

通过时间序列聚类改进互联网流量矩阵预测

Martha Cash
Worcester Polytechnic Institute
Worcester, MA
mcash@wpi.edu

Alexander Wyglinski
Worcester Polytechnic Institute
Worcester, MA
alexw@wpi.edu

摘要——我们提出了一种利用时间序列聚类来改进互联网流量矩阵 (TM) 预测的新框架, 使用深度学习 (DL) 模型。TM 内的流量流通常表现出多样的时态行为, 这会在跨所有流训练单个模型时妨碍预测准确性。为了解决这一问题, 我们提出了两种聚类策略——源聚类和直方图聚类——在模型训练前将具有相似时间模式的流进行分组。聚类创造了更同质的数据子集, 使模型能够更有效地捕捉潜在模式, 并比适应整个 TM 的单一全局预测方法更好地泛化。与现有的 TM 预测方法相比, 我们的方法分别减少了 Abilene 和 GÉANT 的 RMSE 高达 92% 和 75%。在路由场景中, 我们的聚类预测也分别降低了 MLU 偏差 18% 和 21%, 展示了当使用 TMs 进行网络优化时聚类的实际好处。

I. 介绍

A TM 定量描述了网络中源目标 (SD) 节点对之间的交通需求, 并且对于网络规划至关重要, 作为许多流量工程 (TE) 任务的关键输入, 例如路由优化和网络资源分配 [19]。然而, 在实时测量和收集 TM 具有挑战性 [19], 这促使了需要准确的 TM 预测来支持 TE 任务 [5]。

最近关于 TM 预测的研究趋势集中在深度学习 (DL) 模型上 [5], [11], [12], 这些模型利用了 DL 在交通数据中建模复杂非线性关系的能力。这些模型表现出强大的性能, 如通过根均方误差 (RMSE) 和平均绝对误差 (MAE) 等错误指标所显示的那样。大多数方法采用全局预测范式, 训练单个模型来同时预测 TM 中的所有交通需求。正如 [25] 所示, 当数据集特征表现出相似模式时, 全局模型可以胜过独立预测每个交通需求的局部模型。然而, [21] 认为异质性特征会导致全局模型准确性下降。

参照图 1, 我们观察到在一个 TM 内的流量需求在行为上往往存在显著差异。一些需求, SD 对

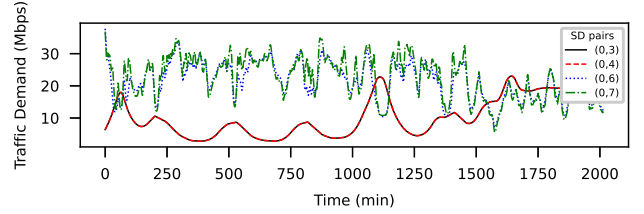


图 1: 来自不同源目的对的流量在 Abilene [16] 数据集中表现出不同的行为。

(0,3) & (0,4) 表现出相似的趋势, 而其他 SD 对 (0,4) & (0,7) 则明显偏离。流量之间的异质性导致了复杂 DL 架构的发展, 如基于注意力的卷积递归神经网络 (ARCNN) [8] 和卷积长短时记忆网络 (CNN-LSTM) [26], 这些设计旨在捕捉 TM 中多样的流动态。

在这项工作中, 我们挑战了为流量矩阵预测设计复杂全局模型的日益增长的趋势。相反, 我们提出通过聚类来利用交通流子集之间的内在相似性。聚类是一种无监督学习技术, 它将时间序列划分为组内相似度高而组间相似度低的组别 [27]。与其拟合一个单一模型来预测整个流量矩阵, 我们为每个流簇训练单独的模型。

本文的主要贡献是:

- 新型的两种主题模型聚类方法: 基于源的聚类和基于直方图的聚类。
- 基于聚类的模型与三种最先进的 (SOTA) TM 预测基线的比较。
- 聚类方法的 TM 预测性能评估使用了定量误差指标 (RMSE 和 MAE) 以及在实际路由环境中进行评估。

论文的其余部分组织如下。第 II 节回顾了与 TM 预测和时间序列聚类相关的研究工作。第 III 节概述了 TMs 和 TM 预测。第 IV 节详细介绍了我们提出的聚类

