THE METRIC SPACE OF CONNECTED SIMPLE GRAPHS

Kiran R. Bhutani Bilal Khan

Abstract

We introduce a sequence $e_n^\ell(\ell = 1, 2, \dots)$ of real-valued functions on $\mathcal{G}_n \times \mathcal{G}_n$, where $\mathcal{G}_n$ is the set of all simple, connected, undirected graphs of order n up to isomorphism. The functions e_n^ℓ arise naturally from the consideration of graph embeddings. The binary relations $\succ_\ell$ on $\mathcal{G}_n$ are then defined to be exactly the zeroes of e_n^ℓ. We observe that the relation $\succ_1$ coincides precisely with the classical subgraph relation, and that $\succ_1, \succ_2, \dots$ is a sequence of weaker binary relations on $\mathcal{G}_n$. Motivated by this, we define $d_n^\ell(H, K) = e_n^\ell(H, K) + e_n^\ell(K, H)$, thereby obtaining a sequence of symmetric, real-valued functions $d_n^\ell(\ell = 1, 2, \dots)$ on $\mathcal{G}_n \times \mathcal{G}_n$. When $\ell = 1$, (G_n, d_n^1) is a totally disconnected metric space of graphs that embodies the classical notion of graph isomorphism: $H \cong K \Leftrightarrow d_n^1(H, K) = 0$, $H \ncong K \Leftrightarrow d_n^1(H, K) = \infty$. As ℓ increases in the range $2, \dots, n-2$, the value of d_n^ℓ decreases for every distinct pair of graphs in $\mathcal{G}_n$ although distinct graphs in G_n do not collapse. Further, we show that the graphs H for which $d_n^\ell(H, K_n)$ is finite, are precisely those graphs $H \in \mathcal{G}_n$ which have diameter $\leqslant \ell$. Finally, when $\ell \geqslant n-1$, the functions d_n^ℓ form a stationary sequence whose elements equip $\mathcal{G}_n$ once again with full metric structure. We refer to this space as (G_n, d_n^*), "The metric space of connected simple graphs." Unlike (G_n, d_n^1), the space (G_n, d_n^*) is connected. A number of open questions concerning the geodesic structure of (G_n, d_n^*) are also presented.

1 Introduction

We begin with an informal, brief description of the concepts and motivation behind this work. The statements here will be formulated precisely in the subsequent sections.

Given two connected, simple graphs H, K of order n, begin by considering a one-to-one mapping $\phi : V[K] \to V[H]$. Upon fixing ϕ, to each edge $e = (u, v) \in E[K]$ we associate a walk p_e between $\phi(u)$ and $\phi(v)$ in H. We view ϕ together with the chosen set of walks $Q = \{p_e | e \in E[K]\}$ as a topological embedding ϕ' of K into H. If all the walks in Q are of length $\leqslant \ell$, then ϕ' is called an ℓ-topological

Subject Classification: 05C99 (Graph theory)

embedding of K into H. Given ϕ', each edge e in $E[H]$ is traversed some number of times by the walks in Q. We call this number "the congestion on e" with respect to the embedding ϕ'—it is, in some sense, the size of the fibre over edge e under ϕ'. For a given embedding ϕ' of K into H, we will be interested in the maximum congestion on the edges of $E[H]$. In general, there are many (or none) ℓ-topological embeddings of K into H, and we shall be seeking those embeddings which minimize the maximum congestion. The logarithm of this minimal value will be later defined as the ℓ-embedding thickness of K in H, denoted $e_n^\ell(H, K)$.

The situation above is a reasonable abstract model of the practical problem of virtual path layout in computer networks [4, 6]: Given a physical computer network H consisting of switches $V[H]$ connected by fiber-optic cables $E[H]$ of fixed capacity, we desire to implement a specific virtual computer network K using the physical hardware of H. This amounts to finding a topological embedding of K into H; the links of the virtual network $E[K]$ are implemented as network connections (i.e. walks) of length $\leqslant \ell$ in the physical network H. Since these connections must utilize the bandwidth of links in the physical network, it is desirable to use an embedding which minimizes the maximum congestion over the physical links.

Above we considered the topological embeddings of K into H, and stated that the embedding which minimizes the maximum congestion of edges in $E[H]$ is what, by definition, determines the value of the function $e_n^\ell(H, K)$. Likewise, one might also consider embeddings H into K, and so determine the value of $e_n^\ell(K, H)$. The sum of these two functions will be later defined as the ℓ-distance between H and K, denoted $d_n^\ell(H, K)$. We will analyze the properties of these functions; in particular, we will show that when $\ell \geqslant n - 1$, d_n^ℓ equips the set of simple connected graphs of order n with a metric structure.

