感谢丁主任的介绍，也非常荣幸为大家分享我们的一些学术成果。我是来自交大人工智能研究院的，同时也是智能汽车所的所长，我今天给大家分享的是新能源汽车的虚拟电池技术，是大数据驱动的。

一、虚拟电池提出的意义。首先，虚拟电池其实只做两件事，第一个是电池的状态估计。因为现在电池老化包括一些环境的影响，温度等，对它的一些估计精度还是影响很大。在我们现在车企像上汽，它有和阿里云的合作，数据量很大。同时新能源汽车数据中心几十万辆车，几年的数据怎么样去给电池赋能，怎么样把数据真正应用起来。第二个是电池安全，咱们也听到了很多的事故，包括特斯拉的自燃等等，其实这些数据的异常表现还是比较明显的，怎么样从这些数据当中发现早期的故障，去预测我们发生自燃的风险，从而给车企有足够的预警，或者是给用户带来足够的安全的时间预料，这就是我们要做虚拟电池的目的。

现在数据中心几十万辆车，根据国家新能源汽车数据上传的要求，电池数据也是比较全的，当然颗粒度不是非常高，一般是二三十秒一次。对于车企来说，他们更关注上传的密度，所以我们现在正在一起合作，从两个维度来做，一个是状态估计，这是需要很高颗粒度的数据上传。二是故障，几十万辆车有什么故障数据，从中发现其规律，我们希望把共同维度的数据用起来，但是主体就只有虚拟电池，只是说虚拟电池可以产生不同功能而已。

现有我们看SOH的状态估计其实是有几种方法，第一个是基于物理模型的SOH估计方法，但是它基本上还是基于标准工况，实车的数据参数辨识比较困难。二是用恒流-恒压充电数据的方法，但是我们也非常理解，像企业或者是数据中心，他们数据涉密或者说有保密性的要求，我们不拿数据，但是可以应用数据形成一个模型，然后形成一个通用结果，大家都可以用。实车数据我们看下来是比较杂乱的，这个跟用户没有没有规律的使用状态是相关的，所以你不可能用实验室的思维去用实车数据来分析。第三个是使用电流电压温度分布的方法，这个比较粗，可以把健康状态的范围估计出来。我们现在使用这种统计学和信息学采集的方法，从数据驱动端是有局限的，我们从2011年开始做电化学的模型，要真正把数据模型和机理模型融合起来，因为电路非常复杂，再加上数据质量不一定很完美，所以怎么真正从数据端把电池本质的东西，电化学的东西给挖掘出来，这是我们的宗旨。

我们说深度学习是一个黑箱模型，不知道规律如何得出的，可解释性比较差，现在从人工智能的角度，我们和交大数学系合作，把可解释性进行提升。从数据端有关这样的弊端和局限，从机理端是可以增加它的可解释性的，因为机理就是电化学的对象，对象必然存在一定的规律，参数变化会体现在表面的电压、电流、温度的变化，以及它的老化状态是和下面的细分电化学参数是相关的。所以如何从数据端来挖掘这些规律，这是我们希望得到的。

1. 虚拟电池。虚拟电池是与实际电池动态特性一样的数据驱动的电池模型，两者接收相同的负载会有相同的端电压响应，且具有相同的拉化过程和状态以及可能存在的故障风险。我们建立这样一个虚拟电池的模型就是希望可以真正的可以代表电池，为每个单体建立虚拟电池。因为现在像宁德时代的电池包，他每个单体的温度、电压、电流数据都有采集，我们希望和每个电池单体形成这样一个虚拟电池模型。

三、模型。 电子化学模型，我们通过神经网络，我们通过卷积神经网络或者是深度模型和电化学模型来做，这些模型结合起来，首先得到一个虚拟电池，再去推电池在不同工况下的反应是什么。

