戏的解决方案，通过百度开放云高性能的对象存储（BOS）和内容分发网络（CDN）把游戏安装包推送给玩家，玩家登录游戏时，经过云安全（BSS）的严格审查。另外游戏服运行在高性能云服务器（BCC）上，每个 BCC 配置一个单独的高性能数据库（RDS），这种方案大大提高了整个游戏系统的稳定性。通过支付服连接各个游戏渠道，保证了不同渠道引入的玩家购买道具等付费行为，支付数据库也记录了交易信息，方便游戏 CP 和运营商对玩家的支付行为再进行二次分析。

### 社交类型手游

图 4 就是偏向于社交类游戏的解决方案，通过游戏加速服务系统将游戏安装包发送到手机端。玩家经过 BSS 的安全检测后，接入到登录服。因为游戏系统相对比较大，会有游戏服的资源管理、版本管理等模块。游戏大厅，类似于传统棋



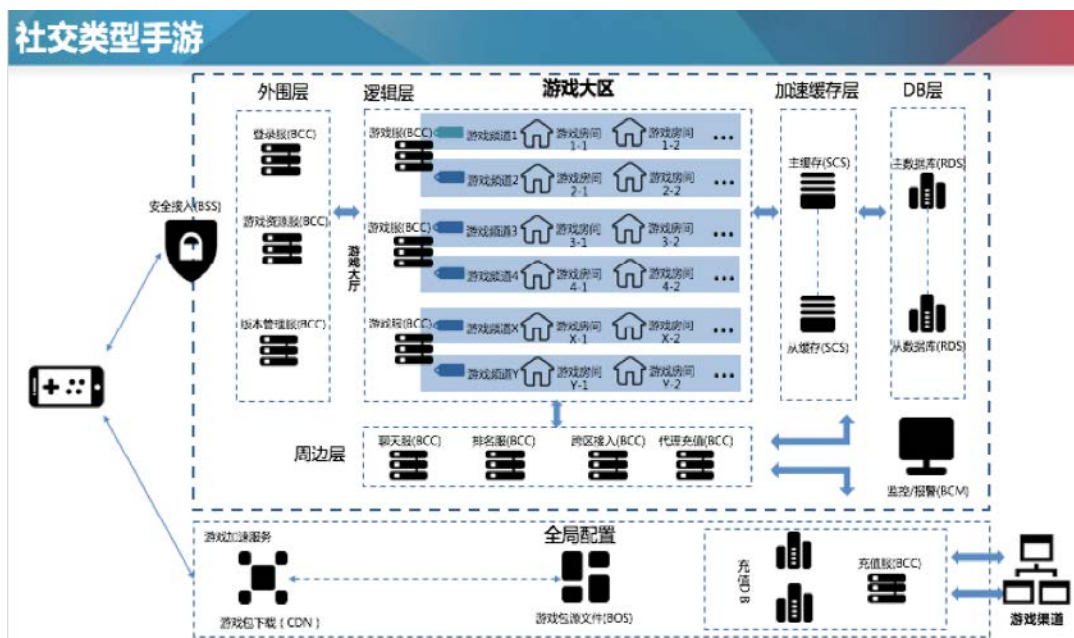


图4

牌大厅的形式，每一个游戏服里有多多个频道，每一个频道有多多个房间。这类游戏的周边系统会把各个频道里边的聊天记录保存，同时因为大家在这里要排名，那排名也会有个专门的服务系统；通过跨区接入功能，也实现了玩家跨区的游戏PK；另外为了增强游戏的流畅度，系统DB之前也会有缓存层。同时所有系统模块都会由智能云监控（BCM）来监测业务运行状况，包括通用运维指标、游戏程序等，游戏CP根据业务场景自定义设置报警规则，及时发现、处理宕机、业务性能压力过大的风险。