In an earlier paper, G. Chartrand, G. Kubicki and M. Schultz [1] presented a different metric on the set of graphs of order n, which they termed ϕ-distance. Their measure was based on embeddings of graphs which minimize the absolute distortion of pairwise distances between vertices. In contrast, the metric proposed here is based on consideration of embeddings between graphs which minimize the maximum size fibre over any edge. For us, the study of such embeddings originated in the practical problem of virtual path layout on computer networks [4, 6].

2 Preliminaries

Given an undirected graph $G = (V, E)$, recall that a **directed walk of length l in G** is a sequence of $l + 1$ vertices $w = (v_0, v_1, \cdots, v_l)$, where $v_i \in V$ for $i = 0, \ldots, l$, and $(v_j, v_{j+1}) \in E$ for $j = 0, \ldots, l - 1$. For a walk $w = (v_0, v_1, \cdots, v_l)$, denote $o(w) = v_0$ and $t(w) = v_l$ to be the **origin** and **terminal** vertices of w, respectively.

Definition 2.1. For each $l \in \mathbb{Z}^+$, we define $\mathcal{W}_G^l$ to be the **set of all directed walks in G of length $\leqslant l$.**

Given a directed walk $w = (v_0, v_1, \cdots, v_l)$, denote it's reverse $(v_l, v_{l-1}, \cdots, v_0)$ as w^R. Define a binary relation $\sim_R$ on $\mathcal{W}_G^l$ as follows: for $w, w' \in \mathcal{W}_G^l$, take $w \sim_R w'$ if and only if $w^R = w'$ or $w = w'$. Clearly, $\sim_R$ is an equivalence relation on $\mathcal{W}_G^l$.

Definition 2.2. Define $\mathcal{P}_G^l = \mathcal{W}_G^l / \sim_R$ and take $\sigma : \mathcal{W}_G^l \to \mathcal{P}_G^l$ to be the canonical projection. Then, an element of $\mathcal{P}_G^l$ is referred to as an **undirected walk of length $\leqslant l$ in G.** Finally, take $\mathcal{P}_G = \bigcup_{l=1}^{\infty} \mathcal{P}_G^l$. For each $p \in \mathcal{P}_G$, define the **boundary** of p as $\partial p = \{o(w), t(w)\}$, where $w \in \sigma^{-1}(p)$ is chosen arbitrarily; clearly ∂p is independent of the choice of w.

Definition 2.3. To each set of undirected walks $Q \subseteq \mathcal{P}_G$, associate **the undirected walk graph** $Q^\circ = (V, E_Q)$, where $(u, v) \in E_Q \Leftrightarrow \exists p \in Q$ such that $\partial p = \{u, v\}$.

Definition 2.4. For $p \in \mathcal{P}_G$, define $\chi_p : E \to \mathbb{N}$, where for $e \in E$, $\chi_p(e) = m$ if and only if e is traversed exactly m times in w, for some $w \in \sigma^{-1}(p)$. It is easy to see that χ_p is well-defined, i.e. it is independent of the choice of $w \in \sigma^{-1}(p)$. For each set of undirected walks $Q \subseteq \mathcal{P}_G$, define $\Phi_Q : E \to \mathbb{N}$ where for each $e \in E$, $\Phi_Q(e) = \sum_{p \in Q} \chi_p(e)$ is called **the congestion of Q at** e. Finally, given any undirected graph G, define $\tau_G : 2^{\mathcal{P}_G} \to \mathbb{N}$, where $\tau_G(Q) = \max_{e \in E} \Phi_Q(e)$ is called **the congestion of Q on G.**

Definition 2.5. Let $\mathcal{G}_n$ be the set of all simple, connected, undirected graphs (up to isomorphism) on n vertices. For each positive integer ℓ, we define $e_n^\ell : \mathcal{G}_n \times \mathcal{G}_n \to \mathbb{R}^{\geqslant 0}$; for each $H, K \in \mathcal{G}_n$ the **ℓ-embedding thickness** of K in H is denoted $e_n^\ell(H, K)$ and defined by

$$e_n^\ell(H, K) \stackrel{\text{def}}{=} \begin{cases} x & \text{if } \exists Q \subseteq \mathcal{P}_H^\ell \text{ s.t. } Q^\circ \simeq K \text{ and } \tau_H(Q) = 2^x \\ & \quad \text{and } \forall R \subseteq \mathcal{P}_H^\ell, R^\circ \simeq K \Rightarrow \tau_H(R) \geqslant 2^x \\ \infty & \text{if } \nexists Q \subseteq \mathcal{P}_H^\ell \text{ s.t. } Q^\circ \simeq K \end{cases}$$

Definition 2.6. A graph K is called an **ℓ-subgraph** of a graph H (denoted $H \succ_\ell K$) if $e_n^\ell(H, K) = 0$.

