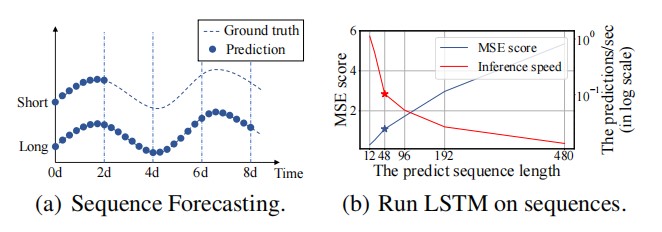
# 概要

许多实际应用需要长序列时间序列的预测，如用电量规划。长序列时间序列预测(LSTF)要求该模型具有较高的预测能力，即能够有效地捕获输出和输入之间精确的长期依赖耦合。最近的研究表明，变压器具有提高预测能力的潜力。然而，变压器存在几个严重的问题，阻止了它直接适用于LSTF，包括二次时间复杂度、高内存使用，以及编码器-解码器体系结构的固有限制。为了解决这些问题，我们为LSTF设计了一个高效的基于变压器的模型，名为告密者，具有三个独特的特点：（i）一种prob稀疏自注意机制，在时间复杂度和内存使用方面实现了O(L\*logL)，在序列依赖对齐方面具有相当的性能。（ii）自我注意力提取的突出部分通过将级联层输入减半来控制注意力，并有效地处理极长的输入序列。（iii）生成式解码器虽然概念简单，但以一个正向操作预测长时间序列，而不是一步一步的方式，这大大提高了长序列预测的推理速度。在四个大规模数据集上进行的大量实验表明，Informer的性能显著优于现有的方法，并为LSTF问题提供了一种新的解决方案。

# 介绍

时间序列预测是许多领域的关键组成部分，如传感器网络监测（帕帕迪米里奥和余2006年）、能源和智能电网管理、经济和金融（朱和沙沙2002年）以及疾病传播分析(Matsubara等2014年)。在这些场景中，我们可以利用大量关于过去行为的时间序列数据来做出长期的预测，即长序列时间序列预测(LSTF)。然而，现有方法大多在短期问题设置下设计，如预测48点或以下（霍克雷特和施米杜伯1997；李等2018；于等2017；刘等2019；秦等2017；温等2017）。越来越长的序列使模型的预测能力紧张，使这一趋势阻碍了对LSTF的研究。作为实证例子，图（1）显示了真实数据集的预测结果，其中LSTM网络预测了变电站从短期（12点0.5天）到长期（480点20天）的每小时温度。当预测长度大于48点(图(1b)中的实星)，MSE性能上升到不理想，推理速度急剧下降，LSTM模型开始失效时，整体性能差距较大。



（图1）

图1：(a)LSTF可以覆盖比短序列预测更长的时间，在政策规划和投资保护方面具有至关重要的区别。(b)现有方法的预测能力限制了LSTF的性能。例如，从长度=48开始，MSE上升得高得令人无法接受，推理速度迅速下降。