方法。第 V 节概述了数据集和实验设置。第 VI 节展示了评估结果。最后，第 VII 节总结了论文并概述了未来工作的方向。

II. 相关工作

互联网流量预测是一个时间序列预测问题。从经典统计方法到深度学习方法，广泛的各种模型已被应用于这一领域。Otoshi 等。将流量分解成长期和短期成分，并应用季节自回归综合移动平均模型进行预测 [9]。然而，该方法需要反复对数据进行差分以达到平稳性，这限制了其实际应用。Valadarsky 等。探索前馈神经网络、卷积神经网络 (CNNs) 和非线性自回归模型用于流量预测，但这些模型未能捕捉到时间依赖关系，导致对 TM [2] 的预测性能较差。

为了更好地建模时间动态特性，最近的工作 [10], [13], [15] 转向了循环神经网络 (RNN)，如长短时记忆网络 (LSTM) 和门控循环单元 (GRU) 模型，以预测整个流量矩阵。其他方法，例如 [6], [7]，将每个流量矩阵元素视为独立的时间序列，并为每个元素训练单独的模型。刘等。提出了一种混合方法，该方法先预测总的流量量，并使用预先计算的比例将其分配到各个流量矩阵元素中，同时通过专门的模型对高流量流进行修正 [14]。然而，一些研究 [4], [8] 已经表明，独立建模流量矩阵元素会忽视它们之间的相关性，这可能会潜在地损害整体预测性能。

TMs 可以解释为一组相互关联的时间序列的结构化集合。在这个更广泛的背景下，一些研究工作已经探讨了聚类时间序列预测。Bandara 等人。解决异构时间序列预测的挑战是通过首先基于统计属性（如均值、方差、自相关）对序列进行聚类，然后为每个聚类拟合 RNN 模型 [21]。Martinez 等。使用几何距离对时间序列进行聚类，并应用 k -最近邻进行预测 [22]。张等人。提出使用流量概率分布对交通网络中的流量进行聚类，而邹等。则应用层次聚类和卷积神经网络来预测交通系统中的车道交通量 [23], [24]。

此前没有相关工作将时间序列聚类应用于互联网流量矩阵。本文旨在通过聚类相似的流量流并在每个聚类上训练专用模型来填补这一空白，从而探索提高预测准确性同时降低模型复杂性的潜力。

III. 交通矩阵预测概述

对于具有 N 个节点的网络，流量矩阵是一个 $N \times N$ 矩阵，其中每个条目 (i, j) 表示从源节点 i 到目的地节点 j 的流量量，通常以字节为单位。在时间间隔 Δt ，TM 表示为 TM_t ，其中 $t \in [t, t + \Delta t] \subset T$ 和 T 是完整的测量周期。一个在时间 t 的 TM 是：

$$\text{TM}_t = \begin{bmatrix} T_{1,1}^t & T_{1,2}^t & \cdots & T_{1,n}^t \\ T_{2,1}^t & T_{2,2}^t & \cdots & T_{2,n}^t \\ \vdots & \vdots & \ddots & \vdots \\ T_{n,1}^t & T_{n,2}^t & \cdots & T_{n,n}^t \end{bmatrix} \quad (1)$$

在整个区间 T 上，每个 TM 条目 $\text{TM}[i, j]$ 变成一个长度为 T 的时间序列，捕捉来自节点 i 到 j 的流量，称为流量，并且每对 (i, j) 称为源-目的 (SD) 对。

TM 预测的目标是根据 k 个之前的矩阵历史来预报 $\text{TM}_{t+1} : (\text{TM}_{t-k}, \dots, \text{TM}_t)$ 。TMs 表现出非线性、非平稳性和周期性的行为，这使得预测变得具有挑战性。由于 DL 模型能够捕捉这种复杂模式，因此它们非常适合这一任务。我们将预测表述为一个监督学习问题，遵循 [2] 中的方法。

为了构建训练数据，我们在连续的 D 个 TM 上应用滑动窗口以生成长度为 L [5] 的 $W = D - L + 1$ 个窗口。在每个窗口中，前 $L - 1$ 个矩阵作为输入，第 L^{th} 个矩阵是预测目标。在整个 TM 上重复此操作可生成 DL 模型的输入-输出对训练集。

TM 预测方法分为两类：全矩阵 (EM) 预测和局部预测 [14], [7], [5]。EM 预测使用单一模型来预测整个 TM，而局部预测则为每一流训练一个模型，对于具有 N 个节点的网络，这将导致 N^2 个模型。尽管局部预测通常能提供更高的准确性 [7]，但它计算成本高昂且扩展性差。EM 模型可以匹配局部性能 [14]，但通常需要复杂、资源密集的架构。