3 Results

When $\ell = 1$ the notion of ℓ-subgraph coincides with the classical notion of a subgraph.

Lemma 3.1. *$H \succ_1 K \Leftrightarrow K$ is a subgraph of H*

Proof. $\Leftarrow$ Suppose K is a subgraph of H via $i : K \hookrightarrow H$. Take $Q = i(E[K])$; since Q consists of walks of length 1, $Q \subseteq \mathcal{P}_H^1$. Note that $Q^\circ \simeq K$ and $\tau_H(Q) = 1$. This shows that $e_n^1(H,K) = 0$, which means $H \succ_1 K$.

$\Rightarrow$ $H \succ_1 K$ means $e_n^1(H,K) = 0$, so $\exists Q \subseteq \mathcal{P}_H^1$ s.t. $Q^\circ \simeq K$ and $\tau_H(Q) = 1$. But $\mathcal{P}_H^1 = E[H]$, so $Q \subseteq E[H]$. It follows that Q° is a subgraph of H, and thus K is a subgraph of H. □

The next lemma is easy to verify.

Lemma 3.2. *$\forall l, l'$, $1 \leqslant \ell < \ell'$ implies $\forall H, K \in \mathcal{G}_n$, $e_n^{\ell'}(H,K) \leqslant e_n^{\ell}(H,K)$.*

As a corollary, the relations $\succ_\ell$, $\ell \in \mathbb{Z}^+$ form an ascending sequence of binary relations on $\mathcal{G}_n$.

Corollary 3.3. *$\forall l, l'$, $1 \leqslant \ell < \ell'$ implies $\forall H, K \in \mathcal{G}_n$, $H \succ_\ell K \Rightarrow H \succ_{\ell'} K$.*

The next proposition shows the functions $\{e_n^\ell\}_{l \in \mathbb{Z}^+}$ satisfy a graded triangle inequality.

Proposition 3.4. *For any $\ell_1, \ell_2 \in \mathbb{Z}^+$, and any $G, H, K \in \mathcal{G}_n$,*

$$e_n^{\ell_1 \ell_2}(G,K) \leqslant e_n^{\ell_1}(G,H) + e_n^{\ell_2}(H,K)$$

Proof. Suppose that (i) $e_n^{\ell_1}(G,H) = x$ and (ii) $e_n^{\ell_2}(H,K) = y$. Then this implies that (i) $\exists Q \subseteq \mathcal{P}_G^{\ell_1}$ such that $Q^\circ \simeq H$, and $\tau_G(Q) = 2^x$, and (ii) $\exists R \subseteq \mathcal{P}_H^{\ell_2}$ such that $R^\circ \simeq K$ and $\tau_H(R) = 2^y$.

Since $H \simeq Q^\circ$ are simple graphs, we may define $\alpha : E_Q \to \mathcal{P}_G^{\ell_1}$ to be the map taking each $e = (u,v) \in E_Q$ to the unique undirected walk $p_e \in Q$ satisfying $\partial p_e = \{u,v\}$. Fix a graph isomorphism $\pi : H \to Q^\circ$. Then π maps the edge set $E[H] = \mathcal{P}_H^1$ bijectively onto E_Q, giving us the composite map $\alpha\pi : \mathcal{P}_H^1 \to \mathcal{P}_G^{\ell_1}$. Since $\tau_G(Q) = 2^x$, we know that $\forall X \subseteq \mathcal{P}_H^1, \tau_G(\alpha\pi X) \leqslant \tau_H(X) \cdot 2^x = 2^x$.

