



基于SCADA系统数据的分层的上下文敏感组串异常诊断技术 在大规模光伏系统中的研究和应用

Hierarchical Context-Aware Anomaly Diagnosis in Large-Scale PV Systems Using SCADA Data

康大海 (Dahai Kang)

能源互联网总监 Energy Internet Director

Email: kangdh@cnegroup.com



协合新能源集团有限公司

Concord New Energy Group Limited – China

Concord New Energy Group (182.HK)



成立于2006年，是一家国际性的可再生能源集团，也是目前香港证券市场上唯一一家具有纵向集成一体化商业发展模式的清洁能源发电上市公司，以风力和太阳能发电厂投资运营、风力和太阳能发电服务业务为主营业务。

Concord New Energy Group Limited (CNE) founded in 2006, specializes in wind and solar power operation. To date, we are the only pure vertical integrated clean energy power company listed on the Hong Kong Stock Exchange.

提纲 (Outline)

- 背景 (Background)
- 挑战 (Challenges)
- 解决方案 (Proposed Solution)
- 结果及分析 (Results)
- 结论及展望 (Conclusion and Future Work)

背景 (Background)

- 光伏运维的广泛需求 (PV O&M Demand)
- 日益增加的单电站装机容量 (Installation Volume per PV Plant)



图 1 中国西藏一个20MW的光伏电站
(Fig.1 A 20 PV Plant in Tibet China)

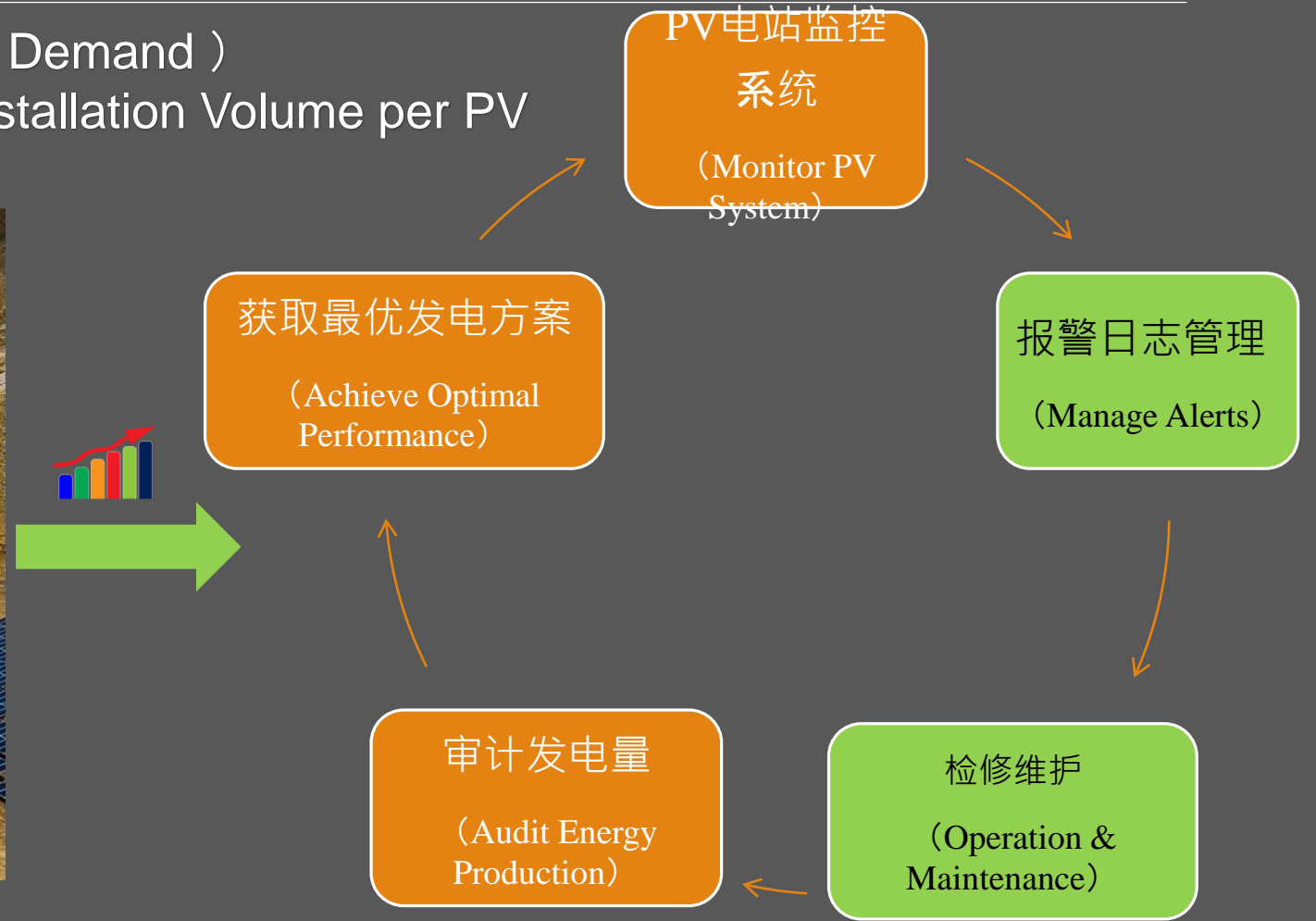


图 2 光伏电站一个典型的运维周期
(Fig.2 An Example of Operation and Maintenance (O&M) Cycle)

挑战1: 光伏系统中异常的多样性和复杂性

- **多样性和复杂性** (Diverse and complex *faults*)

典型异常案例 (Fault Examples) :



光伏板表面破损
(Gunshot Damage)