EM 预测的一个关键限制是其对流量同质性的隐含假设，而这一假设在实践中很少成立。Gao 等人. [8] 表明来自同一源头的流量往往具有相关性，而来自不同源头的流量则可能没有。这表明传输矩阵中包含可用于预测的潜在结构。

IV. 提出的交通矩阵预测聚类方法

我们提出了一种用于 TM 预测的第三种预测方法：基于聚类的预测，使用时间序列聚类。我们不单独建模

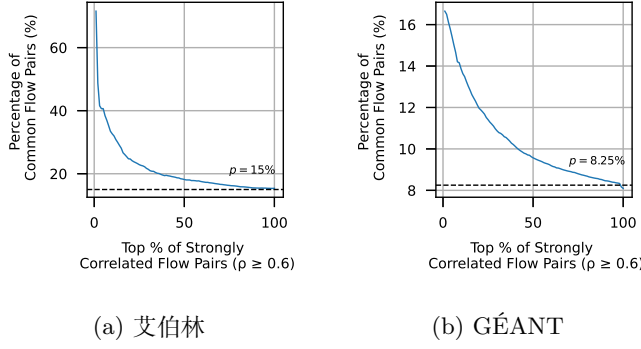


图 2: 在前 $x\%$ 强相关的流量对中, 具有相同源的流量的比例。强相关流量随着包含更多对而减少。虚线表示基线百分比。Abilene 有 15% 的流量对是强相关的基线比例。GÉANT 的基线百分比为 8.25%。

每个流或整个矩阵, 而是将相似的流分组, 并为每个簇训练一个模型, 以减少与局部预测相比的模型数量, 同时比 EM 模型更好地保持流特定的动力学特性。在这项工作中, 我们研究了两种聚类方法: 基于源的方法, 根据起源节点对流进行分组; 以及基于直方图的方法, 根据分布相似性对流进行聚类。

A. 基于源的聚类

第一种聚类方法, 基于源的聚类, 是由 [8] 激发的。来自同一源节点的流量流应该表现出相似的特点, 如周期性, 因为这些流量需求是由相同的来源生成的 [1]。为了证明这一点, 我们计算了流之间的相关性 (ρ)。相关性反映了两个随机变量之间的关系, 皮尔逊相关系数是一种常用的测量相关性的方法 [28]。流之间的相关性是按照以下方式计算的:

$$\rho = \frac{\text{cov}(\mathbf{TM}[i, j], \mathbf{TM}[i', j'])}{\sigma_{\mathbf{TM}[i, j]} \sigma_{\mathbf{TM}[i', j']}} \quad (2)$$

其中 $\text{cov}(\cdot)$ 是协方差函数, $\sigma_{\mathbf{TM}[i, j]}$ 是整个测量区间 T 内流 $\mathbf{TM}[i, j]$ 的方差。变量 ρ 在 $[-1, 1]$ 之间变化。越接近 1 表示两个流量正相关性越高, 越接近 0 则表示负相关性越高。图 2 说明了当流量来自同一源时, 使用两个实际数据集 Abilene 和 GÉANT, 交通流是强相关的 ($\rho \geq 0.6$)。数据集的更多详细信息在第 V 节中提供。

图 2 显示了强相关的流对 ($\rho \geq 0.6$) 在 Abilene 和 GÉANT 中的时间一致性。在 Abilene 中, 大量流对随着时间保持高度相关, 尤其是前 15%, 这表明存在稳定的流量关系核心。相比之下, GÉANT 显示出较低的一致性和更小的稳定核心, 暗示了更多瞬时和变化的流动

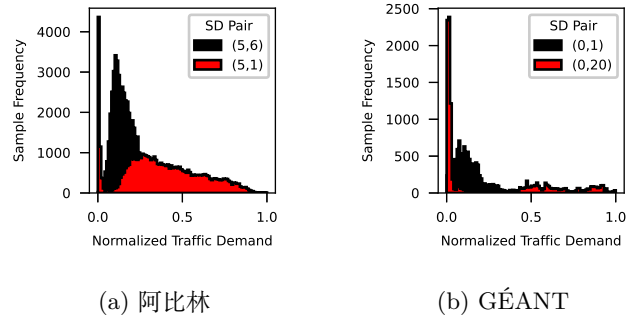


图 3: 归一化交通需求的直方图, 针对两个具有共同源点的目的地流, 在 (a) Abilene 和 (b) GÉANT 中。分布显示, 来自同一来源的流量可以表现出显著不同的统计行为。