## MOBA类型手游

图5是MOBA类型游戏的解决方案，MOBA类型游戏对实时战斗的要求极强，但由于玩家位置不同、接入的运营商不同，导致南方和北方玩家访问质量不一致，游戏体验很差，所以百度开放云推荐此类游戏在南方和北方分别部署游戏集群。根据用户的不同位置，通过智能DNS接入最佳机房，实现用户更流畅的游戏体验；基于百度雄厚的基础网络资源，使用专线打通北京和广州区域机房，实现数据高速的同步和备份；另外游戏CP使用百度开放云业界领先的大数据服务，实现对玩家日志信息、玩家支付信息的精细分析，为游戏精细化运营提供宝贵数据。

## MOBA类型手游



## 客户案例-盛大游戏-沙巴克传奇



### 客户的需求

1. 游戏网卡高PPS转发
2. 数据中心内网线路高速、稳定
3. 高性能服务器、数据库
4. 资源独享、不受其他客户业务影响
5. 外网带宽低延迟
6. 带宽计费灵活
7. 数据分析、统计



### 开放云的解决方案

1. 百度高性能内核优化技术
2. 内网全万兆的冗余链路
3. 高性能计算集群
4. 游戏专用云服务器、物理机
5. 百度自建10线+的 BGP机房
6. 共享带宽功能
7. 百度超强的大数据服务

图 5、图 6

## 五、移动游戏客户案例说明

沙巴克传奇，是今年盛大游戏重磅推出的 MMORPG 类型的手游，沙巴克传奇对服务器性能、网络安全的要求非常高。盛大游戏不仅需要高 PPS 转发的服务器、数据中心内网之间高速连通，还要求资源独享、灵活计费和完整的数据分析等能

力。百度基于高性能的内核优化技术、自建的高质量 BGP 带宽、高性能的计算集群系统、超强的大数据服务，满足了盛大游戏对业务系统的严格要求（见图 6）。

在这个游戏里，不管前期测试、大推阶段还是后期运行管理，我们做了很多事情。沙巴克传奇封测阶段，我们给了很多支持与建议，包括定向内核优化、高可靠和弹性架构方案的推荐、结合上线节奏分配资源专区等。游戏大推阶段，我们提供超高性能的抗 D 服务、网络优化型云服务器、7\*24 小时专人运维、游戏驻场支持、全国网络质量的实时监控和预测。后期运维阶段，我们也会定期的安全巡检，给出资源的生命周期管理的建议等。百度开放云提供了一整套完美的游戏解决方案，也帮助了沙巴克传奇实现了平稳的上线和火爆的大推，满足了沙巴克传奇游戏对性能、稳定性、安全等方面的高要求。

**鲁玮，百度高级产品经理，2015 年加入百度云计算事业部，现作为百度开放云核心产品“计算与网络产品线”的产品负责人，领导、推动了云服务器、专属机、虚拟私有网络、VPN 服务、专线、弹性 IP 的产品化，实现了百度开放云相关产品从无到有，从有到优的明显提升。**

# 百度语音识别和唤醒技术解析

作者 孟夕

随着人工智能的普及，语音已成为了重要的交互方式，尤其是百度语音识别和语音唤醒技术一经推出，便受到了开发者的广泛关注。

由百度开发者中心和 InfoQ 联合举办的第 65 期“百度语音识别和语音唤醒技术解析及实践”沙龙上，百度语音开放平台资深产品经理何荡，分享了百度语音技术最新进展及最新解决方案。同时，百度语音开放平台资深研发工程师魏力凯和唐立亮，也分别介绍了百度语音识别&百度语音唤醒技术细节以及具体实践。最后还设置了演示分享环节，以便与开发者更好地互动。

## 百度语音开放技术最新进展及最新解决方案。

首先，何荡介绍了百度语音技术的最新概况，尤其提到了语音识别和语音合成，并在后面的语音最新效果中进行了展示。语音识别方面，机器识别技术已超过人；语音合成方面，主要的是情感语音合成，即通过大数据拼接的技术，把情感的因素加到声音里面，更合理地说是在声音的采集以及最后合成的效果上，都