当然我们有几种方法，包括前馈神经网络、循环神经网络，各有利弊。前馈神经网络的劣势是它不是动态的，在BMS的条件下，一般来说电压估值的误差很大，因为它的动态性不一定跟得上，你的学习网络还是要兼顾实车真正数据反映响应的情况。循环的神经网络可以刻画动态特性，但是训练比较慢，如果说你算力够强，实时性要求不高，可以用这样的方法，但是我们希望真正在云端可以实时得到一些状态估计。所以，我们现在希望可以用新的方法就是从机器学习到向机器学习，向机器学习就会得到电化学的实质模型，最终考虑通过机器学习模型来获得电化学参数。

 因为电池的老化，膜是它的寿命的核心因素之一，另外现在的新能源电池大部分是石墨是电极，当然松下和索尼的电池有硅碳，因为硅原子进入之后可以提升能量密度，但是硅的介入要非常的精准，我们也比较过很多电池，现在国内真正能商用的电池当中，用硅的不多。但是在日韩，用得非常多，所以加硅的比例就是它的Knowhow，因为硅的热胀冷缩会导致它的负极产生很大变化，这个比例是日韩电池极度机密的东西。我们通过十年时间，慢慢对电池进行电化学端的分析，可以看出它的一些特点。这些因素我不具体展开，是影响电池老化的，我们把这些相关因素对应的电化学参数识别出来，相应的老化状态、模型等等就会相对比较精确，这就是把数据驱动和电化学的负反应相关起来。

老化的模型我不具体讲，因为数学的东西比较多。我们一开始用的都是偏微方方程，怎么样能变成等效电路形式，但是内涵还是化学，我们用了三年时间做数学，真正用数学模型给解开。我们从基础端做这样的研究，就是要解决产业端目前面临的痛点问题，所以基础是根本，最终应用是一种最终路径，两者如何结合，我们也是和数学系和化工系的同事合作一起做这个事。 可以看到，它的内部的这些膜是纳米级的，这个东西非常复杂，因为这个膜不是一种物质，而是很多物质通过化合物的聚集产生的，到底这个薄在什么位置，他长厚是长在哪里，这里和涉及到很多电化学的知识，所以我们的团队里也有很多电化学的博士，要真正和车结合把这些问题搞清楚。

我们有一个博士后过来之后，把它的边界长在什么地方搞清楚了。这个参数得到之后，机器学习，人工智能电化学的初始值就可以得到，因为他是后面云端不断迭代的，但是初始值是非常重要的。包括现在我们有很好的设备，因为原来要看电池内部要拆开，把外面的金属壳剥离，还要在无氧环境当中，切割的时候不能直接切割。现在我们是无损的检测，就是通过X射线，只要把外面的金属壳去掉就可以看到内部。像电化学的孔隙率等等方面是可以看到的。另外如果说你要看到纳米级，比如说上海光源有这样的设备，目前我们交大是可以看到微米级。这样就可以看到不同情况下，里面的状态，而不是说要拆开来，拆开之后你的电池就废掉了，下面再做循环是不可能的。这样的话，在我们刚刚说的机器学习的初始状态设置就有了依据，而不是说只是对电池进行估计，估计这个东西准不准我们不知道，但是我们看到了他真实膜的状态或者是离子沉积后的状态，这就是一个物理量，这个物理量是真实的，因为它可观测。

具体的步骤，通过机器学习，包括初始值的设置，结合电化学模型的相同负载来看偏差，通过偏差再去对我们的情况进行辨识，最终调节数据驱动的电化学模型，然后对其不断进行在线调整。这些辨识的结果通过数据驱动，数据可测的其实真不多，因为我们只知道电池、电压、电流、温度，这样这几个可测值，如何通过几个可测值保持精度，这需要不断在线进行调整或者是更新。

结果还是比较准的，特别是平稳行驶状态的误差很小。在平稳的时候，误差为5毫伏，急刹车的时候误差为40毫伏，算是精度很高的。特别是在低温环境下，我们在黑龙江的黑河，这是上汽的冬测基地，通过数据测下来，平均误差可以达到1%。同时考虑不同的低温下，多温度动态放电，我们主要看一下他的动态误差是否能接受，不同温度下的误差不一样，低温下误差更大，这个时候电化学的东西就起作用了，因为本质上它当中的参数在低温下会产生锻压式的变化，如果说用常规的方法，在低温下是不管用的，还是要真正的把机理融合。通过这种VLT曲线，最后我们把实时数据转为常温的恒温恒压的数据。