行为。这种对比突显了需要在如 Abilene 这样的稳定网络中建模持久依赖性, 同时保持在像 GÉANT 这样的动态网络中的适应性。在 Abilene 中, 一小部分流之间的稳定相关性的集中表明主导来源经常驱动流量。这支持基于源头的聚类, 将流按起源分组以利用共享的时间模式并提高存在稳定关联时的预测稳健性。

B. 基于直方图的聚类

虽然基于源的聚类利用了来自同一源节点的数据流可能共享时间模式这一直觉, 但它未能捕捉到流量大小或变化性上的差异。如图 3 所示, 来自相同源的数据流例如 (0, 6), (0, 7), (0, 3) 和 (0, 4) 可以表现出截然不同的流量分布, 表明结构上的接近并不总是意味着行为相似。

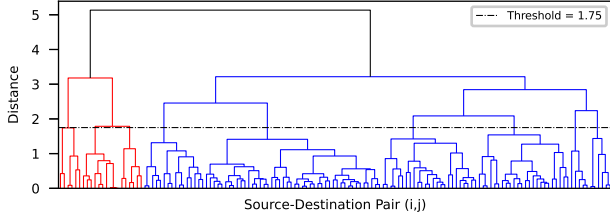
为了解决这一问题, 我们引入了一种基于互补直方图的聚类方法, 该方法根据其归一化流量时间序列的分布特性对流进行分组。通过将每个流表示为一个直方图, 我们可以捕捉更广泛的统计特征, 例如突发性、平坦度或多模态性, 而不受源或目的地的影响。这使得能够将具有相似使用模式的流进行分组, 即使这些流跨越不同的节点。

我们使用 Jensen-Shannon 散度 (JSD) 量化直方图的相似性, 它是 Kullback-Leibler (KL) 散度的一个对称且有界的版本。给定两个流直方图 P 和 Q , JSD 定义为:

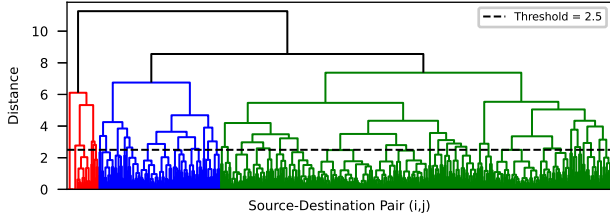
$$\text{JSD}(P||Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M) \quad (3)$$

其中 $D(\cdot)$ 表示 KL 散度。JSD 当 P 和 Q 相同时评估为 0，并在它们完全不相交时达到 1。

我们计算所有流之间的成对 JSD 值，结果得到一个 $N^2 \times N^2$ 距离矩阵。使用层次聚类，每个流最初形成自己的聚类，然后迭代合并具有最小 JSD 的那些聚类。此过程产生一个链接矩阵，该矩阵捕获了合并的层级结构。



(a) 阿比林



(b) GÉANT

图 4: 用于 Abilene 和 GÉANT 数据集的树状图，可视化 Jensen-Shannon 散度距离矩阵的层次聚类。虚线表示确定形成聚类数量的阈值。

为了选择聚类的数量，我们分析了图 4 中的系统树，并在突出的垂直间隙处切割，这些间隙表明自然分割。我们分别为 Abilene 和 GÉANT 选择了 1.75 和 2.5 的切割距离，分别得到 9 和 47 个聚类，从而平衡了碎片化和聚类粒度。

V. 实验设置

本节描述了我们在实验中使用的数据集、预处理、基准和评估指标。

A. 数据集与预处理

我们使用来自 Abilene [16] 和 GÉANT [17] 网络的真实数据来评估我们的预测方法。Abilene 网络有 12 个节点和 30 条链路，流量矩阵在 24 周内每 5 分钟收集一次。GÉANT 网络有 23 个节点和 74 条链路，在 4 个月内每 15 分钟收集一次流量矩阵。我们将数据集分为 80/20 的训练-测试比例，并且使用了 10% 的训练样本作

为验证集。我们使用 10 个流量矩阵作为历史输入进行预测。我们使用最小-最大归一化将流量流缩放到范围 $[0, 1]$ 用于训练。

B. 基准模型

我们将聚类方法与三种代表无聚类基线的 SOTA 模型进行基准测试：Prophet [3]，ARCNN [8] 和独立的 GRU [5]。Prophet 是一种以交通工程 (TE) 为中心的框架，使用具有角度中心损失函数的 GRU 模型进行流量矩阵预测。ARCNN 是将 CNN 结合来捕获流间相关性、RNN 用于流内依赖关系 [15] 以及长范围时间建模注意力机制的 SOTA 时空模型。GRU 是一种广泛用于流量矩阵预测的 RNN 类型，因为它能够保留长期依赖关系。这三个模型均在 EM 预测视角下运行。我们还将与局部预测视角进行比较，并训练 N^2 个 GRU 模型（即每个流量流一个）[7]。

