大家好，我是上海理工大学的郑岳久，刚刚张教授给的报告是非常精彩，这一块我们也有所接触，但是没有那么深入。我今天给大家汇报的是《云端平台的电池数优化与寿命预测》。

谢谢张教授，张教授从电化学和数据打通来做热失控，相信会对大家有所启发。如果说大家对张教授的技术有兴趣，可以和他多交流。再次感谢张教授的演讲。  
 今天同样另一位青年教授，也是研究电池的，他是上海理工大学郑岳久教授，主要研究方向为新能源汽车动力电池成组系统，2009年开始车用锂离子电池组系统的相关研究，发布论文40余篇，并且拥有多项国内外专利，今天他想和在座各位沟通关于云端平台的电池数据优化与寿命预测，有请郑岳久教授。

我本身也是做锂电池的，但是锂电池从材料到最终的应用有很长的内容，我们最主要关注的一个是电池本身，一个是到云端数据结合的桥梁作用。  
 这个是我们认为电池全寿命周期电管理与安全技术整体路线。我们下面可以看到从原材料生产到电池使用是一个整体。电池管理主要的技术包括之前要有一个分选，包括状态估计、均衡管理、寿命估计与预测、故障诊断、残值评估，以及后续回收之后再分选重组的过程。实际上云端的应用包括生命估计、故障诊断与安全预警以及残值评估，这三点是可以放在云端应用的。  
 首先是锂电池云端数据优化。前面各位专家提到了数据问题，目前还存在很大的问题，我们首先就数据问题来讲一下基本情况。 我认为目前海量数据存储面临主要的问题包括两部分。第一个是数据存储方式比较单一，第二个是数据同步性比较差。也就是说我现在的数据存储不管是什么信号，拿过来1秒一个点或者是10秒或者是30秒，都是拍脑袋的。数据同步性差体现在哪里呢？明明在放电，总电压和总电流一般同步，但是单体电压不同步，难以单体建模。如果说我们存在这些问题，那么我们面临的就是后面的数据拿过来之后什么用处都没有。

我们提出来一个方法什么呢？数据要进行分类存储。我们至少要看到，温度和SoC信号以及电压、电流在稳定期或者是充电段的信号都是稳态信号而不是动态信号。针对这种稳态信号，你用1HZ或者是10HZ的频率去存储冗余太大，而且没有必要。针对动态运行，你用10秒或者是30秒的存储频率存储，基本上就是不可用的，是失真的数据。针对低频信号我们要采用事件驱动的存储方式，也就是说当我的值发生了我想要的变化时，我才会进行一个记录。这种记录方式可以大大降低我们的存储量。同时你可以看到，失真的情况基本上很少，因为只要你把它的阈值调小，事件驱动基本上都是可以跟踪到所有低频信号改变的。这个是针对低频信号提出的一个方法，叫做事件驱动的存储。

 对于高频信号，大家担心的问题是什么呢？就是我这个数据到底用多大频率去存储，这是一直以来大家都在问的问题。我们通过实车运行，我们有一辆车用4HZ的频率去记录总电压、总电流以及单体电压的情况，我们认为这是它的原始信号，如果说我们做一个傅立叶变换做出的结果我们抓取的是最高频率德威治，我们分析之后，用还原的技术可以看到，对于我们这辆车的控制体系来说，我们实车动态信号，如果说可以做到2HZ左右的存储，基本上你的电压误差可以控制在5毫伏，在这种情况下，我们认为存储才是有意义的，对于高频信号，特别是对于我们所测试的这辆车，就是应该用2HZ以上的频率去存储。

当然频率越高，失真越小，但是你要考虑到数据量的问题。实际上你用2HZ去存储，实际上还是压力非常大的。我们的车如果说按照2HZ存，我们算了一下，单车如果说跑的话，一年需要50-100G空间，问题是你10秒，对动态工况来说是没有价值的。我们采用的方法是对于总电压总电流按照原有频率存储。比如说10Hz，对于单体电压，每10秒或30秒一个点，通过复原技术，大概的失真如果说10秒一个点，在5毫伏左右，这和你的建模精度相当，因此我们可以通过这样的方式来大大降低总体的数据量。