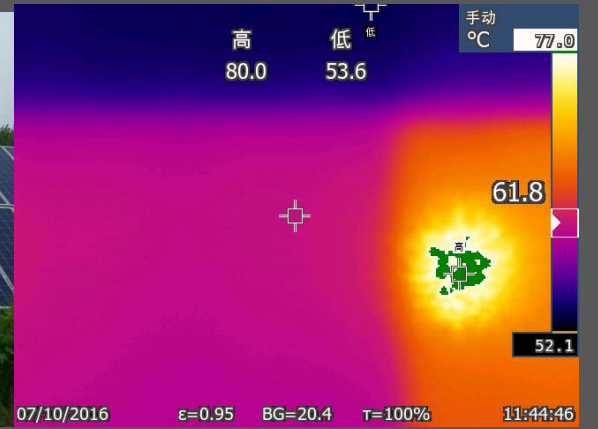
泥土遮挡
(Surface soiling)



杂草遮挡
(Weed Shading)



实物遮挡
(Building Shading)



热斑
(Hotspot)

挑战2： 数据收集系统带来的挑战

- 光伏电站数据收集系统：**SCADA**
(Data collection by supervisory control and data acquisition (SCADA) system)

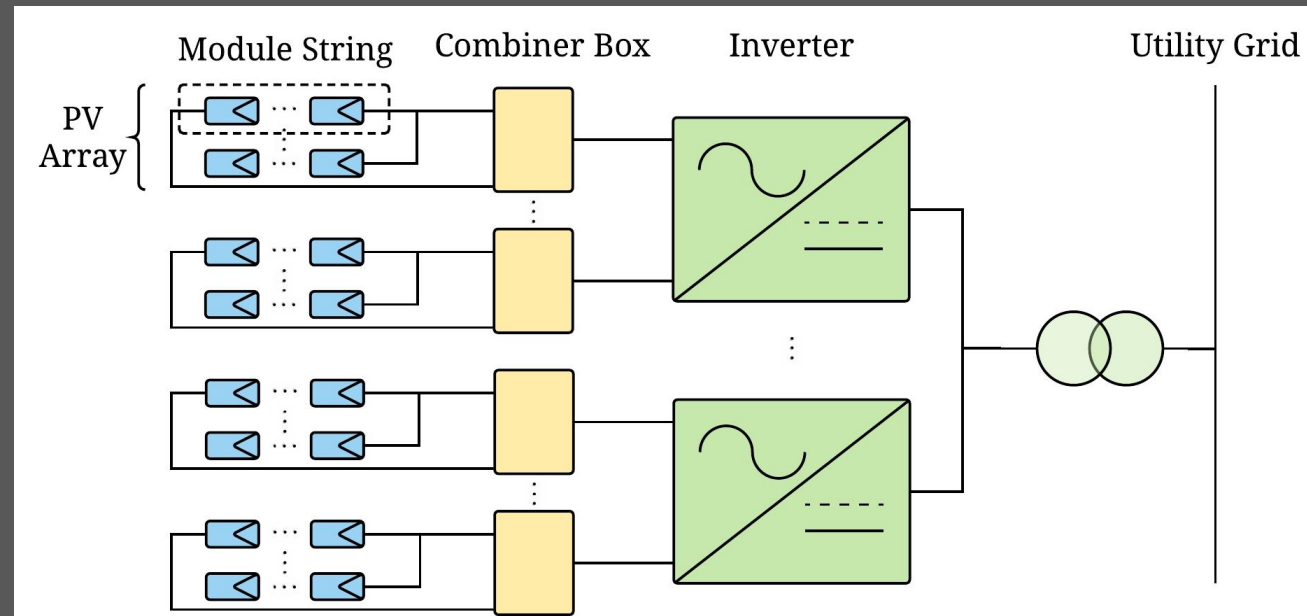


图 3 一个典型的并网光伏系统结构示意图
(Fig.3 Diagram of a grid-connected PV system)

挑战3: 技术上的挑战

基于模型方法

(Model-based)

- 需**增加**额外监测设备 (Extra cost from additional measurement devices)
- 针对特定的故障，使用**受限** (Intolerance to real environment)

基于数据的方法

(Data-driven)

- 有监督的学习方法 (Supervised)
 - **标签数据** (Labeled data)
 - **已知的异常类型** (Known fault types)
- 基于统计的方法 (Statistical)