我们评估聚类方法和基准测试的性能使用均方根误差 (RMSE) 和平均绝对误差 (MAE) 指标，定义为：

C. 性能指标

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (TM_i - \hat{TM}_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

其中， n 表示测试集中的样本数量， TM_i 表示真实的 TM，而 $\hat{TM}_i$ 表示预测的 TM。

我们还在流量工程 (TE) 的背景下评估了预测的交通矩阵 (TM)，该方法通过路由流来最小化最大链路利用率 (MLU) 从而减少拥塞。TE 被公式化为一个多商品流问题，使用 TM 作为输入 [2], [3], [4]。预测的 TM 应产生接近真实值的 MLU 值才有用。由于流量比例的小变化可能会导致显著的 MLU 变动，即使 RMSE 或 MAE 值较低也可能无法准确反映 TE 性能。为了评估这一点，我们按照以下方式计算 MLU 偏差和平均 MLU 偏差：

$$Bias = \frac{\hat{U}}{U} \quad (6)$$

$$\overline{Bias} = \frac{1}{n} \sum_{i=1}^n \frac{\hat{U}_i}{U_i} \quad (7)$$

其中 $\hat{U}$ 是从预测的 TM 得出的 MLU 输出, 而 U 是从真实 TM 得出的 MLU 输出。接近 1 的偏差表示更好的预测性能。我们在计算预测的 TM 的 MLU 时使用 Abilene 和 GÉANT 对应的网络拓扑。

D. 实现细节

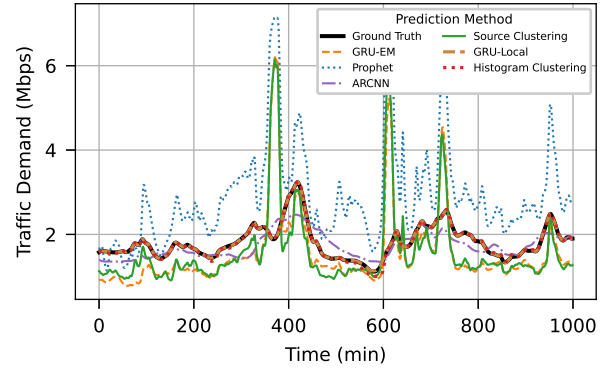
所有深度学习模型均在 PyTorch 中实现, 而 TE 模拟使用 Gurobi。我们对所有聚类方法和基准模型 (不包括 ARCNN) 采用一个具有 30 个隐藏层大小的一层 GRU, 遵循之前工作 [4] 中的配置。模型训练了 100 个周期, 批量大小为 32, 使用 ADAM 优化器和学习率为 0.001。我们使用提前停止策略, 耐心值为 5, 最小 delta 值为 $1e-5$, 以防止过拟合。所有参数选择均基于现有预测方法中常用的设置 [2] - [15]

VI. 结果

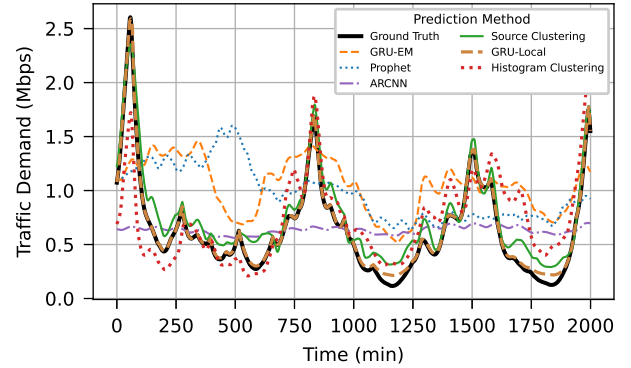
表 I 总结了各种预测方法的 RMSE 和 MAE。基于直方图的聚类在两个数据集上的 EM 模型中实现了最低误差, 相比 Abilene 上的 ARCNN 降低了 92% 的 RMSE, 相比 GÉANT 上的 Prophet 降低了 75% 的 RMSE。这突显了通过分布相似性对流进行分组的好处, 这产生了更为同质化的训练数据, 并改进了时间建模。虽然没有超越局部预测的表现, 但它在 EM 模型上提供了显著的优势, 并且比训练 N^2 个单独模型要远为可扩展。源聚类在 Abilene 上的改善有限, 但在 GÉANT 上相比 Prophet 降低了 67% 的 RMSE, 表明其在网络特定拓扑结构中的潜在应用。

