arXiv:2301.09187v1 [cs.DS] 22 Jan 2023

一种基于邻域聚合的图同构不变量

Alp´ar J¨uttner 和 P´eter Madarasi

摘要 本文提出一种新的图同构不变量，称为 $\mathfrak{w}$-*labeling*，可用于设计多项式时间算法，以解决各种图类中的图同构问题。例如，所提出的组合方法能够区分所有非余谱图对；此外，在满足某些特征子空间性质的假设下，甚至能够区分非同构余谱图。

我们还研究了上述标签的一种改进形式，称为 $\mathfrak{s}^k$-*labeling*，它兼具理论和实际应用价值。其中，它可用于生成图指纹，从而唯一标识所考察数据库中的所有图，包括至多有 64 个节点的所有强正则图，以及至多有 12 个节点的所有图。可以证明，它能够在同构意义下识别所有树和 3-连通平面图，由此也为这两类图提供了一种新的同构算法。这种指纹的实际价值在于显著加快图数据库中的搜索，而这在生物和化学应用中是一项常见需求。

关键词 图同构，图指纹，图哈希，图数据库，强正则图，同构不变量，平面图

1 引言

graph isomorphism problem 是 NP 中少数几个既未被证明属于 P、也未被证明为 NP 完全的自然问题之一。与此同时，针对各种图类，研究者已经开发出多项式时间图同构算法，例如

A. J¨uttner，罗兰大学运筹学系，ELTE，P´azm´any P´eter s´et´any 1/C，1117 布达佩斯，匈牙利。电子邮件：alpar.juttner@ttk.elte.hu P. Madarasi，罗兰大学运筹学系，ELTE，以及 E¨otv¨os Lor´and 匈牙利科学院（ELKH）组合优化研究组，P´azm´any P´eter s´et´any 1/C，1117 布达佩斯，匈牙利。电子邮件：madarasip@staff.elte.hu

树和平面图 [1], 有界度图 [2], 区间图 [3] 或排列图 [4]。此外，最近还针对 *colored hypergraph isomorphism problem* [5]提出了一种固定参数可处理算法。图同构问题已从实用角度得到广泛研究，在大多数情况下都能高效求解。最佳的实用图同构算法包括 Nauty [6], VF2 [7] 及其变体 [8]。

在许多应用中，我们不仅需要比较两个图，还需要在大型图数据库中查找给定图 G 的一个同构副本。与其分别求解 G 和数据库中每个图之间的图同构问题，不如生成所谓的图指纹，使其满足以下性质：如果两个图的指纹不同，那么这两个图就不同构。作为预处理步骤，计算数据库中所有图的指纹，并计算查询图 G 的指纹后，我们可以筛选出数据库中指纹与 G 相同的图。然后，我们可以检查这些筛选出的图中是否有任何一个与 G 同构。

先前的工作 图指纹已得到广泛应用，并且已有多种生成方案。例如，有研究通过考察 [9] 中的短路径（及其节点标签）来生成图指纹。图的谱（即其邻接矩阵的谱）是另一种同构不变量，已从理论角度得到研究 [10, 11], ；此外，研究者还将其与所谓的热核相结合 [12], ，使其也具备了实际应用价值。文献 [13], 对谱唯一的图的数量进行了数值考察，范围截至 12 个节点；结果发现，约 80% 的图可由其谱唯一确定。

近年来，人们开发了多种基于离散时间量子游走（DTQW）或连续时间量子游走（CTQW）的算法，旨在区分非同构图对。众所周知，标准单粒子离散时间量子游走和连续时间量子游走都无法区分参数相同的强正则图（SRG）；此外，不含相互作用的定粒子数连续时间量子游走无法区分任何参数相同的强正则图对，参见 [14] 和 [15]。然而， [14] 中提出的一种单粒子离散时间量子游走变体，其区分能力被证明强于标准离散时间量子游走。也就是说，对于某些参数相同的非同构强正则图对，它能生成不同的特征，但仍有一些强正则图对无法被它区分。 [16], 表明，就图同构问题而言，连续时间量子游走的能力不如离散时间量子游走。另一方面，一种采用相互作用玻色子的先进量子游走方法能够区分所有节点数不超过 64 的强正则图 [17]。相比之下，第 3 节介绍的易于计算的指纹不仅能区分上述所有强正则图，还能为图提供紧凑的指纹。

颜色细化是一种经典方法，用于证明两个图不同构。尽管颜色细化无法处理所有图对，但它以很高的概率将随机图与任何非同构图区分开来 [18]。颜色细化的一种可能扩展是著名的 Weisfeiler-Leman 方法 [19], ，也称为多维颜色细化。本文考虑另一种过程，它可以视为推广经典颜色细化算法的另一种方式。

我们的结果 本文引入了 $\mathfrak{w}$-标记的概念，在某些条件下（这些条件适用于广泛的图对），可借此在多项式时间内解决图同构问题。本文证明，所提出的组合方法能够区分所有非余谱图对，而无需计算图谱。此外，即使图是余谱图，本文也给出了若干确保这些图能够被区分的条件。

本文还介绍了上述标记的一种改进形式，称为 $\mathfrak{s}^k$-标记。其应用包括一种指纹生成方法，该方法能够唯一识别所考虑的图数据库中的所有图——包括所有已知的强正则图以及所有节点数不超过 12 的图。因此，它的性能可与最先进的量子漫步算法相媲美。此外，它能将图的所有信息压缩为一个简短的同构不变指纹。我们还构造了两个它无法区分的非同构图。

记号 通常，集合用花括号表示，多重集则用花括号后加上标井号表示。例如，$\{1,2\}$ 表示由数字 1、2 组成的集合，而 $\{1,1,2\}^{\#}$ 表示由数字 1、1 和 2 组成的多重集。令 $\mathbb{N}$ 和 $\mathbb{Z}_+$ 分别表示正整数集和非负整数集。

在本文中，$G=(V,E)$ $G_1=(V_1,E_1)$ 和 $G_2=(V_2,E_2)$ 表示三个任意的无自环无向图，且每个图至少有两个节点，其中 V、V_1、V_2 分别为节点集，E、E_1、E_2 分别为边集。为简便起见，假设这三个节点集均为 $\{1,\dots,n\}$，即 $V=V_1=V_2=\{1,\dots,n\}$。这些图的邻接矩阵分别为 A、A_1、$A_2\in\{0,1\}^{n\times n}$。令 $N_G(i)$ 表示图 G 中节点 i 的邻居集合。

除非另有说明，本文给出的结果同样适用于含有自环的图。请注意，对于节点标记图，可以通过添加自环来表示标签；显然，即使图同时含有自环和节点标签，也可以仅使用自环对它们进行紧凑编码。

令 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$ 和 $\mu_1\geq\mu_2\geq\cdots\geq\mu_n$ 分别表示 A_1 和 A_2的特征值。若对所有 i都有 $\lambda_i=\mu_i$，则称 G_1 和 G_2 为 *cospectral* 。令U、$V\in\mathbb{R}^{n\times n}$为正交矩阵（即 $U^TU=I$和 $V^TV=I$），使得 $A_1U=U\,\mathrm{diag}(\lambda_1,\lambda_2,\dots,\lambda_n)$ 且 $A_2V=V\,\mathrm{diag}(\mu_1,\mu_2,\dots,\mu_n)$。$U$和 V分别称为 G_1和 G_2的 *eigenmatrices*。令 u_1、u_2、…、u_n和v_1、v_2、…、v_n分别表示 U和 V的列向量。注意，V既表示 G_2 的特征矩阵，也表示 G的节点集，但这不会造成歧义。令 u_{ij} 表示特征向量 u_i的 j^{th} 个元素，换言之，它是 U中第 j^{th}行、第 i^{th}列的元素，其中 i、$j\in\{1,\dots,n\}$。矩阵 Q 的前k 列记作 $Q|_k$。最后，若 $i=j$，则令 $\delta_{ij}=1$，否则令其为 0。

2 通过计数游走进行标记

对于图 $G=(V,E)$，令 $\mathfrak{w}_G\colon V\to\mathbb{Z}_+^{V\times\mathbb{Z}_+}$，使得 $\mathfrak{w}_G(i)_{jl}$ 表示节点 i 与节点 j 之间长度为 l 的游走数量，其中 $l\geq 0$。函数 $\mathfrak{w}_G$称为 *(infinite)* $\mathfrak{w}$*-labeling*。若两个矩阵 Q_1 和 Q_2 满足

permutation-equal 且存在一个置换矩阵 P，使得 $PQ_1 = Q_2$。此等价关系记作 $Q_1 \overset{\mathrm{P}}{=} Q_2$。

根据 $\mathfrak{w}_G$ 的定义，以下命题显然成立。

Claim 若对于两个节点 $u \in V_1$ 和 $v \in V_2$，$\mathfrak{w}(u) \overset{\mathrm{P}}{\neq} \mathfrak{w}(v)$，则不存在将节点 u 映射到节点 v 的 G_1 与 G_2 之间的同构。

定义 **1** G_1 和 G_2 是 $\mathfrak{w}$-*equivalent*，如果可以对节点重新标记，使得对于每个节点 i，$\mathfrak{w}_{G_1}(i) \overset{\mathrm{P}}{=} \mathfrak{w}_{G_2}(i)$。

Claim 如果两个图同构，那么它们是 $\mathfrak{w}$-等价的。

稍后将证明，对于特殊的图对，反向结论也成立。

2.1 只有短游走才重要

$\mathfrak{w}$ 分配给节点的矩阵长度是无限的，因此，检查两个这样的矩阵是否在置换意义下相等并不简单。现在，我们证明只需比较矩阵的前（$n+1$）列即可。

定义 2 对于域上的给定列向量 $q_0, q_1, \ldots$，$\mathsf{span}(q_0, q_1, \ldots)$ 表示由列向量 $q_0, q_1, \ldots$ 生成的线性子空间。

以下引理将在定理 1 的证明中发挥作用。

引理 **1** *For an arbitrary real square matrix* $M \in \mathbb{R}^{n\times n}$ *and a column vector* $q_0 \in \mathbb{R}^n$, $\mathsf{span}(q_0, q_1, q_2, \ldots) = \mathsf{span}(q_0, q_1, \ldots, q_{n-1})$ *where* $q_i := M^i q_0$ *for all* $i \geq 0$.

Proof 通过归纳法可以证明，对于所有 i，$\mathsf{span}(q_0, q_1, \ldots, q_i) = \mathsf{span}(q_0, q_1, \ldots, q_{i+1})$ 蕴含 $\mathsf{span}(q_0, q_1, \ldots, q_i) = \mathsf{span}(q_0, q_1, q_2, \ldots)$。因此，列 $q_0, q_1, \ldots, q_n$ 生成 $\mathsf{span}(q_0, q_1, q_2, \ldots)$。

下面的定理表明，只需考虑 $\mathfrak{w}$ 标签的前 $(n+1)$ 列，也就是说，只有短游走的数量才重要。令矩阵 $\mathfrak{w}_G|_k(i)$ 表示矩阵 $\mathfrak{w}_G(i)$ 的前 k 列。

定理 **1** *For every graph pair* G_1, G_2 *on* n *nodes and for all* $i_1 \in V_1, i_2 \in V_2$, $\mathfrak{w}_{G_1}(i_1) \overset{\mathrm{P}}{=} \mathfrak{w}_{G_2}(i_2)$ *if and only if* $\mathfrak{w}_{G_1}|_{n+1}(i_1) \overset{\mathrm{P}}{=} \mathfrak{w}_{G_2}|_{n+1}(i_2)$.