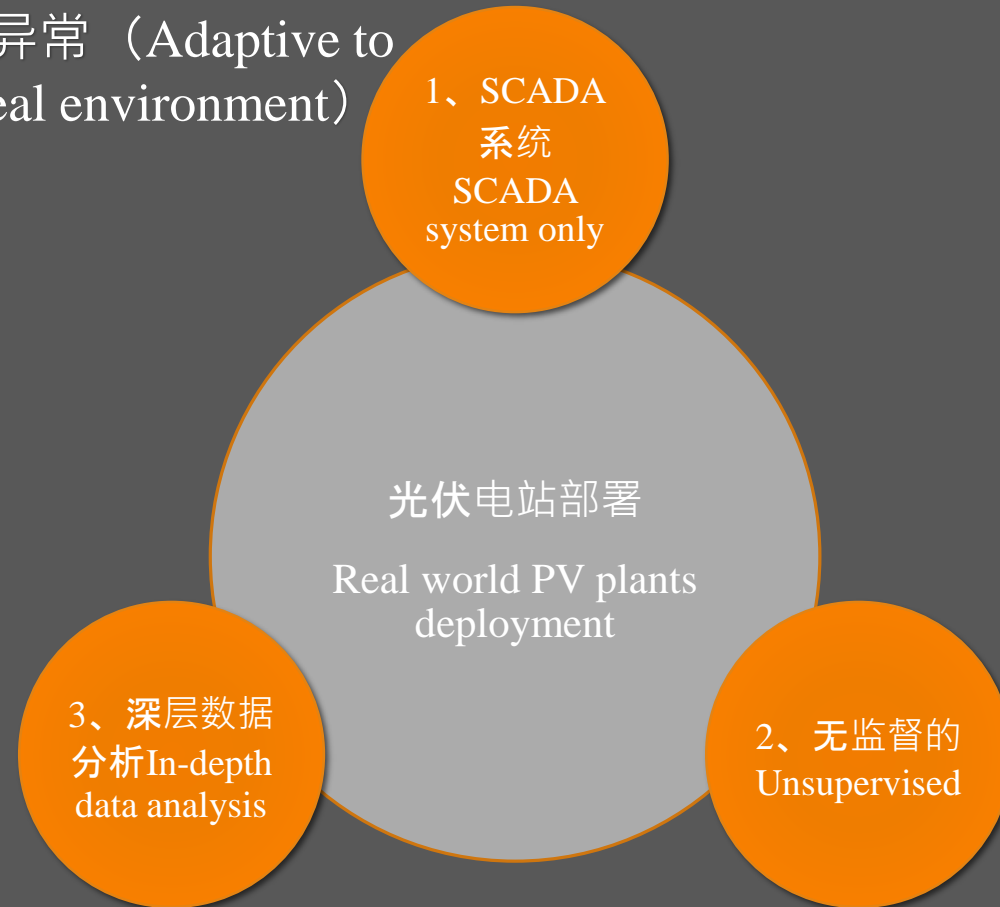
- 以上两种方法都会给光伏系统的运营和维护带来额外的成本，并存在有限的应用范围。这促使了基于SCADA系统的**无监督的学习方法**解决方案的提出。 (If not only the SCADA system is used in anomaly diagnosis solutions, both types of methods can increase deployment cost.)

问题的提出 (Problem)

如何只利用SCADA系统的监测数据去做大规模光伏系统的异常检测？ (How to deploy an anomaly diagnosis solution in large-scale PV systems solely relying on SCADA data?)

解决方案（Solution）

- 经济实用（Cost-effective）
- 可适应现场多种异常（Adaptive to various faults in real environment）



解决方案 (Solution)

■ 数据预处理 (Data preprocessing)

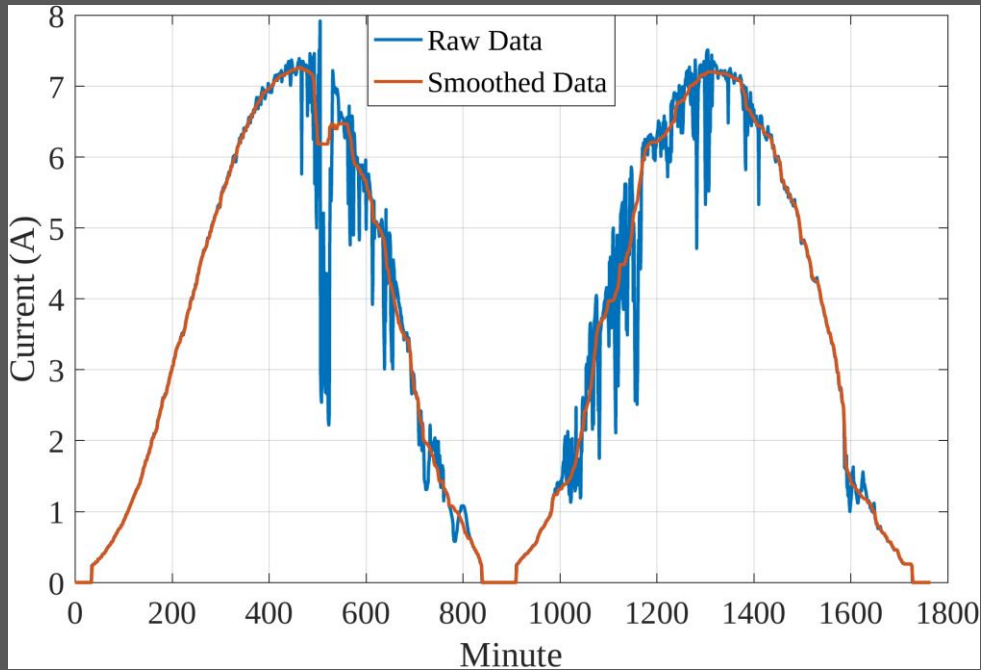


图 4 1分钟采样频率的原始数据和中值滤波之后的数据 (滤波长度为60分钟)

Fig.4 Comparison of 1 minute sampled raw data and 60 minute smoothed data.