图 5 强调了在交通流层面进行聚类的好处。对于采样流, 基于直方图的聚类 (虚线红色) 紧密跟踪真实情况 (实线黑色), 而 EM 方法由于偏向主导流量经常偏离。在 Abilene 上, 源聚类 (绿色实线) 的表现与 EM 方法相似, 表明该流被归组到高容量流中。然而, 在 GÉANT 上, 它更接近真实情况, 并且偶尔胜过直方图聚类。这表明聚类的有效性因拓扑和流量行为而异; 但是, 直方图聚类通常优于 EM 方法, 并与局部 GRU 模型相匹配, 展示了其在泛化和效率方面的优势。

图 6 展示了整个 TM 预测, 说明了比 EM 基线更好的重建质量。直方图聚类在高负载下更好地保持空间模式, 如图 6a 中的亮方形所示, 捕捉到了重要流的强度和布局。源聚类表现相似但稍微平滑了局部热点。在稀疏设置中, 如图 6b 所示, 两种聚类方法通过保持低流量的位置并避免 GRU EM 和 Prophet 中看到的过度平



(a) 对 Abilene 流量流对 (6,3) 的预测。



(b) 对 GÉANT 流量流对 (6,16) 的预测。

图 5: 从源目的地对采样的交通流用于不同的预测方法。滑, 超过了 EM 的表现。这些结果突显了聚类在时间和空间上对建模交通异质性的益处。

表 II 展示了每种方法对路由性能的影响。直方图聚类在两个数据集上都产生了大约 1 的 MLU 偏差, 匹配了局部预测, 并且与 Prophet 模型或 GRU EM 模型相比, 在 Abilene 上的表现优于 ARCNN 模型 18%, 在 GÉANT 上则优于 21%。这表明它保留了关键的空间和时间流特性。虽然源聚类在 Abilene 上的表现较差, 但在更大更复杂的网络 GÉANT 上取得了良好的结果。

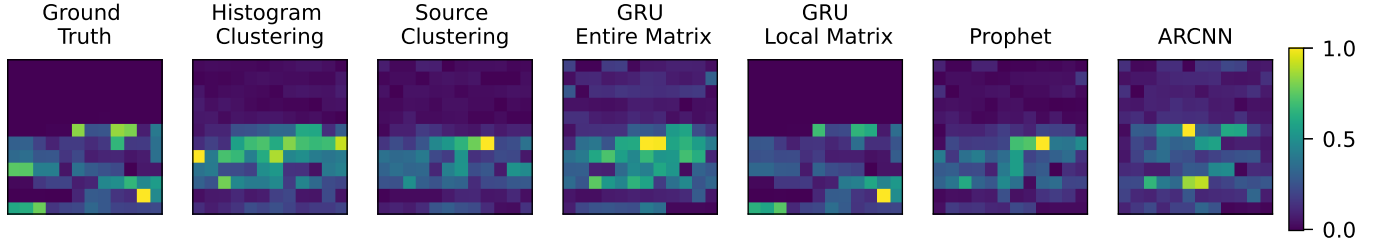
VII. 结论

我们提出了一种基于聚类的互联网流量矩阵预测框架, 引入了两种策略: 源节点聚类, 即将流按照源节点进行分组; 直方图聚类, 即将流按照分布相似性进行分组。我们在 Abilene 和 GÉANT 数据集上对这两种方法进行了评估, 并将它们与三种 EM 模型以及一种每个流拟合一个模型的局部基线进行了比较。