四、虚拟电池在故障诊断上的应用。电池自燃和滥用是没有关系的，外面也不会产生那么高的环境温度。怎么从内部看到电池发生自燃，这就需要知道它内部的东西。我们现在想得到什么呢？对于明后天，它会不会发生自燃有一个量化预测，而不是说等到温度开始上升了，这个时候可能车里的人逃生时间就很少了，因为它的反应是非常迅速的，自燃会由一个单体蔓延扩散，其他单体受期影响升温。我希望通过我们的方法得到第一个单体什么时候自燃，什么时候发生热失控，这是我们希望得到的。

现在故障的数据确实非常少，即便是故障了，我相信很多企业也会稍微修饰一下再上传，真正到中心的肯定是有故障的。我这里获得的故障数据确实比较少，那么如何做故障的辨识和预测，要通过虚拟电池来得到故障的虚拟电池，得到故障虚拟电池之后，如果说认为它可靠就可以获得更多的故障电池，因为你的虚拟电池是故障状态，你就可以认为它可以产生故障数据，你可以得到比现实更多的故障数据，这是我们做故障虚拟电池的理念。因为这里面涉及到热，温度会影响反应，反映又会影响到温度，这是一个耦合关系。我们通过实验室的非故障数据，因为有的时候电池数据有限，所以我们现在通过实验室数据和实车故障数据结合起来，因为实验室做了十年，电池数据还可以，因此要把它结合起来。另外你要先得到无故障的模型，我们的等效电路当中有一个问题，我们在实车做了两阶，再高阶的话，算力跟不上，精度会受到你的阶数限定，如果说你要真正用参数进行表达，就不会受到阶数限制的问题。

这是我们实测的一些故障数据，这里有一些代码，我不详细展开，但是故障可以分很多类，我们进行了编号。比如，同一类是编号1，另外一类是编号2，一直到10等等。最终我们看这些故障属于哪一类就可以了，然后再去分析故障会产生什么结果。所以不一定是说你现在定义的故障就是我们准确的故障，要启发式的调整电化学参数，给不同的组合去变，变化的时候，跟故障电池已有的实测数据能不能匹配，如果说能匹配，就归为同类，最终我们可以得到故障和电流、电压、温度相关的更多数据。

其实热失控很大的因素就是因为热短路，因为锂枝晶的生长，我们可以说大部分是有锂枝晶生长的，但是很多是电化学组合起来的，而锂枝晶只是外部的体现而已，实际上是很多电化学参数变化产生的锂枝晶的结果。这里面我们也通过电化学模型去把它的形貌给模拟出来，现在我们可以模拟到比较真实的情况。通过我们的观测，可以看到它的形貌和我们预测的结果还是比较接近的。我们是不是可以通过数据端去衍生到形貌的预测变化？这个我们正在打通，很快也会有一些成果出来。
 另外，我们说的这些电化学参数，跟生长分解，电池开裂等等都是产生故障的因素，将这些因素推导到这些电化学参数的变化值，这里我不做详细展开。另外故障注入之后，我们也可以得到一些电化学曲线，说明我们的故障虚拟电池是可以信赖的，至少在一些情况下是可以信赖的。

另外，我们说的这些电化学参数，跟生长分解，电池开裂等都是产生故障的因素，将这些因素推导到这些电化学参数的变化值，我们也可以得到一些电化学曲线。这说明我们的故障虚拟电池是可以信赖的，至少在一些情况下是可以信赖的。通过这个时刻产生的电压电流的温度数据，反推过来，刚刚我们建立了关系，虚拟电池这是一个中间层，前端是它的故障类型，后端是你可测的电流电压温度，后端用桥搭起来，这样就可以把训练值放在一起得到一些规律，得到这些规律之后我们可以得到这个故障之后，最后转成实车数据。其实，现在类型相似的时候，它的内在的浅层的规律是一致的，并不是电池类型大小变了，就会产生很大变化。以上是我分享的内容，谢谢大家。