Proof 在此证明中，令 Q_1, Q_2, Q_1' 和 Q_2' 分别表示矩阵 $\mathfrak{w}_{G_1}(v_1)\mathfrak{w}_{G_2}(v_2)$、$\mathfrak{w}_{G_1}|_{n+1}(v_1)$ 和 $\mathfrak{w}_{G_2}|_{n+1}(v_2)$。若 $Q_1 \overset{\mathrm{P}}{=} Q_2$，则根据定义，存在一个置换矩阵 P，使得 $PQ_1 = Q_2$。显然，$PQ_1 = Q_2$ 意味着 $PQ_1' = Q_2'$。为证明反方向，假设 $Q_1' \overset{\mathrm{P}}{=} Q_2'$，并令 $q_0, q_1, q_2, \ldots$ 和 $q_0', q_1', q_2', \ldots$ 分别表示 Q_1 和 Q_2 的各列。回想一下，A_1 和 A_2 分别表示 G_1 和 G_2 的邻接矩阵。由于 $Q_1' \overset{\mathrm{P}}{=} Q_2'$，存在一个置换矩阵 P，使得 $PQ_1' = Q_2'$，因此只需证明对所有 $i \geq n+1$，$Pq_i = q_i'$ 均成立。

由归纳法，假设对于所有 $0 \leq k < i$，$Pq_k = q_k'$ 成立。根据引理 1，存在

$^{n-1}$ $n-1$存在系数 α_0、$\ldots\alpha_{n-1}\in\mathbb{R}$，使得 $q_{i-1}=\sum\alpha_j q_j$ 且 $q'_{i-1}=\sum\alpha_j q'_j$。$j=0$ $j=0$因此，

$$Pq_i=PA_1q_{i-1}=P\sum_{j=0}^{n-1}\alpha_jA_1q_j=\sum_{j=0}^{n-1}\alpha_jPq_{j+1}=\sum_{j=0}^{n-1}\alpha_jq'_{j+1}=\sum_{j=0}^{n-1}\alpha_jA_2q'_j$$
$$=A_2q'_{i-1}=q'_i \quad (1)$$

对所有 $i\geq n+1$ 都成立，这一点需要证明。

以下示例表明，前述定理在这样的意义下是紧的：仅考虑 $\mathfrak{w}$-标签的前 n 列并不总是足够。

Example 1 设 P_n 表示一条包含 n 个节点的路径，设 P'_n 表示一条包含 n 个节点且在其中一个端点上带有自环的路径。要区分两张图中没有自环的两个端点，我们需要考虑 $\mathfrak{w}$-标签的前 $(n+1)$ 列，因为它们的标签不会更早出现差异。

请注意，定理 1 甚至以如下更强的形式成立，该形式根据 $\mathfrak{w}$-标签的秩给出了所需列数。

定理 **2** *For every graph pair G_1,G_2 and for all $i_1\in V_1,i_2\in V_2$, $\mathfrak{w}_{G_1}(i_1)_{\mathrm{P}}=\mathfrak{w}_{G_2}(i_2)$ if and only if $\mathfrak{w}_{G_1}|_{s(i_1)}(i_1)_{\mathrm{P}}=\mathfrak{w}_{G_2}|_{s(i_2)}(i_2)$ where $s(i_1)=r(\mathfrak{w}_{G_1}(i_1))+1$and $s(i_2)=r(\mathfrak{w}_{G_2}(i_2))+1$.*

证明与定理 1 的证明类似，因此省略。结合引理 1 和定理2，可知只需逐列生成 $\mathfrak{w}$-标签，一旦当前列与前面的列线性相关，就立即停止。下面的定理给出了 $\mathfrak{w}$-标签最大秩的上界，因而也给出了需要计算的列数上界。首先，考虑以下记号。令 $\mathrm{diam}(G,i)$ 表示从节点 i 出发的最长最短路径，形式化地说，$\mathrm{diam}(G,i):=\max\{\mathrm{dist}(i,j):\ j\in V_G\}$，其中 $\mathrm{dist}(i,j)$ 是图 G 中节点 i 和 j 之间的距离。令 R、p 和 $\mathrm{diam}(G)$ 分别表示节点标签的最大秩、不同特征值的数量以及 G 的直径。

定理 **3** *If G is connected, then $p\geq R\geq\mathrm{diam}(G)+1$.*

Proof 令 Q 表示秩最大的节点标签，即 $r(Q)=R$。根据引理 1，Q 的前 R 列线性无关，这意味着矩阵 $I,A,A^2,\ldots,A^{R-1}$ 也线性无关。

众所周知，实对称矩阵 A^p ~ ~ ~ ~的最小多项式为 $m_A(x)=\prod_{i=1}(x-\lambda_i)$，其中 $\lambda_1,\lambda_2,\ldots,\lambda_p$ 是 A 的互异特征值。因此，$p=\deg(m_A)$，从而 $I,A,A^2,\ldots,A^p$ 线性相关，这意味着 $p\geq R$。

为了证明 $R\geq\mathrm{diam}(G)+1$，注意到对于任意节点i，都有 $r(\mathfrak{w}_G(i))>\mathrm{diam}(G,i)$；也就是说，$\mathfrak{w}_G(i)$ 的秩大于从节点 i 出发的最长最短路径的长度。将此结论应用于实现图的直径的节点 i（即 $\mathrm{diam}(G)=\mathrm{diam}(G,i)$），可得 $R\geq r(\mathfrak{w}_G(i))>\mathrm{diam}(G,i)=\mathrm{diam}(G)$。

从现在起， $\mathfrak{w}_G$ 可能指 $\mathfrak{w}_G|_{n+1}$ ，也可能指无限的 $\mathfrak{w}$-标签。注意，给定节点 i 的标签 $\mathfrak{w}_G|_{n+1}(i)$ 可以通过直接的动态规划方法，在 $\mathcal{O}(nm)$ 次运算内计算出来。此外，可以证明，出现的数值所需的位数相对于图的规模呈多项式增长。因此，通过对两个图的 $\mathfrak{w}$-标签进行排序，只需 $\mathcal{O}(n^2m + n^3 \log(n))$ 步即可判定这两个图是否 $\mathfrak{w}$-等价。

2.2 谱结果

本节中，我们考察 $\mathfrak{w}$-等价性与图的谱或特征子空间之间的关系。以下技术性观察将有助于我们在证明定理 4 中说明非余谱图对并非 $\mathfrak{w}$-等价。

Claim 如果 $\mathfrak{w}_{G_1}(i)_p = \mathfrak{w}_{G_2}(i')$，那么对于所有 $l \geq 0$，从 $i \in V_1$ 和 $i' \in V_2$ 出发、长度为 l 的闭游走数量相同。

Proof 根据定义，存在一个置换矩阵 P ，使得 $P\mathfrak{w}_{G_1}(i) = \mathfrak{w}_{G_2}(i')$。注意， $\mathfrak{w}_{G_1}(i)$ 和 $\mathfrak{w}_{G_2}(i')$ 的第一列规定了 P 将 $\mathfrak{w}_{G_1}(i)$ 的 i^{th} 行映射到 $\mathfrak{w}_{G_2}(i')$ 的 i'^{th} 行，这意味着对于所有 $l \geq 0$，从 $i \in V_1$ 和 $i' \in V_2$ 出发的闭游走数量相同。

回想一下， $U \in \mathbb{R}^{n \times n}$ 是 G的特征矩阵，而 u_{ij} 表示 U的第 i^{th} 列中的第 j^{th} 个元素。

引理 2 *For all $i, j \in \{1, \dots, n\}$ and for all $l \geq 0$, $(A^l)_{ij} = \sum_{k=1}^{n} u_{ki} u_{kj} \lambda_k^l$ holds, where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of G. The right-hand side of this equation will be referred to as the* 特征分解。

*Proof*根据定义， U是正交归一的，且 $AU = U\,\mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$。显然， $A = U\,\mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)U^{-1}$成立，因此 $A^l = U\,\mathrm{diag}(\lambda_1^l, \lambda_2^l, \dots, \lambda_n^l)U^{-1}$。因此，对于任意节点对 $i, j \in \{1, \dots, n\}$， $(A^l)_{ij} = \sum_{k=1}^{n} u_{ki} u_{kj} \lambda_k^l$。

以下观察可由该引理直接推出。

推论 **1** *For all $i, j \in \{1, \dots, n\}$ and $l \geq 0$, there exist $\beta_1^{ij}, \beta_2^{ij}, \dots, \beta_p^{ij} \in \mathbb{R}$ for which $(A^l)_{ij} = \sum_{m=1}^{p} \beta_m^{ij} \tilde{\lambda}_m^l$, where $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_p$ are the distinct non-zero eigenvalues of G. The right-hand side of this equation will be referred to as the* 聚合特征分解。

*Proof*根据引理2， $(A^l)_{ij} = \sum_{k=1}^{n} u_{ki} u_{kj} \lambda_k^l$对于 $l \geq 0$ 和 $i, j \in \{1, \dots, n\}$.显然， $\beta_m^{ij} := \sum_{k:\lambda_k = \lambda_m} u_{ki} u_{kj}$是一个恰当的选择，其中 $i, j \in \{1, \dots, n\}$ 和 $m \in \{1, \dots, p\}$。

以下定理表明，非余谱图并非 $\mathfrak{w}$ 等价。

定理 **4** *If G_1 and G_2 are $\mathfrak{w}$-equivalent, then the spectra of G_1 and G_2 arethe same.*

Proof 证明分为两个步骤。

步骤 1：我们证明 G_1 和 G_2 的非零特征值集合相同。

引理 **3** *Coefficient β_k^{ii} in the aggregated eigen decomposition is zero if it corresponds to a non-zero eigenvalue of exactly one of G_1 and G_2 for all $i, k \in \{1, \ldots, n\}$.*

Proof 设 $\lambda_1, \lambda_2, \ldots, \lambda_r, \theta_{r+1}, \ldots, \theta_p$ 和 $\lambda_1, \lambda_2, \ldots, \lambda_r, \tilde{\mu}_{r+1}, \ldots, \tilde{\mu}_q$ 分别表示 G_1 和 G_2 的所有不同非零特征值，其中 $\lambda_1, \lambda_2, \ldots, \lambda_r$ 是两个图共有的非零特征值，而 $\theta_{r+1}, \ldots, \theta_p, \tilde{\mu}_{r+1}, \ldots, \tilde{\mu}_q$ 两两不同。

为简便起见，假设节点已重新编号，使得恒等映射给出一个 $\mathfrak{w}$-等价关系，即对每个节点 i，$\mathfrak{w}_{G_1}(i) = \mathfrak{w}_{G_2}(i)$。