如果说把这三个方法结合在一起，我们通过初步计算有一个对比，你可以看到它的存储量大概是8个单位，如果对比1HZ的存储方式是238个单位，差的还是很大的。动态电压失真也比1HZ的要小。  
 第二个是数据不同步性的问题。这个问题体现在哪里呢？你的电压在往上升，但是你现在的情况在放电，这这样导致不能准确反应电压的不一致性，二是电池模型、参数辨识结果不可靠。这就导致了你存在各种误判的可能性。我们说总电压和总电流都可以保持同流，但是单体电压不可能都放在同一个帧里面传，所以必然会存在一定的延时。我们用了什么方法呢？就是主控给你一个命令说，现在有一个帧在这里，大家可以开始采集了。这样后面每个子控制器最终存储的时候，存储的数值是三角形所指的点的位置，而不是说我收到数据的位置，我采用这种方式之后，同步性得到大大改观，基本上不需要考虑存在同步性比较大的问题。

基于数据的优化情况下，我们才可以做一个云端数据的预测和估计。我们在云端上有一个最大的优势，因为我们的算力比较大，我可以把每个电池包当中的每一个东西都算出来，可以进行一个所谓的图形化的显示。我们说大数据最大的一个好处就是可视化技术。刚刚邓总给大家展现了上海市地图上的很多数据，看着很直观。同样在电池里面，你如何直观地看一个电池状态？我们提出一种新的工具，类似于一个地图，把它横坐标和纵坐标得到之后，就可以把电芯和电池组的状态给得出来，真正电池组的状态是逐渐变化的过程，这是我们实际模拟的电池的具体过程。  
 这里还可以给出：如果说我有一个均衡算法可以调整电芯电量，还可以给出我的均衡算法的可视化是怎么样的，均衡能否满足我电池总体容量可以达到最高容量的情况，这些都是我们认为的大数据平台上可以做的可视化的点。

基于目前的数据我们如何做电池的寿命估计？因为我提了几个问题，一个是动态数据很难利用，另外一个是同步性的问题。这两个问题导致了动态模型是没有办法建立的，在这样的情况下，我们只能寻求于它的充电数据，因为充电数据是稳态的，你即便有零点几秒的延时，对我们来说问题不大。现在担心的是单体的数据也是缺失的，所以针对现有数据主要是针对充电工况对整体电池包进行一个容量估计和预测。

这是我们做的一个基本预测的过程。我们知道电池寿命衰减是一个长期的过程，我们必须要考虑到它温度的变化，这个地方要加到温度的修正。同时车上的数据噪声非常大，一般的滤波是做不到那么好的结果的，那么怎么办呢？我就需要提出来控制它的噪声。通过我的模糊逻辑去控制它的噪声之后，实现容量更加精准的估计。这是我们给出的某一批100车的数据实际估计情况，蓝色的部分是第一批获取到的数据。我们做了两次验证，第一次我们依靠蓝色的数据去预测绿色的数据，蓝色的数据本身就是我们左边这个方法得出来的结果。通过蓝色的数据去预测绿色的数据，看它的重合度，从而间接证明我们这个算法是稳定的。其次，我们会在绿色数据以后的一段时间会有一个测试点，在这个测试点实际在现场去测了它的容量，根据这个容量结果，跟我们预测的结果进行了一个对比，大概的误差是3%左右。也就是说我们实际基于充电数据的预测基本上是可信的，这是我们给出的更多数据结果。

另外，我们可以进行内阻预测。我们知道内阻会随着SOC和温度变化，都需要修正，我们只是进行了温度修正，而SoC没有进行修正。经过滤波之后，我们可以看到内阻有一个逐渐增长的趋势，这和实际内阻增长也是比较符合的情况。当然这些工作只是我们前面做的工作的一部分，因为我们目标是在云端的寿命预测。我们还可以做更多工作，比如说基于粒子滤波的方法来估计未来有多大概率到某一个程度；比如说你这辆车跑到1000个循环有多大的可能性。这个技术会对我们未来的保险业务会有一些支撑作用。  
 做一个简单的总结。首先，是数据优化很重要。优质数据是一切云端分析和计算的保障，我们必须要进行数据分类、频率优化、数据同步。第二，云端具有组内电芯状态可估的计算能力，算力和储能能力强，可以使用离线的随机化算法。第三，可视化是云端大数据分析的一大优势，电量散点图是可视化的重要工具。第四，云端车具备“车车比较”以及车内电芯的横向比较功能，也可以利用电芯的历史数据进行纵向比较，这个数据为故障预警提供了重要的支持，至于说故障诊断和预警，时间关系我无法汇报。  
 以上是我的汇报，谢谢大家。