We begin by extending $\alpha\pi$ to a map $\widetilde{\alpha\pi} : \mathcal{P}_H^{\ell_2} \to \mathcal{P}_G^{\ell_1 \ell_2}$ having the property that $\forall X \subseteq \mathcal{P}_H^{\ell_2}, \tau_G(\widetilde{\alpha\pi} X) \leqslant \tau_H(X) \cdot 2^x$. This is accomplished as follows: Given an undirected walk $p \in \mathcal{P}_H^{\ell_2}$, choose an arbitrary $w \in \sigma^{-1}(p)$. Suppose that $w = (v_0, v_1, \ldots, v_k)$, for some $k \leqslant \ell_2$. For each $i = 1, \ldots, k$ define $r_i = \alpha\pi(v_{i-1}, v_i) \in \mathcal{P}_G^{\ell_1}$. Choose $q_i \in \sigma^{-1}(r_i)$ such that q_i is a directed walk starting at $t(q_{i-1})$ and ending at $o(q_{i+1})$. Since each $q_i \in \mathcal{W}_G^{\ell_1}$, the concatenation of directed walks $q_1, q_2, \cdots, q_k$ forms a directed walk w' starting at vertex the vertex $o(q_1)$ and ending at the vertex $t(q_k)$. The length of w' is $\leqslant \ell_1 k \leqslant \ell_1\ell_2$, thus $w' \in \mathcal{W}_G^{\ell_1\ell_2}$. We define $\widetilde{\alpha\pi}(p) \stackrel{\text{def}}{=} \sigma(w') \in \mathcal{P}_G^{\ell_1\ell_2}$.

Now define $T = \widetilde{\alpha\pi}(R)$. Since $R \subseteq \mathcal{P}_G^{\ell_2}$, it follows that $T \subseteq \mathcal{P}_G^{\ell_1\ell_2}$. But $R^\circ \simeq K$, and since by construction, $T^\circ \simeq R^\circ$, it follows that $T^\circ \simeq K$. Now $\tau_G(T) = \tau_G(\widetilde{\alpha\pi}R)$ which is $\leqslant \tau_H(R) \cdot 2^x = 2^y 2^x = 2^{x+y}$. All this shows that there exists $T \subseteq \mathcal{P}_G^{\ell_1\ell_2}$ such that $T^\circ \simeq K$ and $\tau_G(T) \leqslant 2^{x+y}$. Thus $e_n^{\ell_1\ell_2}(G,K) \leqslant x+y = e_n^{\ell_1}(G,H) + e_n^{\ell_2}(H,K)$. □

Remark 3.5. The well-known fact that "If H is a subgraph of K, and K is a subgraph of H then $H \simeq K$" can be extended to the ℓ-subgraph relations $\succ_\ell$, $\forall \ell \in \mathbb{Z}^+$, as the next result shows.

Theorem 3.6. *$\forall \ell \in \mathbb{Z}^+$, $\forall H,K \in \mathcal{G}_n$, $[H \succ_\ell K$ and $K \succ_\ell H$ implies $H \simeq K]$*

Proof. Suppose $H,K \in \mathcal{G}_n$ such that (i) $e_n^\ell(H,K) = 0$ and (ii) $e_n^\ell(K,H) = 0$. Then this implies that (i) $\exists Q \subseteq \mathcal{P}_H^\ell$ such that $Q^\circ \simeq K$ and $\tau_H(Q) = 1$, and (ii) $\exists R \subseteq \mathcal{P}_K^\ell$ such that $R^\circ \simeq H$ and $\tau_K(R) = 1$.

We show first that $|E[H]| = |E[K]|$. Suppose not, for a contradiction. Then, WLOG $|E[H]| > |E[K]|$. But $H \simeq R^\circ$, $R \subseteq \mathcal{P}_K$, implying that $|R| > |E[K]|$. By pigeonhole argument, $\exists e \in E[K]$ such that $\exists r_1, r_2 \in R$ distinct, both traversing the edge e, which contradicts the assumption that $\tau_K(R) = 1$. So $|E[H]| = |E[K]|$.

Now note that $Q^\circ \simeq K \Rightarrow |Q| = |E[K]|$, and $R^\circ \simeq H \Rightarrow |R| = |E[H]|$. Thus $|R| = |E[H]| = |E[K]| = |Q|$. Each $q \in Q$ corresponds to a undirected walk in H; since $\tau_H(Q) = 1$ and $|E[H]| = |Q|$, this implies that each $q \in Q$ must be a undirected walk of length 1 in H. Likewise, each $r \in R$ corresponds to a undirected walk in K; since $\tau_K(R) = 1$ and $|E[K]| = |R|$, this implies that each $r \in R$ must be a undirected walk of length 1 in K.