对于任意 i 和 j，存在系数 $\alpha_1, \alpha_2, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_q$，使得

$$(A_1^l)_{ij} = \sum_{k=1}^{r} \alpha_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} \alpha_k \tilde{\theta}_k^l \tag{2}$$

and

$$(A_2^l)_{ij} = \sum_{k=1}^{r} \beta_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{q} \beta_k \tilde{\mu}_k^l \tag{3}$$

根据推论 1，对所有 $l \geq 1$均成立。由于两个图是 $\mathfrak{w}$-等价的，可得

$$\sum_{k=1}^{r} \alpha_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} \alpha_k \tilde{\theta}_k^l = (A_1^l)_{ii} = (A_2^l)_{ii} = \sum_{k=1}^{r} \beta_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{q} \beta_k \tilde{\mu}_k^l$$

对所有 $i \in \{1, \ldots, n\}$和 $l \geq 1$均成立，其中第二个方程由断言 2.2得出。减去等式右边，得到

$$\sum_{k=1}^{r} (\alpha_k - \beta_k) \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} \alpha_k \tilde{\theta}_k^l - \sum_{k=r+1}^{q} \beta_k \tilde{\mu}_k^l = 0 \tag{4}$$

对所有 $l \geq 1$均成立。令 $m := p + q - r$，并考虑关于$l \in \{1, \ldots, m\}$的以下线性方程。

$$\sum_{k=1}^{r} x_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} x_k \tilde{\theta}_k^l + \sum_{k=r+1}^{q} x_{p+k-r} \tilde{\mu}_k^l = 0, \tag{5}$$

其中

$$x_s := \begin{cases} \alpha_s - \beta_s & \text{if } 1 \leq s \leq r, \\ \alpha_s & \text{if } r+1 \leq s \leq p, \\ -\beta_{r+s-p} & \text{if } p+1 \leq s \leq p+q-r, \end{cases} \tag{6}$$

对于所有 $s \in \{1, \dots, m\}$。该线性方程组的矩阵为

$$M := \begin{bmatrix} \tilde{\lambda}_1^1 & \dots & \tilde{\lambda}_r^1 & \tilde{\theta}_{r+1}^1 & \dots & \tilde{\theta}_p^1 & \tilde{\mu}_{r+1}^1 & \dots & \tilde{\mu}_q^1 \\ \tilde{\lambda}_1^2 & \dots & \tilde{\lambda}_r^2 & \tilde{\theta}_{r+1}^2 & \dots & \tilde{\theta}_p^2 & \tilde{\mu}_{r+1}^2 & \dots & \tilde{\mu}_q^2 \\ \tilde{\lambda}_1^3 & \dots & \tilde{\lambda}_r^3 & \tilde{\theta}_{r+1}^3 & \dots & \tilde{\theta}_p^3 & \tilde{\mu}_{r+1}^3 & \dots & \tilde{\mu}_q^3 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\lambda}_1^m & \dots & \tilde{\lambda}_r^m & \tilde{\theta}_{r+1}^m & \dots & \tilde{\theta}_p^m & \tilde{\mu}_{r+1}^m & \dots & \tilde{\mu}_q^m \end{bmatrix}. \tag{7}$$

注意，$M = M' \mathrm{diag}(\tilde{\lambda}_1^1, \dots, \tilde{\lambda}_r^1, \tilde{\theta}_{r+1}^1, \dots, \tilde{\theta}_p^1, \tilde{\mu}_{r+1}^1, \dots, \tilde{\mu}_q^1)$，其中 M'表示以下范德蒙矩阵。

$$M' := \begin{bmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ \tilde{\lambda}_1^1 & \dots & \tilde{\lambda}_r^1 & \tilde{\theta}_{r+1}^1 & \dots & \tilde{\theta}_p^1 & \tilde{\mu}_{r+1}^1 & \dots & \tilde{\mu}_q^1 \\ \tilde{\lambda}_1^2 & \dots & \tilde{\lambda}_r^2 & \tilde{\theta}_{r+1}^2 & \dots & \tilde{\theta}_p^2 & \tilde{\mu}_{r+1}^2 & \dots & \tilde{\mu}_q^2 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\lambda}_1^m & \dots & \tilde{\lambda}_r^m & \tilde{\theta}_{r+1}^m & \dots & \tilde{\theta}_p^m & \tilde{\mu}_{r+1}^m & \dots & \tilde{\mu}_q^m \end{bmatrix} \tag{8}$$

因此，$\det(M) = \det(M') \prod_{k=1}^{r} \tilde{\lambda}_k \prod_{k=r+1}^{p} \tilde{\theta}_k \prod_{k=r+1}^{q} \tilde{\mu}_k \neq 0$，所以唯一的解是 $x \equiv 0$，也就是说，

$$\begin{cases} \alpha_s = \beta_s & \text{if } 1 \leq s \leq r, \\ \alpha_s = 0 & \text{if } r+1 \leq s \leq p, \\ \beta_{r+s-p} = 0 & \text{if } p+1 \leq s \leq p+q-r \end{cases} \tag{9}$$

对所有 $s \in \{1, \dots, m\}$均成立。

设 $\lambda^* \neq 0$ 为一个特征值，它只对应于其中一个图，例如 G_1。接下来，我们证明存在一个节点 $i \in V_1$，使得在推论 1 给出的 $(A_1^l)_{ii}$ 的聚合特征分解中，λ^* 的系数非零——这与引理 3 矛盾。设 $\tilde{m}$ 为满足 $\tilde{\lambda}_{\tilde{m}} = \lambda^*$ 的唯一索引。根据推论 1，在从节点 i 出发的闭游走数量的情况下，$\lambda_{\tilde{m}}$ 的系数为 $\beta_{\tilde{m}}^{ii} = \sum_{k:\ \lambda_k = \tilde{\lambda}_{\tilde{m}}} u_{ki} u_{ki}$。设 m 为满足$\lambda_m = \lambda_{\tilde{m}}$的一个索引，并设 i 满足 $u_{mi} u_{mi} > 0$（至少存在这样一个索引，因为 $u_m u_m = 1$）。注意，$\beta_{\tilde{m}}^{ii} \geq u_{mi} u_{mi} > 0$ 成立，因此节点 i 满足要求，这与引理 3 矛盾。

步骤 2：我们证明 G_1 和 G_2 中特征值的重数相同。只需证明非零特征值的重数相同，因为这也意味着 G_1 和 G_2 中零特征值的重数相同。令 $\tau_i^{(k)}$ 表示 $\tilde{\lambda}_i$ 在 $G_k (k = 1, 2)$ 中的重数，其中 λ_1、.....、λ_p 是 G_1 和 G_2 的共有特征值。

根据引理2，长度为 l 的 G_k 闭游走数量之和为 $\sum_{j=1}^{p} \tau_j^{(k)} \tilde{\lambda}_j^l$，$(l \geq 1)$。由于 G_1 和 G_2 是 $\mathfrak{w}$-等价的，断言 2.2 适用，因此两个图中长度为 l 的闭游走数量之和

对所有 l 都相同，也就是说，$\sum_{j=1}^{p} \tau_j^{(1)} \tilde{\lambda}_j^l = \sum_{j=1}^{p} \tau_j^{(2)} \tilde{\lambda}_j^l$ 对所有 $l \geq 1$ 都成立。减去等式右侧，得到

$$\sum_{j=1}^{p} (\tau_j^{(1)} - \tau_j^{(2)}) \tilde{\lambda}_j^l = 0 \tag{10}$$

对所有 $l \geq 1$均成立。考虑关于 $l \in \{1, \ldots, p\}$的这些方程，并令 $x_j := \tau_j^{(1)} - \tau_j^{(2)}$ 对所有 $j \in \{1, \ldots, p\}$均成立。与步骤 1类似，该方程组的矩阵行列式非零，因此唯一解为 $x \equiv 0$，这意味着对所有 $j \in \{1, \ldots, p\}$，$\tau_j^{(1)} = \tau_j^{(2)}$ 均成立。因此，两个图中的每个非零特征值都具有相同的重数，从而 0 的重数也相同，证明完毕。

注意，$\mathfrak{w}$-等价性比谱能够区分更多图对。例如，长度为 6 的环与两个不相交三角形是众所周知的余谱图，但它们显然并不具有 $\mathfrak{w}$-等价性。

下面我们将证明，即使两个非同构图是余谱图，如果它们的特征子空间差异足够大，它们也可能不具有 $\mathfrak{w}$-等价性：

定理 **5** *Let G_1 and G_2 be cospectral graphs with single eigenvalues. If one of the eigenmatrices has a row that contains non-zero elements only, then the two graphs are $\mathfrak{w}$-equivalent if and only if the graphs are isomorphic.*

Proof 显然，只需证明若 G_1 和 G_2 是 $\mathfrak{w}$-等价的，则它们同构。为此，我们将证明存在一个置换矩阵 Π，使得 $\Pi A_1 \Pi^T = A_2$。回顾一下，$U = (u_1, \ldots, u_n)$和 $V = (v_1, \ldots, v_n)$ 分别表示 G_1 和 G_2的特征矩阵，即 $A_1 = U \mathrm{diag}(\lambda_1, \ldots, \lambda_n) U^T$以及 $A_2 = V \mathrm{diag}(\lambda_1, \ldots, \lambda_n) V^T$。置换矩阵 Π对应一个同构，当且仅当 $\Pi U \mathrm{diag}(\lambda_1, \ldots, \lambda_n) U^T \Pi^T = V \mathrm{diag}(\lambda_1, \ldots, \lambda_n) V^T$成立；而这又当且仅当存在一个矩阵 $S = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$使得 $\Pi U = VS$，其中 $\sigma_i \in \{-1, 1\}$。因此，只需证明存在这样的矩阵 Π 和 S。

不失一般性，假设 U 的第 i^* 行由非零元素组成。根据 $\mathfrak{w}$-等价的定义，存在一个置换 π，使得 $(A_1^l)_{i^* j} = (A_2^l)_{\pi(i^*)\pi(j)}$，因此对所有 $j \in \{1, \ldots, n\}$，都有 $u_{ki^*} u_{kj} = v_{k\pi(i^*)} v_{k\pi(j)}$。显然，$V$ 的第 $\pi(i^*)$ 行由非零元素组成。令 $S := \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$，其中$\sigma_k := \mathrm{sgn}(u_{ki^*}) \, \mathrm{sgn}(v_{k\pi(i^*)}) \in \{-1, 1\}$，并令 $\Pi = \begin{cases} 1 : \text{若} \\ \end{cases}$ $\pi(j) = i$，否则为 0。接下来，我们证明 $\Pi U = VS$。左右两边在位置 (j, k) 处的值分别为 $u_{k\pi^{-1}(j)}$ 和 $\sigma_k v_{kj}$。注意到，对所有 j，$u_{k\pi^{-1}(j)} = \sigma_k v_{kj}$ 当且仅当对所有 j，$u_{kj} = \sigma_k v_{k\pi(j)}$，而这又等价于对所有 j，$u_{ki^*} u_{kj} = v_{k\pi(i^*)} v_{k\pi(j)}$，因为 $u_{ki^*} = \sigma_k v_{k\pi(i^*)}$ 和 $\sigma_k^2 = 1$，且 π 的选取满足对所有 j，$u_{ki^*} u_{kj} = v_{k\pi(i^*)} v_{k\pi(j)}$。因此 $\Pi U = VS$，定理得证。

下一个引理将有助于证明定理 6 和定理 8。

引理 **4** *If G_1 and G_2 are $\mathfrak{w}$-equivalent graphs and the nodes of G_2 are re-indexed in such a way that $\mathfrak{w}_{G_1}(i)_{\mathrm{p}} = \mathfrak{w}_{G_2}(i)$ for all $i \in \{1, \ldots, n\}$, then for any*

single eigenvalue, the corresponding normalized eigenvectors in the two graphs are element-wise the same up to sign.

