



# IEICE Online Transactions

## —多語言翻譯 試用—

跨越語言障礙的知識交流 開始了！

電子資訊通信學會的英文期刊可以使用母語（中文）閱讀的「多語言翻譯試用」現已啟動。

<https://zh-tw.global.ieice.org/>



每年超過 1000 篇經過審核的高可靠性英文論文，您可以用母語閱讀，讓最前沿的技術資訊更輕鬆地獲取。請積極使用此服務，幫助您的研究和工作。

尚未成為會員的朋友，可以免費註冊成為全球聯合會員，隨時獲取 IEICE 最新的資訊。

[https://www.ieice.org/eng\\_r/member/associate\\_member.html](https://www.ieice.org/eng_r/member/associate_member.html)

### 【聯絡方式】

電子資訊通信學會 多語言翻譯部門 E-mail : [translation@ml.ieice.org](mailto:translation@ml.ieice.org)

### 【電子資訊通信學會 入會指南】

[https://www.ieice.org/eng\\_r/join/individual\\_member.html](https://www.ieice.org/eng_r/join/individual_member.html)

The screenshot shows the IEICE website interface with a navigation bar and a main content area. The main content area features a technical diagram of the Informer model architecture, which is a Transformer-based model. The diagram illustrates the flow from inputs through an Informer Encoder (consisting of N stacked Multi-head Bridge Self-attention blocks) to a Dependency pyramid, and then through an Informer Decoder (consisting of M stacked Multi-head Bridge Self-attention blocks) to produce outputs. The diagram also shows a Fully Connected Layer and a Multi-head Attention block. Below the diagram is a caption: "圖 1 Informer 的框架。"

Inform 模型主要包含 N 個編碼器和 M 個解碼器組成 [14]。在 Encoder 中，採用 Multi-head ProSparse Self-attention 機制取代了 Transformer 中傳統的 Self-attention 機制，在減少資料維度的同時保留了豐富的資訊。在解碼器中，使用產生解碼方法在逐層輸出中一次輸出所有預測結果。

Transformer 傳統的 Self-attention 機制需要  $O(L^2)$  的複雜度與記憶體 [15]。由於在序列的注意力計算過程中，只有小部分 Q 和 K 對注意力分數有貢獻，因此資料的注意力分佈稀疏。KL 散度可以衡量重疊與均勻分佈之間的差異。Informer 利用基於 KL 散度的 Self-attention 的稀疏化來降低模型的計算複雜度，使其記憶體和計算成本均為  $O(L \log L)$ 。

### 通訊論文範例