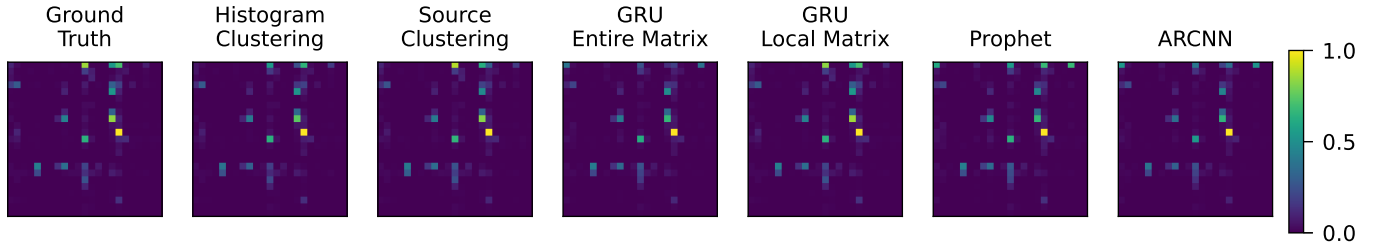
直方图聚类在 RMSE、MAE 和 MLU 偏差方面始终优于 EM 模型, 因为它使模型能够从更同质的流组

表 I: 阿比琳网和 GÉANT 网上的 TM 预测方法的归一化平均测试误差

网络	度量	直方图聚类	源聚类	先知	ARCNN	GRU 门	GRU 本地
阿比琳	RMSE	2.38e-2	0.286	0.249	0.303	0.294	2.89e-3
	MAE	1.80e-2	0.242	0.198	0.253	0.249	1.88e-3
GÉANT	RMSE	8.55e-2	0.114	0.354	0.257	0.244	1.11e-2
	MAE	6.79e-2	0.092	0.306	0.207	0.196	7.79e-3



(a) 阿比琳



(b) GÉANT

图 6: 预测的交通矩阵与两个代表性时间步的真实值相比。

表 II: 预测流量矩阵在 Abilene 和 GÉANT 网络上的平均 MLU 偏差

网络	直方图聚类	源聚类	先知	ARCNN	GRU 门控循环单元	GRU 本地
Abilene	1.00	0.80	0.92	0.82	0.92	1.02
GÉANT	1.07	1.06	0.88	0.89	0.88	1.06

中学习，从而提高泛化能力。虽然聚类并不能完全匹配局部预测的准确性，但它实现了相当的路由性能。它在准确性和可扩展性之间提供了更好的权衡，避免了训练 N^2 模型的成本。

未来的工作包括探索替代的距离度量、模糊聚类方法以及选择最优聚类数量的策略。

致谢

作者感谢 Randy Paffenroth 教授的富有洞察力的对话和宝贵的反馈，这些帮助塑造了这项工作的开发。

参考文献

- [1] P. Tune, M. Roughan, H. Haddadi, and O. Bonaventure, “Internet traffic matrices: A primer,” *Recent Advances in Networking*, vol. 1, pp. 1–56, 2013.
- [2] A. Valadarsky, M. Schapira, D. Shahaf, and A. Tamar, “Learning to route,” in *Proc. 16th ACM Workshop on Hot Topics in Networks*, 2017, pp. 185–191.
- [3] Y. Zhang, N. Han, T. Zhu, J. Zhang, M. Ye, S. Dou, and Z. Guo, “Prophet: Traffic engineering-centric traffic matrix prediction,” *IEEE/ACM Transactions on Networking*, 2023.
- [4] Z. Liu, Z. Wang, X. Yin, X. Shi, Y. Guo, and Y. Tian, “Traffic matrix prediction based on deep learning for dynamic traffic engineering,” in *Proc. 2019 IEEE Symposium on Computers and Communications (ISCC)*, 2019, pp. 1–7.