Proof 根据定理 4，可知 G_1 和 G_2 是余谱的。根据定义和命题2，$\mathfrak{w}_{G_1}|_{n+1}(i) = \mathfrak{w}_{G_2}|_{n+1}(i)$ 意味着

$$\sum_{k=1}^{n} u_{ki}u_{ki}\lambda_k^l = (A_1^l)_{ii} = (A_2^l)_{ii} = \sum_{k=1}^{n} v_{ki}v_{ki}\lambda_k^l \tag{11}$$

对所有 $i \in \{1,\dots,n\}$ 和 $l \geq 0$都成立，这进而意味着 $u_{ki}u_{ki} = v_{ki}v_{ki}$ 成立，其中 $i \in \{1,\dots,n\}$ 是任意的，而 k 满足 λ_k 是单一特征值。也就是说，$|u_{ki}| = |v_{ki}|$ 若 $i \in \{1,\dots,n\}$ 且 k满足 λ_k 是单一特征值，则.....这就完成了证明。

定理 **6** *Let G_1 and G_2 be cospectral with single eigenvalues. If $\{u_{ik} : k \in \{1,\dots,n\}\}^{\#} \neq \{-u_{ik} : k \in \{1,\dots,n\}\}^{\#}$ and $\{v_{ik} : k \in \{1,\dots,n\}\}^{\#} \neq \{-v_{ik} : k \in \{1,\dots,n\}\}^{\#}$ for all $i \in \{1,\dots,n\}$, then the two graphs are $\mathfrak{w}$-equivalent if and only if they are isomorphic.*

Proof 如果 G_1 和 G_2 同构，那么显然它们是 $\mathfrak{w}$-等价的。为证明反方向，设 G_1 和 G_2 是 $\mathfrak{w}$-等价的。设 $w_1, w_2 \in \mathbb{R}^n$ 是满足 $\{w_{1k} : k \in \{1,\dots,n\}\}^{\#} \neq \{w_{2k} : k \in \{1,\dots,n\}\}^{\#}$的向量，并令 $w_1 \overset{\text{L}}{\succ} w_2$表示将坐标按非递增顺序排列后，$w_1$ 的字典序大于 w_2。

不失一般性，可以假设对所有 $i \in \{1,\dots,n\}$，$u_i \overset{\text{L}}{\succ} -u_i$ 和 $v_i \overset{\text{L}}{\succ} -v_i$都成立。设 $\pi : V_1 \to V_2$是一个 $\mathfrak{w}$-等价。由引理4，$|u_{ki}| = |v_{k\pi(i)}|$对所有 $k \in \{1,\dots,n\}$和 $i \in \{1,\dots,n\}$都成立。反设存在一个指标 k^* 和 i^*，使得 $u_{k^*i^*} \neq v_{k^*\pi(i^*)}$。这意味着 $u_{k^*i^*} = -v_{k^*\pi(i^*)}$。令 π^* 表示节点 i^*的双射，使得无论 0 是否为特征值，对所有 $j,k \in \{1,\dots,n\}$，$u_{ki^*}u_{kj} = v_{k\pi^*(i^*)}v_{k\pi^*(j)}$均成立。显然，可以设定 π^* 使得 $\pi^*(i^*) = \pi(i^*)$。因此可得，对所有 $j \in \{1,\dots,n\}$，$u_{k^*i^*}u_{k^*j} = v_{k^*\pi^*(i^*)}v_{k^*\pi^*(j)}$成立，这意味着 $-u_{k^*} = \pi v_{k^*}$。但此时 $u_{k^*} \overset{\text{L}}{\succ} -u_{k^*} = \pi^* v_{k^*} \overset{\text{L}}{\succ} -v_{k^*} = \pi^{*-1}u_{k^*}$，因此 G_1和 G_2 确实同构。

定义 3 若图的每个特征值的重数均为 1，且其特征矩阵 $\mathbb{1}U$ 没有零坐标，则该图为 *friendly* [20]，其中 U 是该图的特征矩阵。

推论 **2** *Two friendly graphs G_1 and G_2 are isomorphic if and only if they are $\mathfrak{w}$-equivalent.*

*Proof*假设 $\mathbb{1}u_i \neq 0$ 和 $\mathbb{1}v_i \neq 0$ 意味着 $\{u_{ik} : k \in \{1,\dots,n\}\}^{\#} \neq \{-u_{ik} : k \in \{1,\dots,n\}\}^{\#}$，且 $\{v_{ik} : k \in \{1,\dots,n\}\}^{\#} \neq \{-v_{ik} : k \in \{1,\dots,n\}\}^{\#}$ 对所有 $i \in \{1,\dots,n\}$ 均成立，因此可以应用定理 6。

回顾以下著名定理。

定理 7（佩龙-弗罗贝尼乌斯） *Let the graph G be connected and have at least two nodes. The largest eigenvalue λ_1 of the adjacency matrix of G is positive, has multiplicity one, and $\lambda_1 \geq |\lambda|$ for every eigenvalue λ. In addition, the eigenvector corresponding to λ_1 can be chosen strictly positive.*

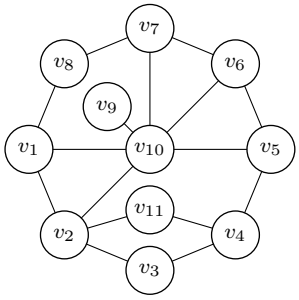


图1： $\mathfrak{s}^k$标记严格强于 $\mathfrak{w}$标记的一个例子。这里， $\mathfrak{w}_G(v_1) = \mathfrak{w}_G(v_5)$，但 $\mathfrak{s}^1_G(v_1) \neq \mathfrak{s}^1_G(v_5)$。

定理 7 中与最大正特征值对应的正归一化特征向量将称为 *Perron-Frobenius eigenvector* of G。下述定理可由引理4直接推出。

定理 **8** *Let G_1 and G_2 be connected cospectral graphs on at least two nodes. If the Perron-Frobenius eigenvectors of G_1 and G_2 are different, then G_1 and G_2 are not $\mathfrak{w}$-equivalent.*

图的佩龙-弗罗贝尼乌斯特征向量决定了其关于无限随机游走的 *invariant distribu-tion* ，因此，前述定理指出，如果两个图的不变分布不同，那么它们就不是$\mathfrak{w}$等价的。

3 完美聚合标记

本节介绍 $\mathfrak{w}$标记的一种细化。首先，令

$$\mathfrak{s}^1_G(i_1)_{il} := \begin{cases} (\emptyset, \{\delta_{ii_1}\}) & \text{if } l = 0, \\ (\mathfrak{s}^1_G(i_1)_{i,l-1}, \{\mathfrak{s}^1_G(i_1)_{i',l-1} : i' \in N_G(i)\}^{\#}) & \text{otherwise} \end{cases}$$

对于一个节点 $i_1 \in V$和 $i \in \{1, \ldots, n\}$, $l \in \mathbb{Z}_+$。本质上，我们将邻居的值收集为一个多重集，而不是像处理 $\mathfrak{w}$ 标签时那样把它们相加。显然，生成这些多重集保留的信息不会少于将这些值相加——即使加数不同，它们也可能得到相同的和。事实上，与 $\mathfrak{w}$ 相比，使用 $\mathfrak{s}^1$ 能收集到严格更多的信息，图1给出了一个例子。

我们更进一步，对 $\mathfrak{s}^1$ 作出推广； $\mathfrak{s}^1$ 用于初始化第一次迭代，也就是情形 $l = 0$，其初始化方式稍微复杂一些。

记号 **9** *Let $\mathfrak{s}^k_G(i_1, \ldots, i_k)$ be an $n \times \mathbb{Z}_+$ matrix whose position (j, l) describes the walks of length at most l between nodes $\{i_1, \ldots, i_k\}$ and i. Formally, defines $\mathfrak{s}^k_G(i_1, \ldots, i_k)_{il}$ as*

$$\begin{cases} (\emptyset, \{\sum_{q=1}^{k} q\delta_{ii_q}\}) & \text{if } l = 0, \\ (\mathfrak{s}^k_G(i_1, \ldots, i_k)_{i,l-1}, \{\mathfrak{s}^k_G(i_1, \ldots, i_k)_{i',l-1} : i' \in N_G(i)\}^{\#}) & \text{otherwise} \end{cases} \tag{12}$$

for all $j \in V$ *and for all* $l \geq 0$.

注意，矩阵 $\mathfrak{s}^k_G(i_1,..,i_k)$ 的第一列对应长度为零的游走，因此其索引为零。列 l 将记为 $\mathfrak{s}^k_G(i_1,..,i_k)_{\bullet l}$。回想一下，$\mathfrak{s}^k_G(i_1,..,i_k)\big|_q$ 表示矩阵 $\mathfrak{s}^k_G(i_1,..,i_k)$ 的前 q 列，也就是说，它描述了长度不超过 $(q-1)$ 的游走。

简单的归纳证明表明，只需考虑前 $(n+1)$ 列，这与 $\mathfrak{w}$-标签的情形类似。

Claim 设给定一个图对 G_1, G_2、互不相同的节点 $i_1,\dots,i_k \in V_1, j_1,\dots,j_k \in V_2$ 以及一个整数 $k \geq 1$。那么，

$$\mathfrak{s}^k_{G_1}(i_1,\dots,i_k) \overset{\mathrm{p}}{=} \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)$$

成立当且仅当

$$\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)\Big|_{n+1} \overset{\mathrm{p}}{=} \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)\Big|_{n+1},$$

其中 $n = |V_1| = |V_2|$。

从现在起，$\mathfrak{s}^k_G(i_1,\dots,i_k)$ 也可能指 $\mathfrak{s}^k_G(i_1,\dots,i_k)\big|_{n+1}$，也就是说，指其前 $(n+1)$ 列。例1表明，前述命题在以下意义上是紧的：考虑前 n 列并不足够。

请注意，$\mathfrak{s}^k_G(i_1,\dots,i_k)_{jl}$ 的大小相对于 n 可能呈指数级增长。实际上，可以通过使用 SHA512 对出现的数据进行哈希来解决这个问题——这也能生成简短的图指纹。还可以通过如下方式将标签替换为小整数，在多项式时间内验证两个给定图的 $\mathfrak{s}^k$-标签是否相同。令空集对应的数为零，而集合 $\{p\}$ 对应的数对所有 $p \in \{1,\dots,k\}$ 均为 p。在递归（12）中，我们用对应的数替换所有已分配数值的标签。在标记过程中某个标签首次出现时，就令其表示数为下一个自然数。由于计算 $\mathfrak{s}^k_G(i_1,\dots,i_k)_{\bullet l}$ 时，所有标签 $\mathfrak{s}^k_G(i_1,\dots,i_k)_{\bullet l-1}$ 都已对应一个数——因此这些标签都是由小整数构成的多重集——整个过程可在多项式时间内完成。我们强调，整个过程中每个出现的标签都必须且只能对应一个数，并且在两个图中该标签每次出现时都必须始终使用这个数。

记法 **10** *For* $q = k-1,\dots,0$, *let* $\mathfrak{s}^k_G(i_1,\dots,i_q) := \{\mathfrak{s}^k_G(i_1,\dots,i_q,i) :\ i \in V \setminus \{i_1,\dots,i_q\}\}^{\#}$.