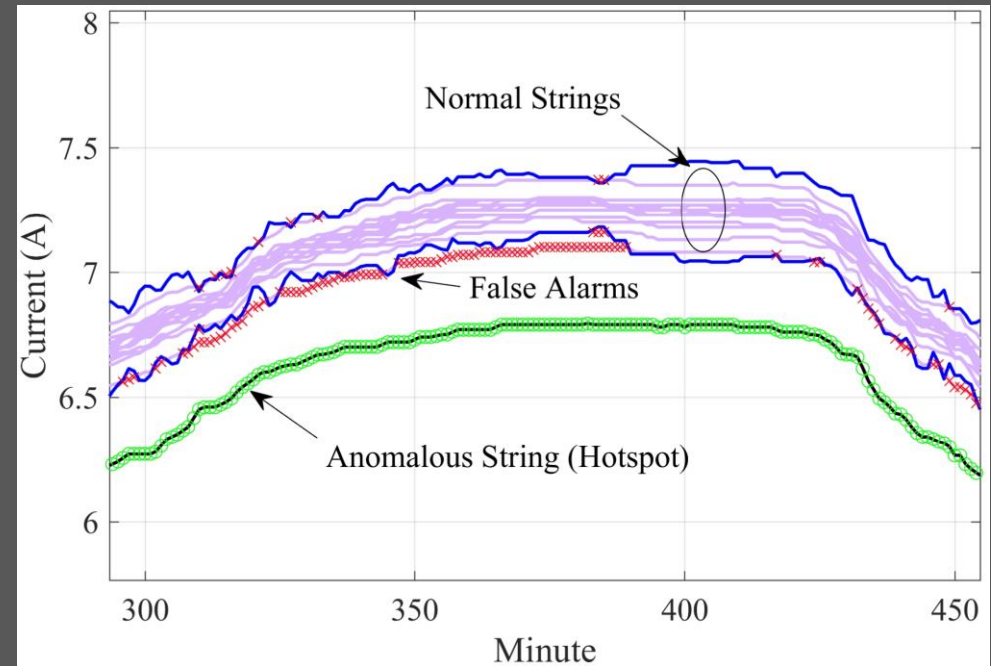


图 5 同一汇流箱下基于统计方法的异常串检测

Fig.5 Anomaly diagnosis using a statistical method for strings in the same combiner box

重要经验信息 1：数据清洗-数据过滤-数据降采样 (Data cleaning – Data filtering – Data down sampling)

解决方案 (Solution)

- 数据分析 (Data analysis)

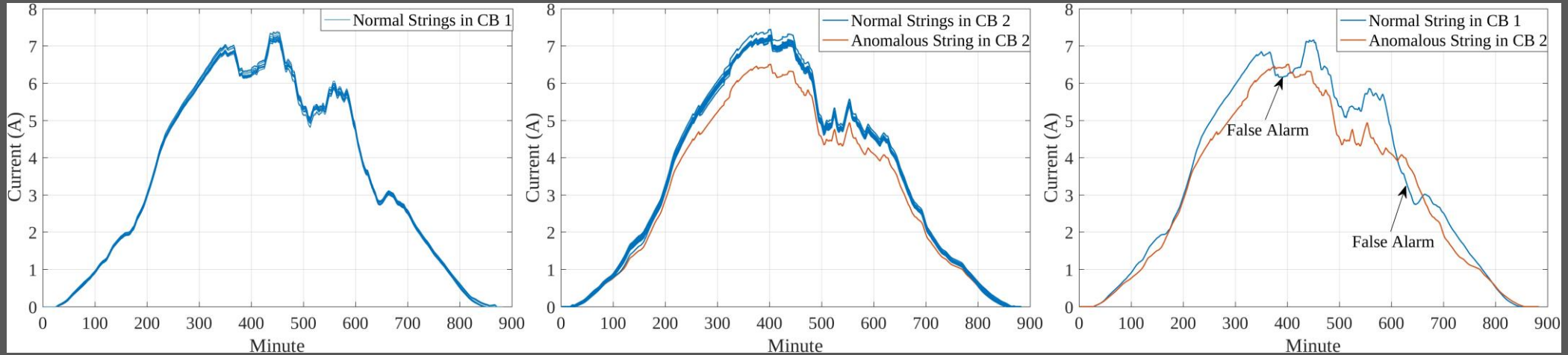


图 6 局部和全局上下文信息

Fig.6 Illustration of local and global context

重要经验信息 2：同时考虑局部和整体信息，可以更深入的理解异常检测的设计方案。 (Taking both local and global context into consideration can provide more insights into anomaly diagnosis solution design.)