## 百度语音技术能力

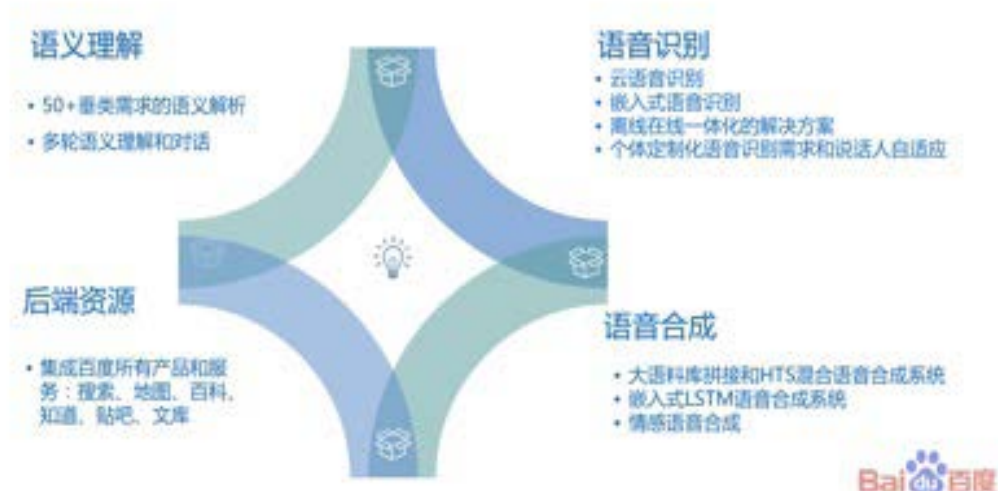


图 1

已经注入了情感（见图 1）。

另外，何荡还提到了百度语音平台的开放计划。

- 第一个是远场识别，计划在今年年底的时候，把远场技术开放出来。这技术也是免费开放的，让大家可以去做一些解放双手的应用出来。
- 第二个是情感语音，把情感的因素加到声音里面，会代替原有机械的声音，比较接近真人的声音，这也是预计在年底的时候开放。
- 第三，被美国麻省理工评为2016年10大突破技术的Deep Speech，今年年底的时候会有更大的技术提升和优化，会有更新版本的Deep Speech出来，放到语音平台上。

## 个性化语音识别 – 离线命令词识别和自定义语义

来自百度语音开放平台的资深研发工程师魏力凯，目前负责百度开放平台的离在线语音、一体化唤醒和自定义语义等技术。他的分享主要分为以下四个部分：

1. 在线自定义
2. 离线自定义
3. 自定义语义

#### 4. 语法编辑器

在线自定义，可以将不常见、不容易识别正确或者希望识别的更加准确的内容枚举到一个称为热词表的文本文件里，使得热词表里的内容能够精确识别，有了在线自定义，每一个开发者，每一个应用，每一台机器都可以有不同的识别策略；而离线自定义则提供了命令词识别的能力，这项能力使得在网络不好甚至完全没有网络的情况下，拥有高准确率的语音识别能力，比如车载环境；自定义语义则允许开发者定义想要的垂类，为了在没网的情况下也能使用，这项技术开始就是基于离线的。

新开放的这三项功能，一个解决在线识别不准确问题，一个解决了没网络的情况下不能识别的问题，而语义自定义则解决了所说的内容没法解析，或者说解析到错误的领域的问题。

最后，魏力凯还介绍了为上述新功能定制的语法编辑器，有了这款编辑器，可以使开发者更加便捷的使用上述技术。

## 百度语音唤醒技术解析及实践

百度语音唤醒内核技术有哪些，其原理和实现方式是怎样的？

唐立亮介绍说，语音唤醒技术常见的解决方案主要有：基于置信度、基于识别的唤醒系统和基于垃圾词网络的技术，百度语音唤醒技术是吸取了这三种方案的精华，即基于垃圾音素，采用模型统计的手段，使几个因素可以代表全部的发音，后边会接一个置信度系统，通过这个判断极大地降低误报率。