这种递归记法旨在导出同构不变量指纹 $\mathfrak{s}^k_G$，它是针对 $q=0$ 得到的（我们省略 $\mathfrak{s}^k_G$ 后面的空括号）。例如，若 $k=2$，则对于 $q=1$，可得到 $\mathfrak{s}^k_G(v) = \{\mathfrak{s}^k_G(v,v') : v' \in V \setminus \{v\}\}^{\#}$，而对于 $q=0$，指纹为 $\mathfrak{s}^k_G = \{\mathfrak{s}^k_G(v) :\ v \in V\}^{\#}$。

以下两个命题可由定义直接得出。

Claim 对于所有整数 $k \geq 1$ 和 $q \in \{0,\dots,k-1\}$，若 $\mathfrak{s}^k_{G_1}(i_1,\dots,i_q) \neq \mathfrak{s}^k_{G_2}(j_1,\dots,j_q)$ 对某些节点 $i_1,\dots,i_q \in V_1$ 和 $j_1,\dots,j_q \in V_2$ 成立，则 G_1 与 G_2 之间不存在同构，使得对每个 $r \in \{1,\dots,q\}$，节点 i_r 映射到节点 j_r。

Claim 对于所有整数 $k \geq 1$、 $q \in \{0, \ldots, k\}$ 以及任意 $i_1, \ldots, i_q \in V_1$ 和 $j_1, \ldots, j_q \in V_2$，若 $\mathfrak{s}^k_{G_1}(i_1, \ldots, i_q) \neq \mathfrak{s}^k_{G_2}(j_1, \ldots, j_q)$，则 $\mathfrak{s}^{k+1}_{G_1}(i_1, \ldots, i_q) \neq \mathfrak{s}^{k+1}_{G_2}(j_1, \ldots, j_q)$。

定义 4 两个图 G_1 和 G_2 是 $\mathfrak{s}^k$-*equivalent* 当且仅当 $\mathfrak{s}^k_{G_1} = \mathfrak{s}^k_{G_2}$。

作为前一个命题的特例，可以得到：如果两个图不是$\mathfrak{s}^k$等价的，那么它们也不是 $\mathfrak{s}^{k+1}$等价的。也就是说，随着 k 增大，能够区分的非同构图对越来越多。注意，当 k达到节点数时，图就能由 $\mathfrak{s}^k$唯一确定。

Remark 1 对于任意给定的常数 $k \in \mathbb{N}$，都可以在多项式时间内验证两个图是否是 $\mathfrak{s}^k$等价的。

3.1 完美聚合的区分能力

本节考察上述概念在树、平面图以及受节点连通性参数约束的图上的区分能力。

定理 **11** *Two trees are isomorphic if and only if they are* $\mathfrak{s}^1$*-equivalent.*

Proof 给定两棵 $\mathfrak{s}^1$-等价树 $G_1 = (V, E_1)$ 和 $G_2 = (V, E_2)$，我们证明它们同构。对于一条边 $(r, p) \in E_i$，令 $T_i(r, p) = (V_i(r, p)\ E_i(r, p))$ 表示 G_i 中由 $(V, E_i \setminus \{(r, p)\})$ 的连通分量构成且包含节点 r 的子树。

通过归纳法，我们证明：对于任意两条边 $(r_1, p_1) \in E_1$ 和 $(r_2, p_2) \in E_2$，若$\mathfrak{s}^1_{G_1}(r_1)\big|_{n'+1} \mathrm{P} = \mathfrak{s}^1_{G_2}(r_2)\big|_{n'+1}$ 且 $\mathfrak{s}^1_{G_1}(p_1)\big|_{n'+1} \mathrm{P}= \mathfrak{s}^1_{G_2}(p_2)\big|_{n'+1}$ 对于 $n' = |V_1(r_1, p_1)|$成立，则 $T_1(r_1, p_1)$ 和 $T_2(r_2, p_2)$ 同构。显然，如果 $n' = 1$—— 换句话说，r_1 是 G_1 中的一个叶节点 —— 那么 r_2 也必定是 G_2 中的一个叶节点。

否则，可得 $\{\mathfrak{s}^1_{G_1}(i)\big|_{n'} : i \in N_{G_1}(r_1)\}^{\#} = \{\mathfrak{s}^1_{G_2}(i)\big|_{n'} : i \in N_{G_2}(r_2)\}^{\#}$。因此，$r_1$ 和 r_2 拥有相同数量的邻居，并且存在一一映射 $\phi : N_{G_1}(r_1) \to N_{G_2}(r_2)$，使得对于每个 $v \in N_{G_1}(r_1)$，v 和 $\phi(v)$的标签在前 n' 列上相同。因此，根据归纳法假设，对于所有 $v \in N_{G_1}(r_1) \setminus p_1$，$T_1(v, r_1)$ 和 $T_2(\phi(v)\ r_2)$ 是同构子树。由此立即可得 $T_1(r_1, p_1)$ 和 $T_2(r_2, p_2)$ 同构。

为了完成定理的证明，任取一个叶节点 $r_1 \in V_1$ 和一个满足 $\mathfrak{s}^1_{G_1}(r_1) \mathrm{P} = \mathfrak{s}^1_{G_2}(r_2)$ 的节点 $r_2 \in V_2$。节点 r_2 也是叶节点，并且对于它们的邻居 $p_1 \in V_1$ 和 $p_2 \in V_2$，有 $\mathfrak{s}^1_{G_1}(p_1)\big|_{|V_1|-1} \mathrm{P}= \mathfrak{s}^1_{G_2}(p_2)\big|_{|V_1|-1}$。将上述命题应用于 r_1, p_1, r_2, p_2，即可证明 G_1 和 G_2 同构。

请注意，这个证明为树提供了一种新的多项式时间同构算法，不过，已有一种更高效的算法 [1]。

定义 5 相对于 $\mathfrak{s}^k_G(i_1, \ldots, i_k)$ 的 *pattern of a walk* $(h_1, \ldots, h_p) \in V^p$ 为 $(\mathfrak{s}^k_G(i_1, \ldots, i_k)\ h_1 0, \ldots, \mathfrak{s}^k_G(i_1, \ldots, i_k) h_p 0)$。

定理 **12** *Let* $\mathfrak{s}^k_{G_1}(i_1, \ldots, i_k) = \mathfrak{s}^k_{G_2}(j_1, \ldots, j_k)$ *and let* i, j *be two nodes suchthat* $\mathfrak{s}^k_{G_1}(i_1, \ldots, i_k)_i = \mathfrak{s}^k_{G_2}(j_1, \ldots, j_k)_j$*. Then the numbers of walks with any given pattern ending at node* i *in* G_1 *and at node* j *in* G_2 *are the same.*

Proof 首先注意，$\mathfrak{s}^k_{G_1}(i_1,\dots,i_k) = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)$ 意味着存在一个双射 $\pi\colon V_1 \to V_2$，使得对所有 $i' \in V_1$，$\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i'} = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{\pi(i')}$ 成立。由于 $\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_i = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_j$，我们还可以规定 $\pi(i) = j$。我们通过对游走长度进行归纳，证明以 i'（位于 G_1 中）和 $\pi(i')$（位于 G_2 中）为终点、具有任意给定模式的游走数量相同。根据定义，对于任意长度为一的模式，长度为零的游走数量相同。根据归纳法，假设该结论对长度为 $(l-1)$ 的游走成立。注意，$\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i'} = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{\pi(i')}$ 意味着 $\{\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i'',l-1}\colon i'' \in N_{G_1}(i')\}^{\#} = \{\mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{j'',l-1}\colon j'' \in N_{G_2}(\pi(i'))\}^{\#}$。根据归纳假设，这意味着以 i'（位于 G_1 中）的邻居和 $\pi(i')$（位于 G_2 中）的邻居为终点、长度为 $(l-1)$ 且具有任意给定模式的游走数量相同。以节点 t 为终点、长度为 $l \geq 1$ 的任意游走，都由一条与 t 相连的边 tt' 和一条以 t' 为终点、长度为 $(l-1)$ 的游走组成。将这一观察分别应用于 $t = i'$ 和 $t = j'$，可得以节点 i'（位于 G_1 中）和 $\pi(i')$（位于 G_2 中）为终点、长度为 l 且具有任意给定模式的游走数量相同，这正是需要证明的结论。

定义 6 如果节点的子集 X 是 *k-separator*，且 $|X| \leq k$ 和 $G \setminus X$ 的连通分量数多于 G，则称其为 *k-separator*。

定理 **13** *If $i \in V_1$ is contained in a $(k-1)$-separator of G_1 and $j \in V_2$ is not contained in any k-separators of G_2, then $\mathfrak{s}^k_{G_1}(i) \neq \mathfrak{s}^k_{G_2}(j)$。*

Proof 根据命题 3，可以假设 $|X| = k-1$。令 $i_1,\dots,i_{k-1} \in V_1$ 表示 X 的节点，并令 i_k 为 $V_1 \setminus X$ 中任意一个节点。我们可以假设 $i = i_1$。反设若 $\mathfrak{s}^k_{G_1}(i) = \mathfrak{s}^k_{G_2}(j)$，则存在 $j_1,\dots,j_k \in V_2$，使得 $\mathfrak{s}^k_{G_1}(i_1,\dots,i_k) = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)$，其中 $j_1 = j$。令 C 为 $G_1 \setminus X$ 中不包含 i_k 的任意一个连通分量，并令 i' 为 C 的一个节点。注意，G_1 中连接节点 i_k 和 i' 的任意游走，都会经过 $i_1,\dots,i_{k-1}$ 中至少一个节点。与此同时，G_2 中存在一条游走，连接 j_k 与 $V_2 \setminus \{j_1,\dots,j_{k-1}\}$ 中的任意其他节点，且避开节点 $j_1,\dots,j_{k-1}$，这与定理 12 矛盾。

根据前述定理可知，若 $\mathfrak{s}^k_{G_1}(i) = \mathfrak{s}^k_{G_2}(j)$，则 i 属于一个 $(k-1)$-分隔集当且仅当 j 属于该分隔集，因此可立即得出以下推论。

推论 **3** *If G_1 is k-connected and G_2 is not, then $\mathfrak{s}^k_{G_1} \neq \mathfrak{s}^k_{G_2}$*

定理 **14** *Let G_1, G_2 be two graphs, and let $\overline{G}_1, \overline{G}_2$ be their complements, respectively. Then $\mathfrak{s}^k_{G_1} = \mathfrak{s}^k_{G_2}$ if and only if $\mathfrak{s}^k_{\overline{G}_1} = \mathfrak{s}^k_{\overline{G}_2}$.*