解决方案（Solution）

- 分层的上下文敏感异常诊断方案（Hierarchical context-aware anomaly diagnosis）

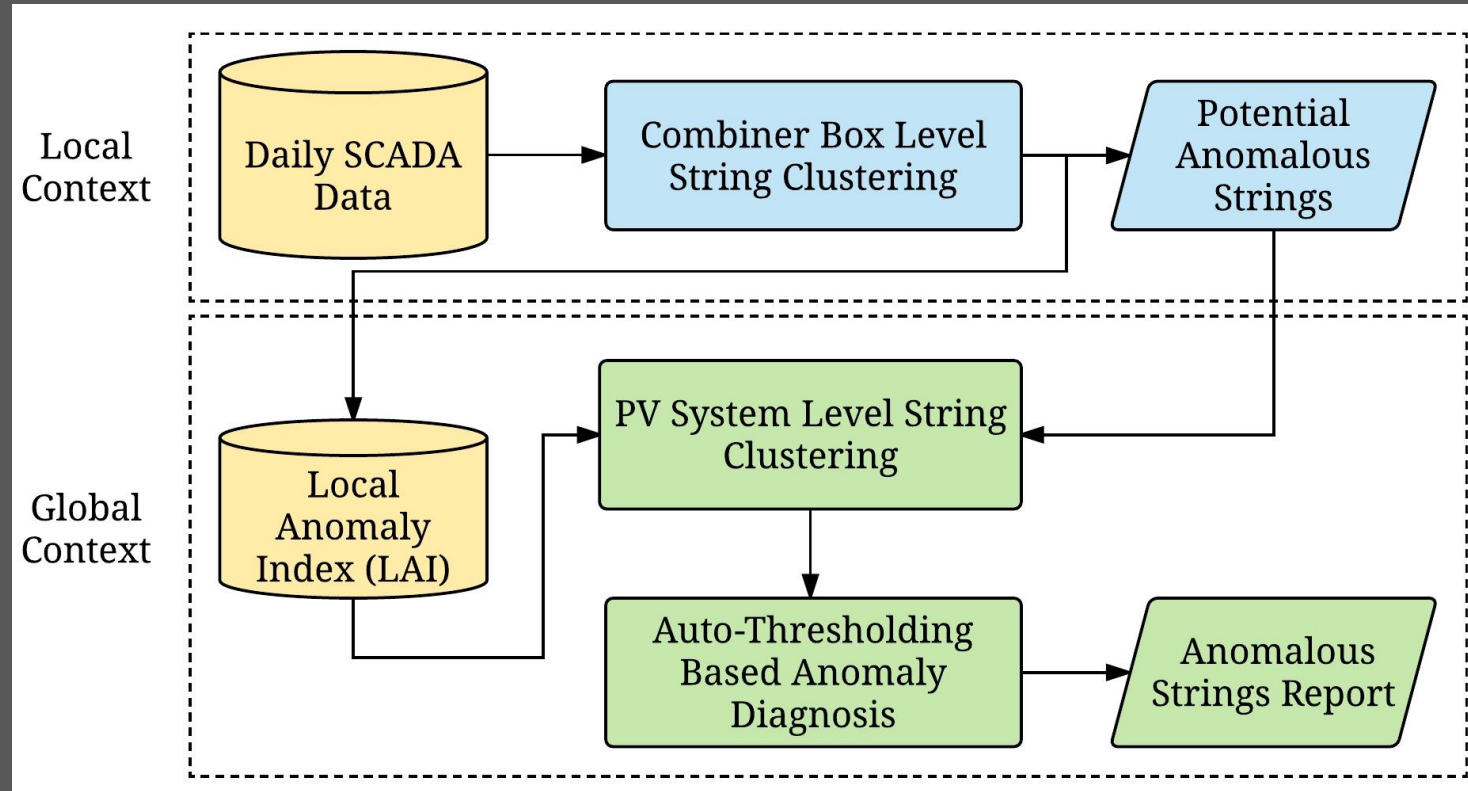


图 7分层的上下文敏感异常诊断方案流程图

Fig.7 Flowchart of the proposed hierarchical context-aware anomaly diagnosis approach.

解决方案 (Solution)

- **LCAD**—局部上下文敏感异常诊断 (Local context-aware anomaly diagnosis – LCAD)

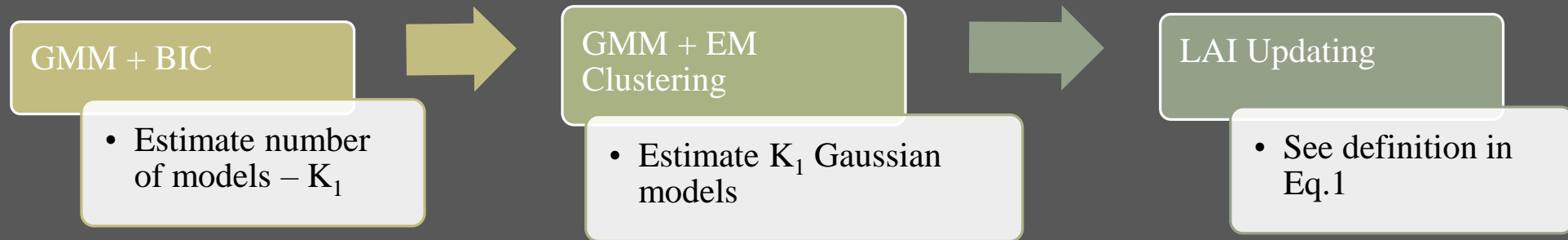


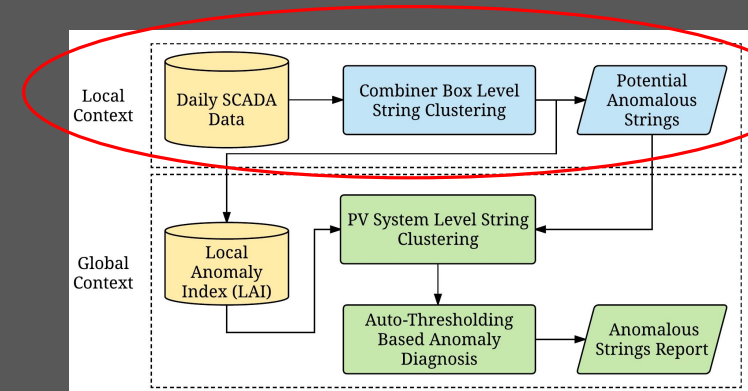
图 8 局部上下文敏感异常诊断层步骤

Fig.8 Steps in local context-aware anomaly diagnosis layer

- Eqn.1, LAI-局部异常系数 (Definition of Local anomaly index)