- [5] A. Azzouni and G. Pujolle, "NeuTM: A neural network-based framework for traffic matrix prediction in SDN," in *Proc. NOMS 2018 - IEEE/IFIP Network Operations and Management Symposium*, 2018, pp. 1–5.
- [6] D.-H. Le, H.-A. Tran, S. Souihi, and A. Mellouk, "An AI-based traffic matrix prediction solution for software-defined network," in *Proc. ICC 2021 - IEEE Int. Conf. Communications*, 2021, pp. 1–6.
- [7] W. Zheng, Y. Li, M. Hong, X. Fan, and G. Zhao, "Flow-by-flow traffic matrix prediction methods: Achieving accurate, adaptable, low cost results," *Computer Communications*, vol. 194, pp. 348–360, 2022.
- [8] K. Gao, D. Li, L. Chen, J. Geng, F. Gui, Y. Cheng, and Y. Gu, "Incorporating intra-flow dependencies and inter-flow correlations for traffic matrix prediction," in *Proc. 2020 IEEE/ACM 28th Int. Symposium on Quality of Service (IWQoS)*, 2020, pp. 1–10.
- [9] T. Otsuchi, Y. Ohsita, M. Murata, Y. Takahashi, K. Ishibashi, and K. Shiimoto, "Traffic prediction for dynamic traffic engineering," *Computer Networks*, vol. 85, pp. 36–50, 2015.
- [10] J. Zhao, H. Qu, J. Zhao, and D. Jiang, "Towards traffic matrix prediction with LSTM recurrent neural networks," *Electronics Letters*, vol. 54, no. 9, pp. 566–568, 2018.
- [11] D. Aloraifan, I. Ahmad, and E. Alrashed, "Deep learning based network traffic matrix prediction," *International Journal of Intelligent Networks*, vol. 2, pp. 46–56, 2021.
- [12] W. Jiang, "Internet traffic prediction with deep neural networks," *Internet Technology Letters*, vol. 5, no. 2, pp. e314, 2022.
- [13] S. Troia, R. Alvizu, Y. Zhou, G. Maier, and A. Pattavina, "Deep learning-based traffic prediction for network optimization," in *Proc. 2018 20th Int. Conf. Transparent Optical Networks (ICTON)*, 2018, pp. 1–4.
- [14] W. Liu, A. Hong, L. Ou, W. Ding, and G. Zhang, "Prediction and correction of traffic matrix in an IP backbone network," in *Proc. 2014 IEEE 33rd Int. Perf. Comp. and Comm. Conf. (IPCCC)*, 2014, pp. 1–9.
- [15] N. Ramakrishnan and T. Soni, "Network traffic prediction using recurrent neural networks," in *Proc. 2018 17th IEEE Int. Conf. Machine Learning and Applications (ICMLA)*, 2018, pp. 187–193.
- [16] Y. Zhang, "Abilene Network Traffic Matrices," <https://www.cs.utexas.edu/~yzhang/research/AbileneTM/>, 2004. [Accessed: Jan. 16, 2025].
- [17] S. Uhlig, B. Quoitin, J. Lepropre, and S. Balon, "Providing public intradomain traffic matrices to the research community," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, pp. 83–86, Jan. 2006.
- [18] A. A. Assad, "Multicommodity network flows—a survey," *Networks*, vol. 8, no. 1, pp. 37–91, 1978.
- [19] A. Soule *et al.*, "Traffic matrices: balancing measurements, inference and modeling," *SIGMETRICS Perform. Eval. Rev.*, vol. 33, no. 1, pp. 362–373, Jun. 2005.
- [20] Kyunghyun Cho, Bart van Merriënboer, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *arXiv preprint*, arXiv:1406.1078, 2014.
- [21] K. Bandara, C. Bergmeir, and S. Smyl, "Forecasting across time series databases using long short-term memory networks on groups of similar series," *CoRR*, vol. abs/1710.03222, 2017. [Online]. Available: <http://arxiv.org/abs/1710.03222>
- [22] F. Martínez, M. P. Frías, M. D. Pérez, and others, "A methodology for applying k-nearest neighbor to time series forecasting," *Artificial Intelligence Review*, vol. 52, pp. 2019–2037, 2019. [Online]. Available: <https://doi.org/10.1007/s10462-017-9593-z>
- [23] P. Zhang, W. Ma, and S. Qian, "Cluster analysis of day-to-day traffic data in networks," *Transportation Research Part C: Emerging Technologies*, vol. 144, p. 103882, 2022. [Online]. Available: <https://doi.org/10.1016/j.trc.2022.103882>
- [24] X. Zou and E. Chung, "Network-wide lane-level traffic flow prediction via clustering and deep learning with limited data," in *Proc. 8th Int. Conf. on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 2023, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/MT-ITS56129.2023.10241694>
- [25] P. Montero-Manso and R. J. Hyndman, "Principles and algorithms for forecasting groups of time series: Locality and globality," *arXiv preprint arXiv:2008.00444*, 2020. [Online]. Available: <https://arxiv.org/abs/2008.00444>
- [26] W. Jiang, "Internet traffic matrix prediction with convolutional LSTM neural network," *Internet Technology Letters*, vol. 5, e322, 2022. [Online]. Available: <https://doi.org/10.1002/itl2.322>
- [27] J. Paparrizos, F. Yang, and H. Li, "Bridging the gap: A decade review of time-series clustering methods," *arXiv preprint arXiv:2412.20582*, 2024. [Online]. Available: <https://arxiv.org/abs/2412.20582>
- [28] R. Taylor, "Interpretation of the correlation coefficient: a basic review," *Journal of Diagnostic Medical Sonography*, vol. 6, no. 1, pp. 35–39, 1990.