Distance Between Graphs Using Graph Labelings

Kiran R. Bhutani* Bilal Khan†
Center for Computational Science
Naval Research Laboratory, Washington D.C.

Abstract

In [**?**], Fan Chung Graham investigates notion of graph labelings and related bandwidth and cut-width of such labelings when the host graph is a path graph. Motivated by problems presented in [**?**], and our investigation of designing efficient virtual path layouts for communication networks, we investigate in this note, labeling methods on graphs where the host graph is not restricted to a particular kind of graph. In [**?**], authors introduced a metric on the set of connected simple graphs of a given order which represents load on edges of host graph under some restrictions on bandwidth of such labelings. In communication networks this translates into finding mappings between guest graph and host graph in a way that minimizes the congestion while restricting the delay. In this note, we present optimal mappings between special n-vertex graphs in $\mathcal{G}_n$ and compute their distances with respect to the metric introduced in [**?**]. Some open questions are also presented.

1 Background

We recall some definitions from Bhutani and Khan [**?**]. Given an undirected graph $G = (V, E)$, recall that a **path** of length l in G is a sequence of

* Department of Mathematics, The Catholic University of America, Washington DC 20064 USA (bhutani@cua.edu).
† Advanced Engineering Sciences ITT at the Center for Computational Science, Naval Research Laboratory, Washington DC 20375 USA.

Subject Classification: 05C99, 05C12

$l+1$ distinct vertices $p=(v_0,v_1,\cdots,v_l)$, where $v_i \in V$ for $i=0,\ldots,l$, and $v_jv_{j+1} \in E$ for $j=0,\ldots,l-1$ [**?**]. We define the **boundary** of p as $\partial p=\{v_0,v_l\}$, and will denote the set of all unordered pairs of vertices in G as $V \times V$.

`Definition 1.1.` Define $\mathcal{P}_G^l$ to be the set of all paths in G of length at most l, and take $\mathcal{P}_G=\bigcup_{l=1}^{\infty}\mathcal{P}_G^l$.

`Definition 1.2.` To each set of paths $Q \subseteq \mathcal{P}_G$, associate the **path multigraph** $Q^\circ=(V,E_Q)$, where $uv \in E_Q \Leftrightarrow \exists p \in Q$ such that $\partial p=\{u,v\}$.

`Definition 1.3.` For $p \in \mathcal{P}_G$ and $e \in E$, let $m(p,e)$ denote the multiplicity of e in p. For $Q \subseteq \mathcal{P}_G$ define the congestion of Q at e as $m(Q,e)=\sum_{p\in Q} m(p,e)$. Finally, given an undirected graph G, we define the **congestion** of Q on G as $\tau_G(Q)=\max_{e\in E} m(Q,e)$.

`Definition 1.4.` Let $\mathcal{G}_n$ be the set of all simple, connected, undirected graphs (up to isomorphism) on n vertices. For each positive integer ℓ and graphs H,K in $\mathcal{G}_n$, the **ℓ-embedding thickness** of K in H, denoted $e_n^\ell(H,K)$, was defined as follows: If there exists a set $Q \subseteq \mathcal{P}_H^\ell$ such that $Q^\circ \simeq K$ then consider Q for which $\tau_H(Q)=2^x$ is minimal, and set $e_n^\ell(H,K)=x$, otherwise set $e_n^\ell(H,K)=\infty$. We will write $H \succ_\ell K$ if $e_n^\ell(H,K)=0$.

The **embedding thickness** of K in H, denoted by $e_n^*(H,K)$, is obtained as above except that $Q \subseteq \mathcal{P}_H$; that is, Q is a set of paths of arbitrary lengths. We will write $H \succ_* K$ when $e_n^*(H,K)=0$.

We say that a bijection $\alpha : V(K) \to V(H)$ is optimal for mapping graph K onto H if the set of paths Q^* in H between $\{\alpha(u),\alpha(v)\}|(u,v)\in E(K)\}$ satisfying $(Q^*)^\circ \simeq K$ is such that $\tau_H(Q^*)$ is minimal. We call such set of paths as optimal set of paths in H for a given K.