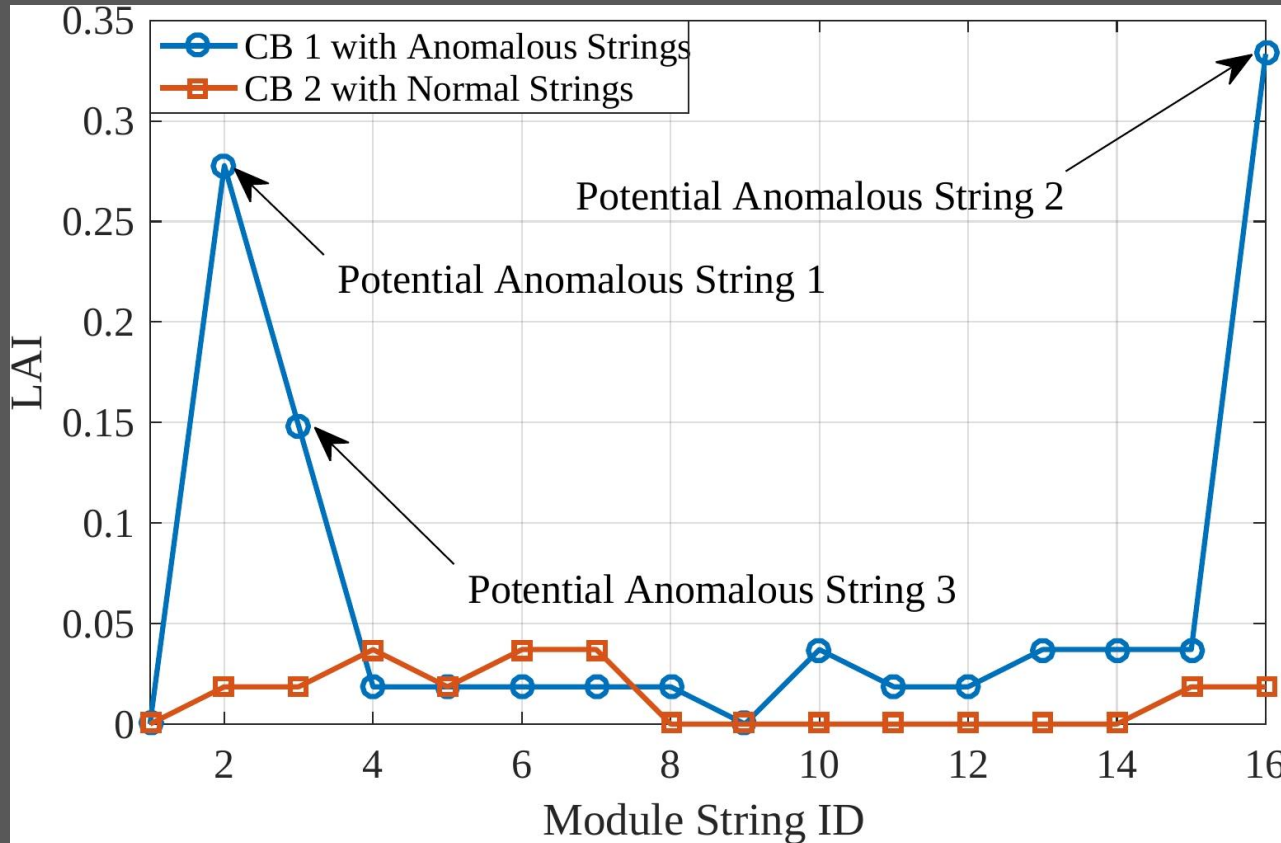
$$LAI(\widetilde{I}_{i,j}) = \frac{\sum_{i=1}^m f(i)}{m}$$

$$f(i) = \begin{cases} 1 & \text{当 } i \text{ 被归到异常分布时 (if } i \text{ is assigned to an anomalous model.)} \\ 0 & \text{其他 (otherwise.)} \end{cases}$$



解决方案 (Solution)

- LCAD 局部上下文异常敏感诊断 (Local context-aware anomaly diagnosis)



- 三个可能异常的串在 LCAD 层被发现 (Three potential anomalous strings identified in LCAD layer)
- 通过 **GCAD** 层可以确定第一, 第二个串为真正的异常串, 串三为假阳性的。 (Potential anomalous string 3 is actually normal, and determined in GCAD layer)

图 9 两个不同汇流箱下 (CB1和CB2) 32串的LAI值

Fig.9 An example: LAIs for 32 strings in two different combiner box (CB1 and CB 2)

解决方案 (Solution)

■ 全局上下文异常敏感诊断-GCAD (Global context-aware anomaly diagnosis)