Proof 只需证明若 $\mathfrak{s}^k_{G_1} = \mathfrak{s}^k_{G_2}$，则 $\mathfrak{s}^k_{\overline{G}_1} = \mathfrak{s}^k_{\overline{G}_2}$，因为这也会立即推出反方向成立。设 $i_1,\dots,i_k \in V_1$ 和 $j_1,\dots,j_k \in V_2$ 满足 $\mathfrak{s}^k_{G_1}(i_1,\dots,i_k) = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)$。

首先，我们对 l 进行归纳法证明：若 $i \in V_1$ 和 $j \in V_2$ 满足 $\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{il} = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{jl}$，则 $\mathfrak{s}^k_{\overline{G}_1}(i_1,\dots,i_k)_{il} = \mathfrak{s}^k_{\overline{G}_2}(j_1,\dots,j_k)_{jl}$ 成立。基例 $l = 0$ 显然成立，因为四个图中的初始标签都相同。根据归纳假设，假设该命题对某个 $l \geq 1$ 的 $(l-1)$ 成立。首先注意到，$\mathfrak{s}^k_{G_1}(i_1,\dots,i_k) = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)$ 意味着

$$\{\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i',l-1} : i' \in V_1\}^{\#} = \{\mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{j',l-1} : j' \in V_2\}^{\#}. \tag{13}$$

根据定义，$\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{il} = \mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{jl}$ 意味着

$$\{\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i',l-1} : i' \in N_{G_1}(i)\}^{\#} = \{\mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{j',l-1} : j' \in N_{G_2}(j)\}^{\#} \tag{14}$$

因此，

$$\begin{aligned}
&\{\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i',l-1} : i' \in N_{\overline{G}_1}(i)\}^{\#} \\
&\quad = \{\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i',l-1} : i' \in V_1\}^{\#} \setminus \{\mathfrak{s}^k_{G_1}(i_1,\dots,i_k)_{i',l-1} : i' \in N_{G_1}(i)\}^{\#} \\
&\quad = \{\mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{j',l-1} : j' \in V_2\}^{\#} \setminus \{\mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{j',l-1} : j' \in N_{G_2}(j)\}^{\#} \\
&\quad = \{\mathfrak{s}^k_{G_2}(j_1,\dots,j_k)_{j',l-1} : j' \in N_{\overline{G}_2}(j)\}^{\#}, \qquad (15)
\end{aligned}$$

其中第二个等式由 (13) 和 (14) 得出。由此，$\{\mathfrak{s}^k_{\overline{G}_1}(i_1,\dots,i_k)_{i',\ l-1} : i' \in N_{\overline{G}_1}(i)\}^{\#} = \{\mathfrak{s}^k_{\overline{G}_2}(j_1,\dots,j_k)_{j',\ l-1} : j' \in N_{\overline{G}_2}(j)\}^{\#}$ 由归纳法可得，因此 $\mathfrak{s}^k_{\overline{G}_1}(i_1,\dots,i_k)_{il} = \mathfrak{s}^k_{\overline{G}_2}(j_1,\dots,j_k)_{jl}$ 成立，从而证明了该命题。

我们对 $q = k,\dots,0$ 使用归纳法证明：若 $\mathfrak{s}^k_{G_1}(i_1,\dots,i_q) = \mathfrak{s}^k_{G_2}(j_1,\dots,j_q)$，则 $\mathfrak{s}^k_{\overline{G}_1}(i_1,\dots,i_q) = \mathfrak{s}^k_{\overline{G}_2}(j_1,\dots,j_q)$——当 $q = 0$ 时，这就是定理的陈述。基例 $q = k$ 可由上述陈述立即推出。根据归纳假设，假设该命题对 $(q+1)$ 成立，并证明它对 q 也成立。假设 $\mathfrak{s}^k_{G_1}(i_1,\dots,i_q) = \mathfrak{s}^k_{G_2}(j_1,\dots,j_q)$ 意味着 $\{\mathfrak{s}^k_{G_1}(i_1,\dots,i_q,i) : i \in V_1 \setminus \{i_1,\dots,i_q\}\}^{\#} = \{\mathfrak{s}^k_{G_2}(j_1,\dots,j_q,j) : j \in V_2 \setminus \{j_1,\dots,j_q\}\}^{\#}$。由此，通过归纳法可得 $\{\mathfrak{s}^k_{\overline{G}_1}(i_1,\dots,i_q,i) : i \in V_1 \setminus \{i_1,\dots,i_q\}\}^{\#} = \{\mathfrak{s}^k_{\overline{G}_2}(j_1,\dots,j_q,j) : j \in V_2 \setminus \{j_1,\dots,j_q\}\}^{\#}$，即 $\mathfrak{s}^k_{\overline{G}_1}(i_1,\dots,i_q) = \mathfrak{s}^k_{\overline{G}_2}(j_1,\dots,j_q)$，这正是需要证明的结论。

若图类 $\mathcal{G}$ 满足以下条件：其中任意图的补图也属于 $\mathcal{G}$，则称图类 $\mathcal{G}$ 为 *self-complementary* 。前述定理立即推出以下观察，这将在第3.2节中派上用场。

推论 **4** *Let $\mathcal{G}$ be a self-complementary graph class. Then $\mathfrak{s}^k$ identifies all graphs in $\mathcal{G}$ if and only if it identifies all graphs in $\mathcal{G}' := \{G \in \mathcal{G} : |E_G| \leq \lceil \binom{|V_G|}{2}/2 \rceil\}$.*

接下来，我们将证明，两个三连通平面图同构，当且仅当它们是 $\mathfrak{s}^3$-等价的。

引理 **5** *Let G be a 3-connected planar graph. If $i_1, i_2, i_3 \in V$ are three distinct nodes sharing a common face, then $\mathfrak{s}^3_G(i_1,i_2,i_3)_i \neq \mathfrak{s}^3_G(i_1,i_2,i_3)_j$ for all distinct $i,j \in V$.*

Proof 对于所有 $k \in \mathbb{Z}_+$，令 γ_k 为一个 $V \to \mathbb{R}^2$ 函数，定义如下。若 $k = 0$，则令

$$\gamma_0(i) := \begin{cases} (0,0) & \text{if } i = i_1, \\ (0,1) & \text{if } i = i_2, \\ (1,0) & \text{if } i = i_3, \\ (1,1) & \text{otherwise.} \end{cases} \tag{16}$$

对于 $k \geq 1$，令

$$\gamma_k(i) := \begin{cases} \gamma_{k-1}(i) & \text{if } i \in \{i_1, i_2, i_3\}, \\ \frac{1}{\delta_G(i)} \sum_{i' \in N_G(i)} \gamma_{k-1}(i') & \text{otherwise.} \end{cases} \tag{17}$$

当 k 趋于无穷大时，γ_k 收敛到平面嵌入 [21], ，因此对于充分大的 k，γ_k 是一个单射。因此，只需证明

$$\gamma_k(i) \neq \gamma_k(j) \Longrightarrow \mathfrak{s}^3_G(i_1, i_2, i_3)_{ik} \neq \mathfrak{s}^3_G(i_1, i_2, i_3)_{jk} \tag{18}$$

对所有 i、$j \in V$均成立，我们通过对 k进行归纳法证明这一点。

基例 $\gamma_0(i) \neq \gamma_0(j) \Longrightarrow \mathfrak{s}^3_G(i_1, i_2, i_3)_{i0} \neq \mathfrak{s}^3_G(i_1, i_2, i_3)_{j0}$由定义可轻易得出。由归纳法，假设 (18) 对 $(k-1)$ 成立，其中 $k \geq 1$。

如果 $i \in \{i_1, i_2, i_3\}$ 或 $j \in \{i_1, i_2, i_3\}$，则 (18) 成立，因为 $\mathfrak{s}^3_G\ i_1\ i_2\ i_3\ |_k$ $\{i_1, i_2, i_3\}$ $\{i_1, i_2, i_3\}$ $i\ j \notin (\ ,\ )$ 中与节点对应的所有行都是唯一的。假设 , 。根据定义，$(\ )\ (\ )$ 意味着 $\gamma_k\ i \neq \gamma_k\ j$

$$\frac{1}{\delta_G(i)} \sum_{i' \in N_G(i)} \gamma_{k-1}(i') \neq \frac{1}{\delta_G(j)} \sum_{j' \in N_G(j)} \gamma_{k-1}(j'). \tag{19}$$

若 $\delta_G(i) \neq \delta_G(j)$，则根据 $\mathfrak{s}^3_G(i_1, i_2, i_3)$ 的定义，(18) 成立。否则，由 (19) 可得

$$\{\gamma_{k-1}(i') : i' \in N_G(i)\}^{\#} \neq \{\gamma_{k-1}(j') : j' \in N_G(j)\}^{\#}$$

由归纳法可知，

$$\{\mathfrak{s}^3_G(i_1, i_2, i_3)_{i',k-1} : i' \in N_G(i)\}^{\#} \neq \{\mathfrak{s}^3_G(i_1, i_2, i_3)_{j',k-1} : j' \in N_G(j)\}^{\#} \tag{20}$$

成立，因此 $\mathfrak{s}^3_G(i_1, i_2, i_3)_{ik} \neq \mathfrak{s}^3_G(i_1, i_2, i_3)_{jk}$。

定理 **15** *Two 3-connected planar graphs, G_1 and G_2 are isomorphic if and only if $\mathfrak{s}^3_{G_1} = \mathfrak{s}^3_{G_2}$.*

Proof 只需证明，若 $\mathfrak{s}^3_{G_1} = \mathfrak{s}^3_{G_2}$，则 G_1 和 G_2 同构。令 i_1、i_2、$i_3 \in V_1$ 为 G_1 的某个平面嵌入中同一公共面上的三个不同节点。根据定义，$\mathfrak{s}^3_{G_1} = \mathfrak{s}^3_{G_2}$ 意味着 $\{\mathfrak{s}^3_{G_1}(i) : i \in V_1\}^{\#} = \{\mathfrak{s}^3_{G_2}(j) :\ j \in V_2\}^{\#}$，因此存在 $j_1 \in V_2$，使得 $\mathfrak{s}^3_{G_1}(i_1) = \mathfrak{s}^3_{G_2}(j_1)$。类似地，可得存在 $j_2 \in V_2$，使得 $\mathfrak{s}^3_{G_1}(i_1, i_2) = \mathfrak{s}^3_{G_2}(j_1, j_2)$；并且存在 $j_3 \in V_2$，使得 $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3) \overset{\text{P}}{=} \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)$。下面的命题给出了所需的双射。

Claim 存在唯一的双射 $\mathfrak{s}^3_{G_2}\ j_1\ j_2\ j_3\ \pi(i)$ $i \in V_1$ π $_i =$: $V_1 \to V_2$，使得对所有 , 都成立 $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_{\text{P}}(\ ,\ )$，并且该双射是保边的。

Proof 根据引理 5，G_1 中的标签是唯一的，也就是说

$$\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_i \overset{\text{P}}{=} \mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_{i'} \Longleftrightarrow i = i' \tag{21}$$