接着，唐立亮通过一张图片，介绍了百度语音唤醒的流程（见图2）。

首先用户要输入他们的语音，然后进行端点检测，把人说话的部分给检测出来，之后就是一个信号处理的过程，对信号进行一个非常好的噪音的处理，或者是其他方面的处理。接下来就是提取声学特征，进行识别解码，之后就是置信度判别，由于现在是一个唤醒+识别的系统，那唤醒成功之后需要送到服务器进行在线解码，最后获取到识别结果。

另外，唐立亮也提到，如何评价唤醒技术的好坏？非常重要的两个指标就是唤醒的正确率和误报率。好的唤醒技术，唤醒的正确率很高，唤醒的误报率很低。



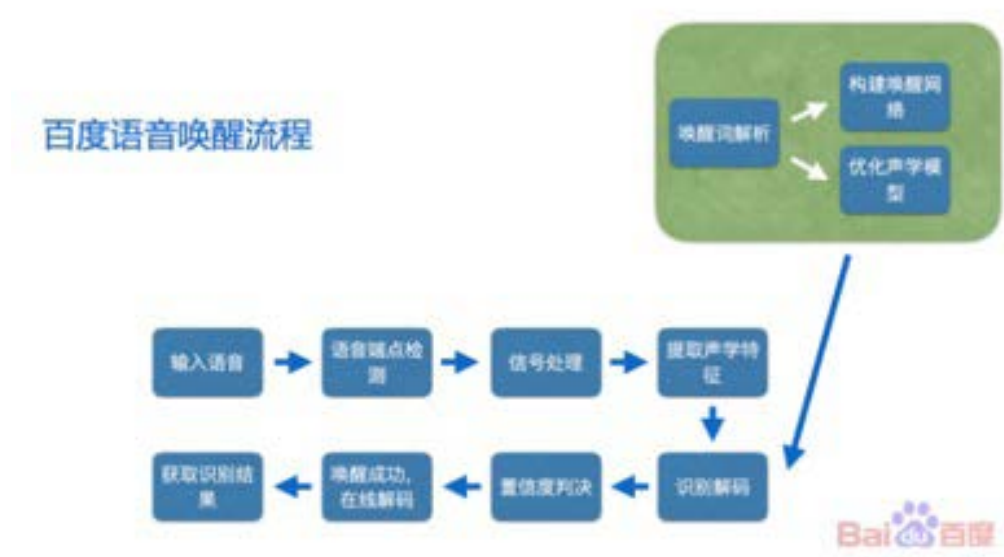


图 2

接着，唐立亮分享了百度语音唤醒的一些应用场景，包括手机 APP 替换用户常见操作、拍照、机器人、车载场景、智能家居、智能硬件等。

关于如何选择唤醒词，他也给出了以下建议：

- 唤醒词可根据应用的个性化需求来订制
- 每个词在3个字到5个汉字之间，4个字最佳
- 音节覆盖尽量多，差异大，响亮
- 建议选择不常用词语
- 唤醒词评估系统，帮助合理选择您的唤醒词

最后，唐立亮介绍说，未来规划上，会考虑开发英文唤醒、打断唤醒、常用指令唤醒和远场唤醒这些优秀的技术，这些技术完成后也会争取第一时间放到平台上给大家使用。

# 版权声明

InfoQ 中文站出品

## 百度迷你书：技术开拓新领域

©2016 极客邦控股（北京）有限公司

本书版权为极客邦控股（北京）有限公司所有，未经出版者预先的书面许可，不得以任何方式复制或抄袭本书的任何部分，本书任何部分不得用于再印刷，存储于可重复使用的系统，或者以任何方式进行电子、机械、复印和录制等形式传播。

本书提到的公司产品或者使用到的商标为产品公司所有。

如果读者要了解具体的商标和注册信息，应该联系相应的公司。

出版：极客邦控股（北京）有限公司

北京市朝阳区洛娃大厦 C 座 1607

欢迎共同参与 InfoQ 中文站的内容建设工作，包括原创投稿和翻译，请联系  
editors@cn.infoq.com。

网 址：[www.infoq.com.cn](http://www.infoq.com.cn)

TECHNOLOGY  
TO EXPLORE  
NEW  
TERRITORY

# 技术开拓 新领域

迷你书