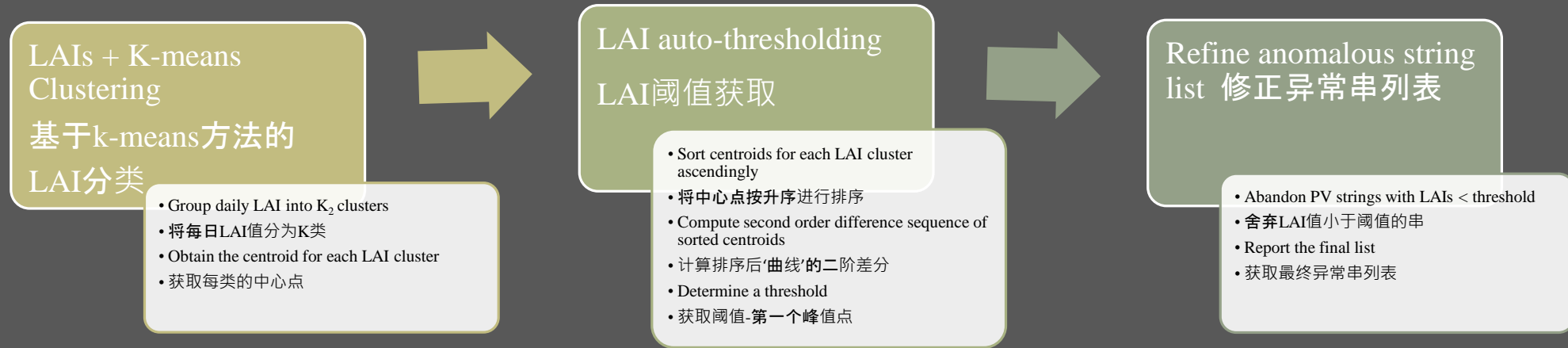
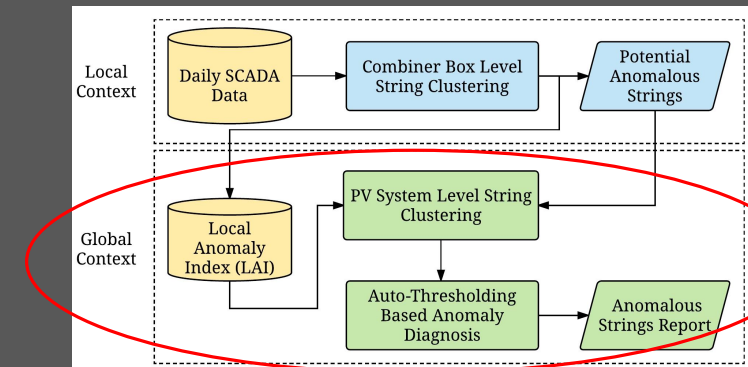
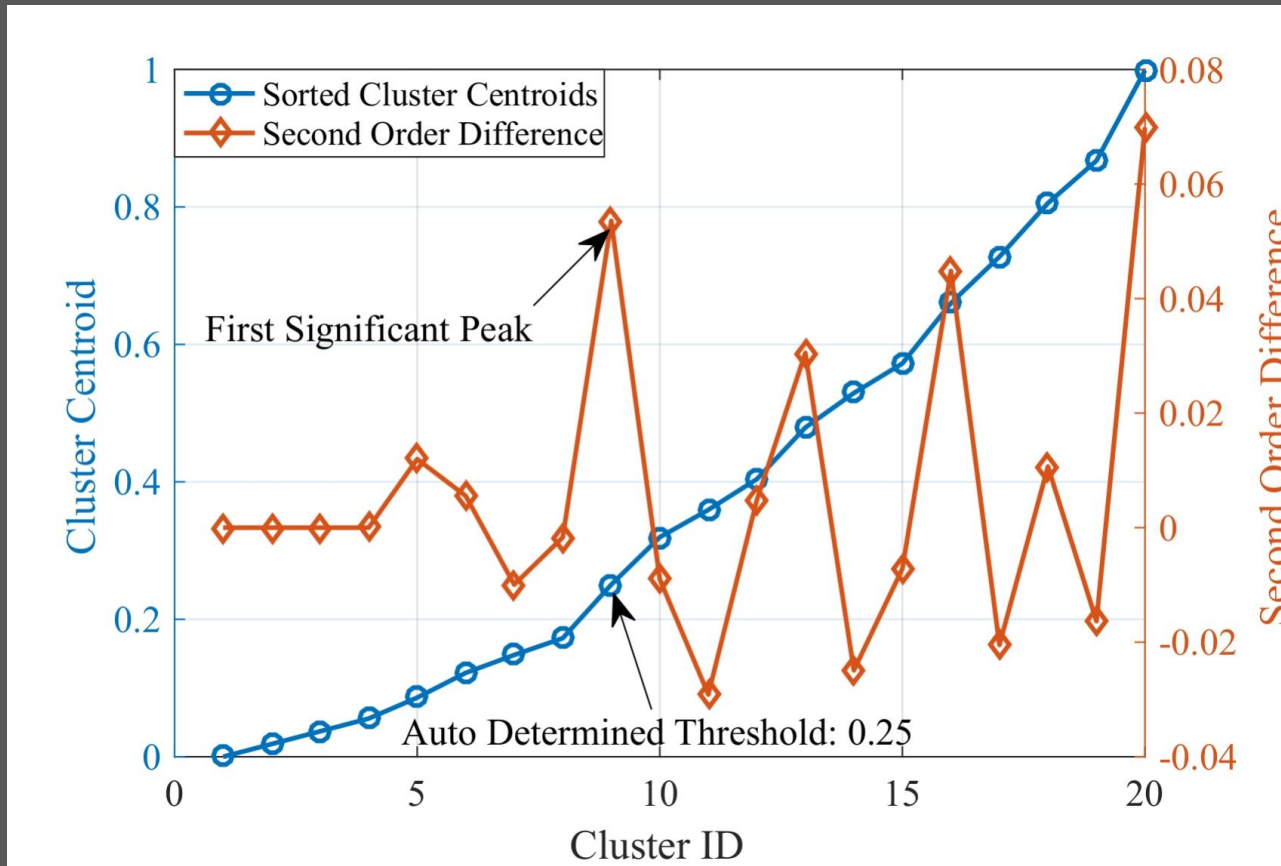


图 10 全局上下文异常敏感诊断层步骤
Fig.10 Steps in global context-aware anomaly diagnosis layer



解决方案 (Solution)

- 全局上下文异常敏感诊断-GCAD (Global context-aware anomaly diagnosis)



- 在这个例子中阈值为**0.25** (In this example, threshold is 0.25)
- LAI值小于0.25的串被认为是正常的串, 这**降低**了错误报警率。
(Strings with LAIs < 0.25 are treated as normal, reducing false alarms)

图 11 自动发现LAI阈值的实例
Fig.11 An illustration of identifying LAI threshold automatically