由此可得。由于 $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_{\text{P}} = \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)$，$G_2$ 中的标签也都是唯一的，也就是说，我们得到

$$\mathfrak{s}^3_{G_2}(j_1, j_2, j_3)_j \overset{\text{P}}{=} \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)_{j'} \Longleftrightarrow j = j'. \tag{22}$$

鉴于 $\mathfrak{s}^3_{G_1}(i_1,i_2,i_3)_{\mathrm{P}} = \mathfrak{s}^3_{G_2}(j_1,j_2,j_3)$，由 (21) 和 (22) 可直接推出 π 的唯一存在性。为了证明 π 是保边的，注意到根据命题 3，(21) 和 (22) 对矩阵 $\mathfrak{s}^3_{G_1}(i_1,i_2,i_3)$ 和 $\mathfrak{s}^3_{G_2}(j_1,j_2,j_3)$ 的前 $(n+1)$ 列也成立。因此，第 $(n+2)$ 列中不会出现两行不同的情况。更确切地说，

$$\{\mathfrak{s}^3_{G_1}(i_1,i_2,i_3)_{i',n+1} : i' \in N_{G_1}(i)\}^{\#} = \{\mathfrak{s}^3_{G_2}(j_1,j_2,j_3)_{j',n+1} : j' \in N_{G_2}(\pi(i))\}^{\#} \tag{23}$$

对所有节点 $i \in V_1$均成立。注意，对于所有节点 $i \in V_1$

$$\{\pi(i') : i' \in N_{G_1}(i)\}^{\#} = \{j' : j' \in N_{G_2}(\pi(i))\}^{\#} \tag{24}$$

由式 (23) 可得，因为矩阵 $\mathfrak{s}^3_{G_1}(i_1,i_2,i_3)\big|_{n+1}$ 和 $\mathfrak{s}^3_{G_2}(j_1,j_2,j_3)\big|_{n+1}$的行能够唯一确定相应的节点。式 (24) 意味着 π 保持边不变，由此完成证明。

3.2 实验结果

我们验证了 $\mathfrak{s}^2$能够识别至多含 12 个节点的全部 165 078 520 805 个图；也就是说，这类图同构当且仅当它们是 $\mathfrak{s}^2$-等价的。我们还考察了较大节点集上所有 $r \in \{3,\dots,10\}$的 r-正则图（见表 1），它们都能被 $\mathfrak{s}^2$识别。它还识别了至多含 64 个节点的全部 43 753 个强正则图。

实验在名为 Atlasz [22], 的高性能计算平台上运行。该平台是一个计算机集群，包含 45 个计算节点，每个节点配备一颗 18 核英特尔至强金牌 6240 处理器和 90GB 内存。从本质上说，这意味着该系统由 45 台独立计算机组成，它们可以通过网络连接相互通信。我们利用这些计算节点计算所有待研究图的哈希值，各节点并行处理这些图。小图使用 Nauty 软件包 [23]生成。请注意， n个节点上的图构成一个自补图类，因此根据推论 4，可以大幅减少需要考察的图的数量。我们将哈希值存储在硬盘上——因为无法将约四太字节的数据存入内存。为尽量缩短运行时间并减少所需空间，我们使用了一种名为 MurmurHash [24]的非密码学 64 位哈希函数。我们发现，只有几千对图具有相同的哈希值；对于这些图对，我们可以轻松计算更强但速度更慢的 SHA512 哈希值，并成功区分所有图。计算耗时超过三周。

3.3 不可区分图对

鉴于这些积极结果和计算实验，自然会提出这样的问题：是否存在一对无法通过 $\mathfrak{s}^k$ 区分的非同构图对？在本节中，我们构造具有相同 $\mathfrak{s}^1$ 或 $\mathfrak{s}^2$ 指纹的非同构图对。

我们需要以下众所周知的定义。

度 (r)	节点数 (n)	r-正则图的数量 至多包含 n 个节点
3	≤ 26	2 220 297 316
4	≤ 18	1 081 035 905
5	≤ 16	2 588 603 970
6	≤ 15	1 492 278 977
7	≤ 14	21 610 854
8	≤ 15	1 473 763 950
9	≤ 14	88 203
10	≤ 16	2 585 942 872

表 1：所测试正则图的最大节点数，以及此类图的数量。

定义 7：若存在整数 λ 和 μ，使得任意两个相邻节点恰有 λ 个公共邻居，且任意两个不相邻节点恰有 μ 个公共邻居，则具有 n 个节点且度为 d 的正则图 $G = (V, E)$称为 *strongly regular* 。

定理 **16** *If G_1 and G_2 are connected strongly regular graphs with the same parameters (n, d, λ, μ) then $\mathfrak{s}^1_{G_1} = \mathfrak{s}^1_{G_2}$.*

*Proof*回忆一下，$\mathfrak{s}^1_{G_1} = \{\mathfrak{s}^1_{G_1}(i) \colon\ i \in V_1\}$ 且 $\mathfrak{s}^1_{G_2} = \{\mathfrak{s}^1_{G_2}(j) \colon\ j \in V_1\}$。只需证明，对于所有的 $i \in V_1, j \in V_2$，$\mathfrak{s}^1_{G_1}(i) = \mathfrak{s}^1_{G_2}(j)$。任取 $i \in V_1, j \in V_2$。首先注意，G_1 和 G_2 中的任意节点，分别都能从 i 和 j出发，至多经过两步到达。令 $L^1_q \subseteq V_1$ 和 $L^2_q \subseteq V_2$ 分别表示在 G_1 中距 i 为 q 的节点，以及在 G_2中距 j 为 q 的节点，其中 $q \in \{0, 1, 2\}$。

根据定义，$L^1_0 = \{i\}, L^2_0 = \{j\}, L^1_1 = N_{G_1}(i), L^2_1 = N_{G_2}(j)$，$L^1_2 = V_1 \setminus (L^1_0 \cup L^1_1)$ 和 $L^2_2 = V_2 \setminus (L^2_0 \cup L^2_1)$。注意，对于 $p \in \{1, 2\}$，L^p_0 与 L^p_2 之间没有边，因此 L^p_2 中的任意节点在 L^p_1 中恰有 μ 个邻居，在 L^p_2 中恰有$(d - \mu)$个邻居。

接下来将证明，对于所有 $l \geq 0$，$\mathfrak{s}^1_{G_1}(i)\big|_l = \mathfrak{s}^1_{G_2}(j)\big|_l$。若 $l = 0$，则该结论成立，因为初始标签相同。对于 $l = 1$，L^p_0、L^p_1和 L^p_2中节点的标签分别为$(1, \{\underbrace{0, \ldots, 0}_{d}\}^{\#})$、$(0, \{1, \underbrace{0, \ldots, 0}_{d-1}\}^{\#})$和$(0, \{\underbrace{0, \ldots, 0}_{d}\}^{\#})$，分别，也就是说，它们在两张图中相同。

对于 $l \geq 2$，我们证明，对于所有 $q \in \{0, 1, 2\}$，L^1_q 和 L^2_q 中节点的标签在 G_1 和 G_2 中保持不变。令 h_q 表示对于 $q \in \{0, 1, 2\}$，L^1_q 中节点的（$l - 1$）标签。根据定义，对于 $p = 1$ 和 $p = 2$，L^p_0、L^p_1 和 L^p_2 中节点的新标签分别为（h_0，$\{\underbrace{h_1, \ldots, h_1}_{\lambda}\}^{\#}$）、（$h_1$，$\{h_0, \underbrace{h_1, \ldots, h_1}_{d-\lambda-1}, \underbrace{h_2, \ldots, h_2}_{d}\}^{\#}$）和（$h_2$，$\{\underbrace{h_1, \ldots, h_1}_{\lambda}, \underbrace{h_2, \ldots, h_2}_{d}\}^{\#}$）。因此可得，对于所有 $i \in V_1$、$j \in V_2$，$\mathfrak{s}^1_{G_1}(i) = \mathfrak{s}^1_{G_2}(j)$ 成立，这意味着 $\mathfrak{s}^1_{G_1} = \mathfrak{s}^1_{G_2}$，从而完成证明。

由于存在两个参数相同但不同构的强正则图（这类图中最小的两个图的参数为 (16, 6, 2, 2)），前述定理立即推出以下结论。

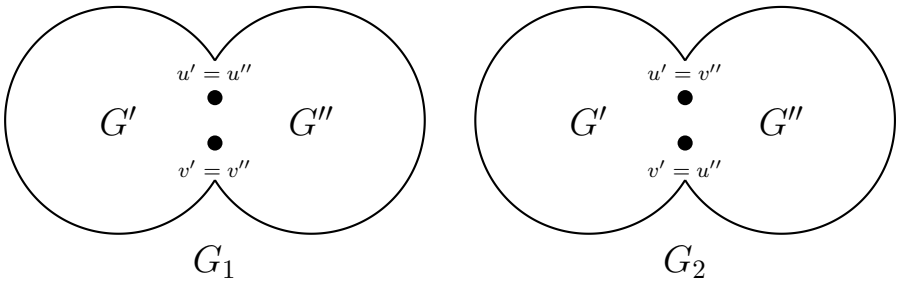


图 2：定理 17 证明中所述构造的示意图。

推论 **5** *There exist two non-isomorphic graphs G_1 and G_2 which are not distinguished by $\mathfrak{s}^1$, that is, $\mathfrak{s}^1_{G_1} = \mathfrak{s}^1_{G_2}$.*

作者长期以来一直在探究一个问题： $\mathfrak{s}^2$ 能否识别所有图。现在，我们给出了两个非同构的图，它们具有相同的 $\mathfrak{s}^2$ 指纹。

定理 **17** *There exist two non-isomorphic graphs G_1, G_2 such that $\mathfrak{s}^2_{G_1} = \mathfrak{s}^2_{G_2}$.*

Proof (Sketch of the proof) 令 $G = (V, E)$ 表示附录 A 中给出的参数为 (35, 18, 9, 9) 的强正则图，并令 u, v 表示分别对应于两个高亮列的两个节点。令 $G' = (V', E')$ 和 $G'' = (E'', V'')$ 为 G 的两个不相交副本，并令 $u', v' \in V'$ 分别表示 $u, v \in V$ 在 G' 中的副本，$u'', v'' \in V''$ 分别表示 $u, v \in V$ 在 G'' 中的副本。考虑如下构造。首先，令 G_1 表示通过合并以下对象得到的图：1） u 的两个副本；2） v 的两个副本。其次，令 G_2 表示通过合并以下节点得到的图：1）节点 $u \in V'$ 和 $v \in V''$；2）节点 $v \in V'$ 和 $u \in V''$。图 2 展示了这一构造。我们通过计算机验证了 $\mathfrak{s}^2_{G_1} = \mathfrak{s}^2_{G_2}$ 和 $\mathfrak{s}^3_{G_1} \neq \mathfrak{s}^3_{G_2}$。后者意味着这两个图不同构，因此 $\mathfrak{s}^2$ 无法区分图 G_1 和 G_2。

Remark 2 设 $G = (V, E)$ 是附录 A 中给出的图。附录 A 所给出的 G 中两个高亮节点 u 和 v 的选取满足

$$\{\mathfrak{s}^2_G(i) : i \in N_G(u)\}^{\#} \neq \{\mathfrak{s}^2_G(i) : i \in N_G(v)\}^{\#}$$

成立。然而，这个条件不足以构造一个反例。

3.4 聚合的另一种方式

本文最后给出一种自然的替代方案，以取代记号 10，并将其与 $\mathfrak{s}^k$ 进行比较。

记号 **18** *For $q = k, \ldots, 0$ and $i \in V$, let*

$$\mathfrak{t}^k_G(i_1, \ldots, i_q; i) := \begin{cases} \mathfrak{s}^k_G(i_1, \ldots, i_k)_i & \text{if } q = k, \\ \{\mathfrak{t}^k_G(i_1, \ldots, i_q, j; i) : j \in V \setminus \{i, i_1, \ldots, i_q\}\}^{\#} & \text{otherwise,} \end{cases}$$

并令

$$\mathfrak{t}_G^k := \{\mathfrak{t}_G^k(;i) : i \in V\}^{\#}. \tag{25}$$

根据以下定理，由记号 10 和 18 给出的两种聚合方式 $\mathfrak{s}_G^k$ 和 $\mathfrak{t}_G^k$ 密切相关。

定理**19** *If* $\mathfrak{s}_{G_1}^{k+1} = \mathfrak{s}_{G_2}^{k+1}$、*then* $\mathfrak{t}_{G_1}^k = \mathfrak{t}_{G_2}^k$. *Similarly*、*if* $\mathfrak{t}_{G_1}^{k+1} = \mathfrak{t}_{G_2}^{k+1}$、*then* $\mathfrak{s}_{G_1}^k = \mathfrak{s}_{G_2}^k$.