Suppose now, towards contradiction, that $H \not\simeq K$. Then, for every bijective map $\pi : V[H] \to V[K]$, there exists $e_\pi \in E[H]$ such that $\pi(e_\pi) \notin E[K]$. But edges $E[H]$ are in bijective correspondence with a set of undirected walks R in K. Thus, $\forall \pi, \exists r_\pi \in R$ corresponding to a undirected walk of length > 1 in K. This contradicts the fact that each $r \in R$ must be a undirected walk of length 1 in K. We have shown then that $\exists \pi : V[H] \to V[K]$ which respects the edge relations $E[H]$, $E[K]$. Hence $H \simeq K$. □

The previous lemma leads us to define:

Definition 3.7. The **ℓ-distance** function $d_n^\ell : \mathcal{G}_n \times \mathcal{G}_n \to \mathbb{R}^{\geqslant 0}$, is defined as

$$d_n^\ell(H,K) \stackrel{\text{def}}{=} e_n^\ell(H,K) + e_n^\ell(K,H)$$

Then d_n^ℓ is a symmetric function, and by theorem 3.6, $d_n^\ell(H,K) = 0 \Leftrightarrow H \simeq K$.

The next lemma can be easily verified:

Lemma 3.8. *Given $G \in \mathcal{G}_n$, and $Q \subseteq \mathcal{P}_G^\ell$ such that $\tau_G(Q) = 2^x$, then $\exists \hat{Q} \subseteq \mathcal{P}_G^\ell$ for which $\tau_G(\hat{Q}) \leqslant 2^x$, $\hat{Q}^\circ \simeq Q^\circ$ and $\forall \hat{q} \in \hat{Q}$, $\hat{q}$ is not self-intersecting.*

Lemma 3.9. *$\forall \ell, \ell' \geqslant n-1$, $e_n^\ell \equiv e_n^{\ell'}$, that is the relations e_n^ℓ are constant.*

Proof. Given $G, H \in \mathcal{G}_n$, $e_n^\ell(G,H) = x$ implies $\exists Q \subseteq \mathcal{P}_G^\ell$ such that $\tau_G(Q) = 2^x$. By lemma 3.8, $\hat{Q} \subseteq \mathcal{P}_G^{n-1}$, $Q^\circ \simeq \hat{Q}^\circ$, and $\tau_G(\hat{Q}) \leqslant \tau_G(Q) = 2^x$. Thus, $e_n^{n-1}(G,H) \leqslant e_n^\ell(G,H)$. On the other hand, by lemma 3.2, $\ell \geqslant n-1$ implies $e_n^\ell(G,H) \leqslant e_n^{n-1}(G,H)$. It follows that $e_n^\ell \equiv e_n^{n-1}$. An analogous argument shows that $e_n^{\ell'} \equiv e_n^{n-1}$. □

Corollary 3.10. *If $\ell = 1$ or $\ell \geqslant n-1$, then e_n^ℓ satisfies the triangle inequality.*

Proof. If $\ell_1 = \ell_2 = 1$ then proposition 3.4 implies that e_n^1 satisfies the triangle inequality. If $\ell_1, \ell_2 \geqslant n-1$, then $\ell_1\ell_2 \geqslant n-1$, so lemma 3.9 implies $e_n^{\ell_1\ell_2} \equiv e_n^{\ell_1} \equiv e_n^{\ell_2} \equiv e_n^{n-1}$, and so all four functions coincide and satisfy the triangle inequality. □

Remark 3.11. As a corollary we extend the well-known assertion that "If G is a subgraph of H, and H is a subgraph of K, then G is a subgraph of K" to the relations $\succ_\ell$, ($\ell \geqslant n-1$).

Corollary 3.12. *If $\ell = 1$ or $\ell \geqslant n-1$, then $\forall\, G, H, K \in \mathcal{G}_n$, $G \succ_\ell H$ and $H \succ_\ell K$ implies $G \succ_\ell K$*

Proof. By definition, $G \succ_\ell H, H \succ_\ell K$ imply $e_n^\ell(G,H) = e_n^\ell(H,K) = 0$. By Corollary 3.10, $e_n^\ell(G,K) = 0$, so $G \succ_\ell K$. □

Theorem 3.13. *If $\ell = 1$ or $\ell \geqslant n-1$, $(\mathcal{G}_n, d_n^\ell)$ is a metric space.*

Proof. Clearly, d_n^ℓ is a symmetric function. By lemma 3.10, e_n^ℓ satisfies the triangle inequality when $\ell = 1$ or $\ell \geqslant n-1$. Finally, to see reflexivity, observe that $d_n^\ell(H,K) = 0$ implies that $e_n^\ell(H,K) = 0$ and $e_n^\ell(K,H) = 0$, so by theorem 3.6, $H \simeq K$. □