R

w/m^2

Experiment Setup

实验设置

Effectiveness

效用

Efficiency

效率

Conclusion and Future Work 总结与展望

Conclusion总结

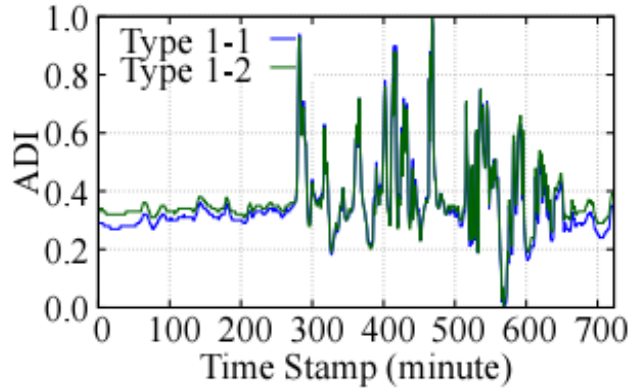
- 提出了一个**经济实用**的分层的上下文敏感异常诊断方法
- A cost-effective hierarchical context-aware anomaly diagnosis approach was proposed
- 基于SCADA系统**每日组串级**的异常检测方案
- Daily string-level anomaly detection using the SCADA system
- **异常检测率高**
- High detection accuracy
- 已部署**10+个**光伏站点，总装机容量：**300+ MWp**
- Deployed to 10+ real world solar farms

Future Work展望

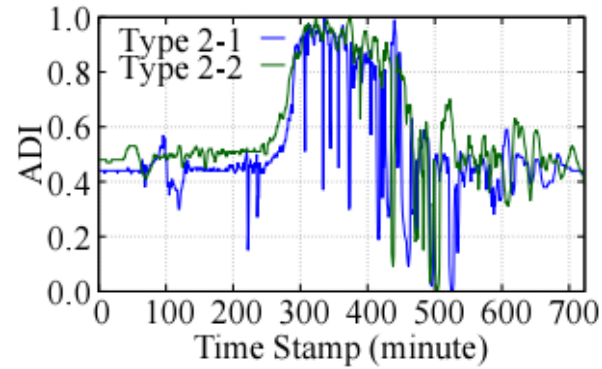
将部署到**更多**不同的光伏电站上来验证模型，并且做准确的分类

Validate the proposed approach using data from multiple solar farms

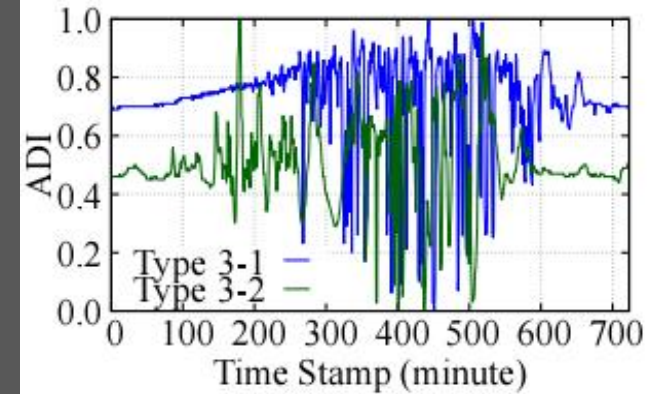
Future Work 展望



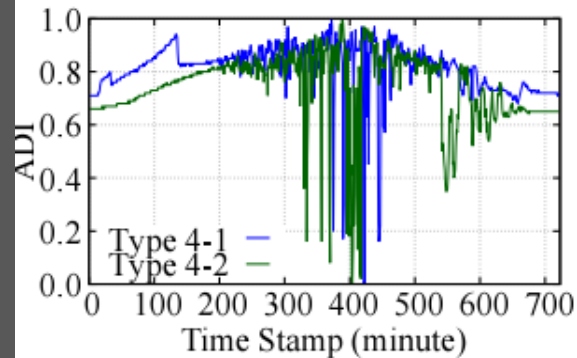
(a) Type 1



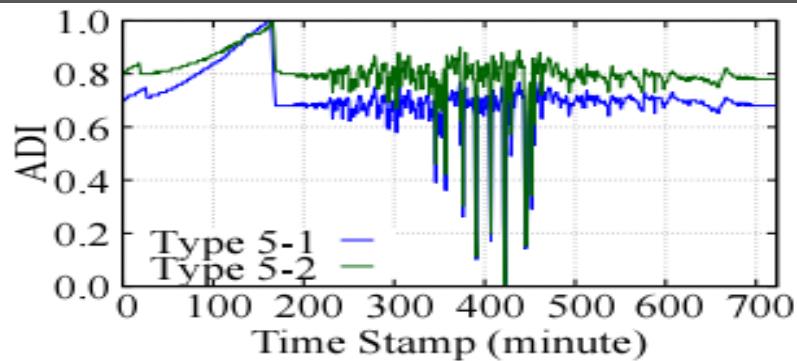
(b) Type 2



(c) Type 3




(d) Type 4



(e) Type 5

POWER

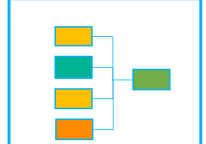


传感器



应用DB
文件

设备及系统



协议适配器




脚本引擎




传输网关

边缘计算服务

设备连接和管理



设备管理




数据接入

数据接入服务




机器学习服务

异常状态检测
故障诊断分析
故障发生预测
能源消耗预测



业务分析服务


时间可利用率分析
性能可利用率分析
能量可利用率分析
故障根本原因分析



数据处理服务

influxdb
docker

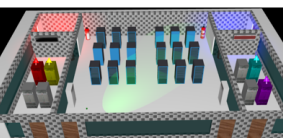
数据管理、业务洞察及预测



多屏展现



大屏展现



3D可视化



业务展现服务

业务展现服务

业务展现





Thank you!