Proof 我们只处理第一部分，因为第二部分可以用类似的方式证明。设 $i \in V_1$、$j \in V_2$ 满足 $\mathfrak{s}_{G_1}^{k+1}(i) = \mathfrak{s}_{G_2}^{k+1}(j)$。首先，我们证明

$$\mathfrak{s}_{G_1}^{k+1}(i, i_1, \ldots, i_q) = \mathfrak{s}_{G_2}^{k+1}(j, j_1, \ldots, j_q) \Longrightarrow \mathfrak{t}_{G_1}^k(i_1, \ldots, i_q; i) = \mathfrak{t}_{G_2}^k(j_1, \ldots, j_q; j) \tag{26}$$

对所有 i_1、...、$i_q \in V_1$ 和 j_1、...、$j_q \in V_2$ 以及所有 $q = k$、...、0 均成立。证明对 $q = k$、...、0 进行归纳。在基例中，当 $q = k$时，如果 $\mathfrak{t}_{G_1}^k(i_1$、...、i_q; $i) \neq \mathfrak{t}_{G_2}^k(j_1$、...、$j_q$; $j)$，那么 $\mathfrak{s}_{G_1}^{k+1}(i$、$i_1$、...、$i_q) \neq \mathfrak{s}_{G_2}^{k+1}(j$、$j_1$、...、$j_q)$，因为后一种情况下节点的标签细化了前一种情况中的标签。（注意，如果我们考虑 $\mathfrak{s}^k$ 而不是 $\mathfrak{s}^{k+1}$，则情况不一定如此！）对于 $q < k$，根据归纳假设，假设该命题对 q 的所有更大取值均成立。根据定义，$\mathfrak{s}_{G_1}^{k+1}(i$、$i_1$、...、$i_q) = \mathfrak{s}_{G_2}^{k+1}(j$、$j_1$、...、$j_q)$ 意味着

$$\{\mathfrak{s}_{G_1}^{k+1}(i, i_1, \ldots, i_q, i') : i' \in V_1 \setminus \{i, i_1, \ldots, i_q\}\}^{\#} = \{\mathfrak{s}_{G_2}^k(j, j_1, \ldots, j_q, j') : j' \in V \setminus \{j, j_1, \ldots, j_q\}\}^{\#}, \tag{27}$$

而 $\mathfrak{t}_{G_1}^k(i_1, \ldots, i_q; i) = \mathfrak{t}_{G_2}^k(j_1, \ldots, j_q; j)$ 表示

$$\{\mathfrak{t}_{G_1}^k(i_1, \ldots, i_q, i'; i) : i' \in V_1 \setminus \{i, i_1, \ldots, i_q\}\}^{\#} = \{\mathfrak{t}_{G_2}^k(j_1, \ldots, j_q, j'; j) : j' \in V \setminus \{j, j_1, \ldots, j_q\}\}^{\#}. \tag{28}$$

根据归纳假设，(27) 蕴含 (28)，由此证明了 (26)。

为完成该定理的证明，请注意，对于 $q = 0$，(26) 表示 $\mathfrak{t}_{G_1}^k(; i) = \mathfrak{t}_{G_2}^k(; j)$ 成立，当且仅当 $\mathfrak{s}_{G_1}^{k+1}(i) = \mathfrak{s}_{G_2}^{k+1}(j)$，鉴于 $\mathfrak{s}_{G_1}^{k+1} = \mathfrak{s}_{G_2}^{k+1}$，这意味着 $\mathfrak{t}_{G_1}^k = \mathfrak{t}_{G_2}^k$。

用于证明定理 17 的图

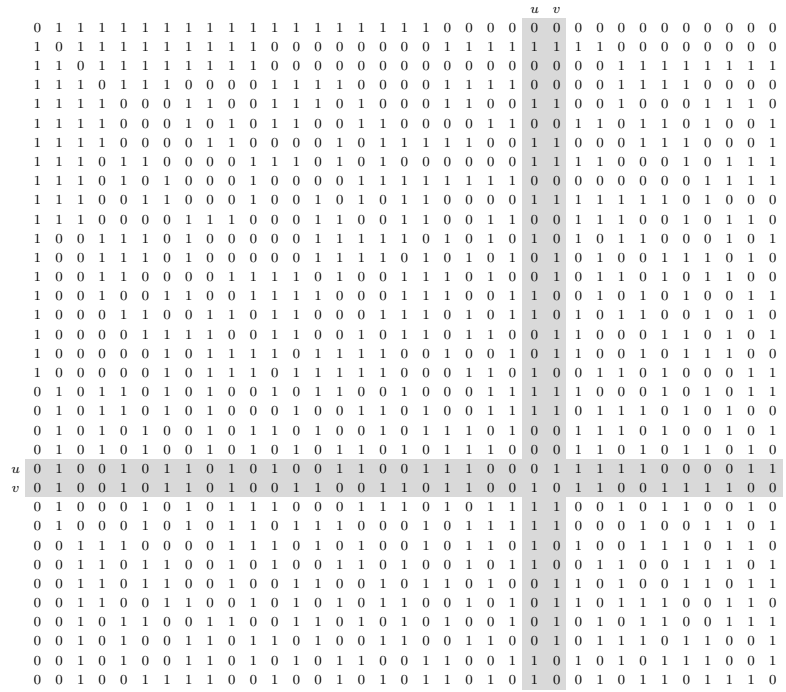


表 2：用于证明定理 17 的图的邻接矩阵。

参考文献

1. J. E. Hopcroft 和 J. K. Wong。平面图同构的线性时间算法。*Proceeding STOC '74 Proceedings of the sixth annual ACM symposium on Theory of computing Pages 172-184* 1974 , ,

April 。2. E. M. Luks。有界度数图的同构可在多项式时间内判定。
Journal of Computer and System Sciences, Volume 25, Issue 1, Pages 42-65, August 1982。
3. G. S. Lueker 和 K. S. Booth。判定区间图同构的线性时间算法。 *Journal of the ACM (JACM), Volume 26, Issue 2, Pages 183-195*, 1979, April 1979。4. C. J. Colbourn。关于排列图同构判定。 *Networks, Volume 11,*
Issue 1 1981 1981 ,
Pages 13-21, , March 。5. V. Arvind、B. Das、J. K¨obler 和 S. Toda。彩色超图同构是固定参数可处理的。 *Algorithmica Volume 71, Pages 120-138*, 2015, January 2015。6. B. D. McKay。实用图同构。
Congressus Numerantium Volume 30 Pages , , , , 。
45-87 1981 1981
7. L. P. Cordella、P. Foggia、C. Sansone 和 M. Vento。用于匹配大图的（子）图同构算法。*IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 26 Issue 10 Page 1367-1372 2004* 2004 , , , 。

8. A. J¨uttner 和 P. Madarasi。VF2++——一种改进的子图同构算法。*Discrete Applied Mathematics*, 2018。

9. D. Shasha、J. T.-L. Wang 和 R. Giugno。树与图搜索的算法学及其应用。载于 *Proceedings of the Twenty-first ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* 02, 39 52, , P ODS ' pages – 纽约，NY，美国，2002。ACM。10. E. R. van Dam 和 W. H. Haemers。哪些图由其谱确定？*Linear Algebra and its Applications*，373（增刊 C）：241 – 272，2003。组合矩阵理论会议（浦项，2002）。11. R. C. Wilson 和 P. Zhu。利用图谱比较图和树的研究。*Pattern Recognition*，41(9)：2833 – 2841，2008。12. D. Raviv、R. Kimmel 和 A. M. Bruckstein。通过谱签名研究图同构与自同构。 *IEEE Transactions on Pattern Analysis and Machine Intelligence*，35 8 1985 2013 ( ): – 1993，8 月。13. A. E. Brouwer 和 E. Spence。12 个顶点上的同谱图。 *Electr. J. Comb.*，16，2009 年 6 月。14. B. L. Douglas 和 J. B. Wang。利用量子游走解决图同构问题的经典方法。 *Journal of Physics A: Mathematical and Theoretical* 41 7 075303, , ( ): 。2008
15. K. Rudinger、J. K. Gamble、M. Wellons、E. Bach、M. Friesen、R. Joynt 和 S. Coppersmith。将非相互作用多粒子量子随机游走应用于强正则图的图同构问题。 *Phys. Rev. A*，86，2012 年 8 月。16. A. Mahasinghe、J. A. Izaac、J. B. Wang 和 J. K. Wijerathna。相位修正连续时间量子游走无法高效地区分强正则图。 *Journal of Physics A: Mathematical and Theoretical* 48 26 , ( ): 265301，2015。17. J. K. Gamble、M. Friesen、D. Zhou、R. Joynt 和 S. N. Coppersmith。应用于图同构问题的双粒子量子游走。 *Physical Review A*，81：052313，2010 年 5 月。18. L. Babai、P. Erd ″ os 和 S. M. Selkow。随机图同构。, ( ): – 。19. B. Weisfeiler 和 A. Leman。将图约化为规范形式及其中出现的代数。 *NTI*， *Series*，2(9)：12–16，1968。20. Y. Aflalo、A. Bronstein 和 R. Kimmel。关于图同构的凸松弛。*Proceedings of the National Academy of Sciences*，112(10)：2942–2947，2015。21. W. T. Tutte。《如何绘制图》。 *Proceedings of the London Mathematical Society*，s3-13(1)：743–767，1963 年 1 月。22. Atlasz。 `https://hpc.iig.elte.hu`。23. B. D. McKay 和 A. Piperno。实用图同构，第二部分。 *Journal of Symbolic Computation* 60 94 112, 2014 , : – 。
24. MurmurHash。 `https://sites.google.com/site/murmurhash`。