Definition 3.14. Define $e_n^* : \mathcal{G}_n \times \mathcal{G}_n \to \mathbb{R}^{\geqslant 0}$ as follows; the **embedding thickness** of K in H is given by

$$e_n^*(H,K) = x \overset{\text{def}}{\Leftrightarrow} \exists Q \subseteq \mathcal{P}_H \text{ s.t. } Q^\circ \simeq K \text{ and } \tau_H(Q) = 2^x$$
$$\text{and } \forall R \subseteq \mathcal{P}_H, R^\circ \simeq K \Rightarrow \tau_H(R) \geqslant 2^x$$

Note that e_n^* is defined in terms of a set of undirected walks $Q \subseteq \mathcal{P}_H$ whose lengths are arbitrary, whereas e_n^ℓ was defined in terms of a set of walks $Q \subseteq \mathcal{P}_H^\ell$ of length $\leqslant \ell$. We write $H \succ_* K$ when $e_n^*(H,K) = 0$, and define $d_n^* : \mathcal{G}_n \times \mathcal{G}_n \to \mathbb{R}^{\geqslant 0}$ to be $d_n^* : (H,K) \mapsto e_n^*(H,K) + e_n^*(K,H)$.

In lemma 3.1, we showed that the relation $\succ_1$ precisely captures the notion of a subgraph, i.e. $H \succ_1 K \Leftrightarrow K$ is a subgraph of H. The next two lemmas show that for n sufficiently large, the relation $\succ_*$ is strictly weaker than $\succ_1$:

Lemma 3.15. $\forall H, K \in \mathcal{G}_n,\ H \succ_1 K \Rightarrow H \succ_* K$

Proof. $H \succ_1 K$ implies K is a subgraph of H via some map $i : K \hookrightarrow H$. Take $Q = i(E[K])$. Note that $Q^\circ \simeq K$ and $\tau_H(Q) = 1$. This shows that $e_n^*(H, K) = 0$, which means $H \succ_* K$. □

The converse of lemma 3.15 is false:

Lemma 3.16. $\forall n \geqslant 5,\ \exists H, K \in \mathcal{G}_n$ *s.t.* $H \succ_* K$ *and* $H \not\succ_1 K$

Proof. Fix $n \geqslant 5$, and take graphs $H, K \in \mathcal{G}_n$ to be the graphs depicted in figure 1. Let $Q \subseteq \mathcal{P}_H$ be the set of paths in H chosen as indicated in the center figure. Then $Q^\circ \simeq K$, and $\tau_H(Q) = 1$. This shows $e_n^*(H, K) = 0$, i.e. $H \succ_* K$. It is an easy exercise to check that K is not a subgraph of H, i.e. $H \not\succ_1 K$. □

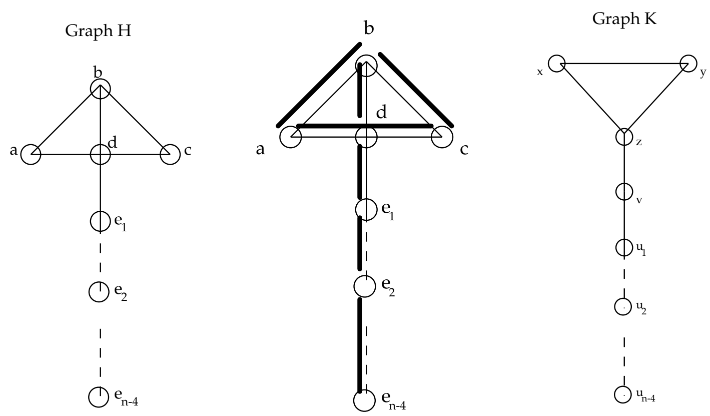

Figure 1: A set of paths certifying that $e_n^*(H, K) = 0$.

Lemma 3.17. *Let* $H = (V, E) \in \mathcal{G}_n$*, and* $e = (u, v) \in V \times V$*. Let* $H' = (V, E \sqcup e)$*. Then* $d_n^*(H, H') \leqslant 1$.

Proof. $H' \succ_1 H$, so by lemma 3.15, $H' \succ_* H$, i.e. $e_n^*(H', H) = 0$. Let $Q = E \cup p$ where p is any undirected walk in H such that $\partial p = \{u, v\}$. Without loss of generality, we may take p to be non self-intersecting. Then $Q^\circ \simeq H'$ and $\tau_H(Q) \leqslant 2$. This shows that $e_n^*(H, H') \leqslant \log_2 2 = 1$. By definition, $d_n^*(H', H) = e_n^*(H', H) + e_n^*(H, H') \leqslant 1$. □

Lemma 3.18. *$\forall H, K \in \mathcal{G}_n$, $d_n^1(H,K)$ is 0 if $H = K$ and infinite otherwise.*

Proof. If $H = K$ then H is a subgraph of K and K is a subgraph of H, hence $e_n^1(H,K) = e_n^1(K,H) = 0$. It follows that $d_n^1(H,K) = 0$. If $H \neq K$, then either H is not a subgraph of K or K is not a subgraph of H. Suppose, WLOG that K is not a subgraph of H. Then, $\forall Q \subseteq \mathcal{P}_H^1$, $Q^\circ \not\simeq K$, so by definition, $e_n^1(H,K) = \infty$. It follows that $d_n^1(H,K) = \infty$. □

Lemma 3.19. *$\forall H, K \in \mathcal{G}_n$, $d_n^*(H,K)$ is finite.*

Proof. One can transform H into K by a finite sequence of transformations: $H = G_0 \rightsquigarrow G_1 \rightsquigarrow \ldots \rightsquigarrow G_m = K$, where for $i = 0, \ldots, m-1$, the transformation $G_i \rightsquigarrow G_{i+1}$ is the addition or deletion of a single edge. It is possible to construct such a transformation sequence by first sequentially adding edges to H until the complete graph K_n is attained, and then deleting edges as needed to arrive at the graph K. Clearly, all the intermediate graphs G_i are simple, connected, undirected graphs on n vertices, hence $\forall i = 0, \ldots, m, G_i \in \mathcal{G}_n$. By lemma 3.17, $d_n^*(G_i, G_{i+1}) \leqslant 1$, and theorem 3.13 asserts that d_n^* satisfies the triangle inequality; hence $d_n^*(H,K) \leqslant m < \infty$. □

Definition 3.20. Given a graph $G \in \mathcal{G}_n$, we define the ℓ-neighborhood of G as

$$\ell\text{-}nbd(G) = \{H \in \mathcal{G}_n \mid d_n^\ell(G,H) \text{ is finite. }\}$$

Since $l \geqslant n-1$ implies $d_n^\ell \equiv d_n^*$, we refer to ℓ-*nbd* as $*$-*nbd* whenever $l \geqslant n-1$. In this language, lemma 3.18 can be restated as $\forall G \in \mathcal{G}_n$, $1\text{-}nbd(G) = \{G\}$, while lemma 3.19 is seen to assert that $\forall G \in \mathcal{G}_n$, $*\text{-}nbd(G) = \mathcal{G}_n$.

The next lemma characterizes $\ell\text{-}nbd(G)$ when G is a complete graph K_n.

Lemma 3.21. *$\ell\text{-}nbd(K_n) = \{H \in \mathcal{G}_n \mid Diameter(H) \leqslant \ell\}$*

Proof. To see that $\ell\text{-}nbd(K_n) \subseteq \{H \in \mathcal{G}_n \mid Diameter(H) \leqslant \ell\}$: Suppose $H = (V,E) \in \ell\text{-}nbd(K_n)$. Then $\exists Q \subseteq \mathcal{P}_H^\ell$ such that $Q^\circ \simeq K_n$. It follows that $\forall u, v \in V$, $\exists q_{u,v} \in Q$ with $\partial q_{u,v} = \{u, v\}$. But $Q \subseteq \mathcal{P}_H^\ell$, so $q_{u,v}$ has length $\leqslant \ell$. This shows that $Diameter(H) \leqslant \ell$.

To see $\ell\text{-}nbd(K_n) \supseteq \{H \in \mathcal{G}_n \mid Diameter(H) \leqslant \ell\}$: Suppose $H = (V,E) \in \mathcal{G}_n$ and $Diameter(H) \leqslant \ell$. Let Q be the set of $n(n-1)$ shortest paths between all pairs of vertices $u, v \in V$. Clearly, $Q^\circ \simeq K_n$ and $Q \subseteq \mathcal{P}_H^{Diam(H)} \subseteq \mathcal{P}_H^\ell$. It follows that $e_n^\ell(H, K_n) \leqslant \tau_Q(H) < \infty$. Now H is a subgraph of K_n, so by lemmas 3.1 and 3.2, $e_n^\ell(K_n, H) = 0$. Since $d_n^\ell(K_n, H) = e_n^\ell(K_n, H) + e_n^\ell(H, K_n)$, it follows that $d_n^\ell(K_n, H) < \infty$, that is $H \in \ell\text{-}nbd(K_n)$. □

4 Remarks and Open Questions

Many interesting graph theoretic properties are defined in terms of the presence of forbidden subgraphs, i.e. "G has property P iff $G \succ_1 H$ for some $H \in \mathcal{H}$", where $\mathcal{H}$ is a family of forbidden subgraphs. Graph properties that are defined in this manner can be naturally weakened to their $*$-analogues by declaring that "G has property $*$-P iff $G \succ_* H$ for some $H \in \mathcal{H}$". By lemma 3.15, the set of graphs with property P is a subset of the set of graphs with property $*$-P; by lemma 3.16 this containment may be proper.

As a concrete example, $G \in \mathcal{G}_n$ is said to be Hamiltonian if $G \succ_1 C_n$, where C_n is the cycle of n vertices. Analogously, we say that $G \in \mathcal{G}_n$ is $*$-Hamiltonian if $G \succ_* C_n$. The reader may check that the property of being $*$-Hamiltonian is weaker than the property of being Hamiltonian; in particular, every Eulerian graph is also $*$-Hamiltonian.

We now present several open questions concerning the structure of $(\mathcal{G}_n, d_n^*)$. We would like to know:

(1) Does $(\mathcal{G}_n, d_n^*)$ ***have any non-trivial isometries?***
(2) Are there isometric embeddings of $(\mathcal{G}_{n-1}, d_{n-1}^*) \hookrightarrow (\mathcal{G}_n, d_n^*)$ ***?***

Recall that in any metric space one can define a Gromov inner product [2] relative to a base point. For the proposed metric space $(\mathcal{G}_n, d_n^*)$ we get the Gromov inner product as follows. Relative to $G \in \mathcal{G}_n$,

$$\forall H, K \in \mathcal{G}_n, \quad (H, K)_G \overset{\text{def}}{=} \frac{1}{2}\left[d_n^*(H, G) + d_n^*(K, G) - d_n^*(H, K)\right]$$

The set of graphs G for which $(H, K)_G = 0$ are said to lie on a geodesic between H and K. Specifically, given two graphs $H, K \in \mathcal{G}_n$, a map $\gamma : I \to \mathcal{G}_n$ from a closed (possibly finite) subset $I \subseteq \mathbb{R}$ into $\mathcal{G}_n$ is called a **geodesic** connecting H to K if $\gamma(\inf(I)) = H$, $\gamma(\sup(I)) = K$, and $\forall x, y \in I, |x - y| = d_n^*(\gamma(x), \gamma(y))$. A geodesic γ is said to be **non-trivial** if $|Dom(\gamma)| > 2$. A graph $L \in \mathcal{G}_n$ is said to lie on the geodesic γ if there exists an $x \in Dom(\gamma)$ for which $\gamma(x) = L$.

(3) Is there is a characterization of those pairs of graphs in $\mathcal{G}_n$ ***which possess no non-trivial geodesics between them?***

Given geodesics γ and γ', γ' is called a **refinement** of γ if $Dom(\gamma') \supsetneq Dom(\gamma)$, $\sup(Dom(\gamma)) = \sup(Dom(\gamma'))$ and $\inf(Dom(\gamma)) = \inf(Dom(\gamma'))$. If γ and γ' are two geodesics from H to K that do not possess a common refinement, then γ, γ' are said to be **independent** geodesics from H to K. We are interested to know

(4) Which pairs of graphs have more than one non-trivial and independent geodesics between them?

5 Acknowledgements

The authors gratefully acknowledge the Center for Computational Science at the Naval Research Laboratory, Washington DC, where this work began, in the context of virtual path layout for high-speed computer networks. The authors would like to thank Daniel Tunkelang for many helpful discussions and comments.

References

[1] G. Chartrand, G. Kubicki, and M. Schultz. Graph similarity and distance in graphs. *Aequationes Mathematicae*, 55:129–145, 1998.

[2] M. Coornaert, T. Delzant, and A. Papdopoulos. *Geometrie et theorie des groupes: Les groupes hyperboliques de Gromov (Lecture Notes in Mathematics No. 1441)*. Springer Verlag, 1990.

[3] R. Diestel. *Graph Theory*. Springer Verlag, 1997.

[4] O. Gerstel, I. Cidon, and S. Zaks. The layout of virtual paths in ATM networks. *IEEE/ACM Transactions on Networking*, 4(6):873–884, 1996.

[5] F. Harary. *Graph Theory*. Addison-Wesley, 1969.

[6] S. Zaks. Path layout in ATM networks—a survey. *The DIMACS Workshop on Networks in Distributed Computing, DIMACS Center, Rutgers University*, October 1997.

Kiran R. Bhutani
Department of Mathematics, The Catholic University of America.
Washington D.C. 20064, U.S.A.

Bilal Khan
Department of Mathematics, City University of New York Graduate Center,
365 5th Avenue, New York City, NY 10016, U.S.A.