`Remark 1.5.` Note that $e_n^1(H,K)=0$ implies that there is a set of edge-disjoint paths $Q \subseteq \mathcal{P}_H^1=E[H]$ such that $Q^\circ \simeq K$, and so $|E[K]| \leqslant |E[H]|$ and K is a connected spanning subgraph of H.

Furthermore, since $\tau_H(Q) \leqslant |E[K]|$ which is at most $\frac{n(n-1)}{2}$, it follows that $e_n^\ell(H,K)$ is bounded above by $2\log n$. Any fixed real number $> 2\log n$ can be used in place of ∞ in the above definition.

Subject Classification: 05C99, 05C12

Definition 1.6. For any graphs H, K in $\mathcal{G}_n$, we define their **ℓ-distance** and **distance**, respectively, as follows:

$$\begin{aligned} d_n^\ell(H,K) &= e_n^\ell(H,K) + e_n^\ell(K,H) \\ d_n^*(H,K) &= e_n^*(H,K) + e_n^*(K,H). \end{aligned}$$

Note that $d_n^\ell(H,K)$ may be infinity—for example, when $\ell = 1$ and K is a proper connected spanning subgraph of H. In [?] authors have shown that when $\ell = 1$ or $\ell \geqslant n-1$, d_n^ℓ is a metric on $\mathcal{G}_n$. While $(\mathcal{G}_n, d_n^1)$ is a totally disconnected metric space that embodies the classical notion of graph isomorphism, $(\mathcal{G}_n, d_n^\ell)$ is a connected metric space whenever $\ell \geqslant n-1$. Remark 1.5 implies:

Lemma 1.7. *$H \succ_1 K \Leftrightarrow K$ is a connected spanning subgraph of H.*

2 Distance between special graphs in $\mathcal{G}_n$

Let $K_n, S_n, C_n, P_n \in G_n$ be the complete graph, the star, the cycle, and the chain on n vertices.

Proposition 2.1. $d_n^*(P_n, K_n) = \log_2(\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil)$

Proof. Since P_n is a subgraph of K_n, lemma 1.7 implies $e_n^*(K_n, P_n) = 0$. To compute $e_n^*(P_n, K_n)$, by Definition 1.4 it suffices to find a set of *paths* $Q \subset \mathcal{P}_{P_n}$ such that $Q^\circ \simeq K_n$ and $\tau_{P_n}(Q)$ is minimal. Let P_n consist of vertices $v_1, \ldots, v_n$, with edges $v_i v_{i+1}$, $i = 1, \ldots, n-1$; then a *path* between v_i and v_j, $(i < j)$ is just the subchain of P_n induced by the vertices $v_i, v_{i+1}, \ldots, v_{j-1}, v_j$; we denote this subchain as $P_{i,j}$. Since $Q^\circ \simeq K_n$, it follows that Q must be the set $\{P_{i,j} \mid 1 \leqslant i < j \leqslant n\} \subseteq \mathcal{P}_{P_n}$. Observe that $\tau_{P_n}(Q) = \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$, so $e_n^*(P_n, K_n) = \log_2(\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil)$. Finally, $d_n^*(P_n, K_n) = e_n^*(P_n, K_n) + e_n^*(P_n, K_n) = \log_2(\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil)$. $\square$

Proposition 2.2. $d_n^*(S_n, K_n) = \log_2(n-1)$

Proof. Since S_n is a subgraph of K_n, lemma 1.7 implies $e_n^*(K_n, S_n) = 0$. To compute $e_n^*(S_n, K_n)$, we find a set of *paths* $Q \subset \mathcal{P}_{S_n}$ such that $Q^\circ \simeq K_n$ and $\tau_{S_n}(Q)$ is minimal. Let S_n consists of vertices $v_0, \ldots, v_{n-1}$, with edges

---

Subject Classification: 05C99, 05C12

(v_0, v_i), $i = 1, \ldots, n-1$, then a *path* between v_i and v_j, $(i < j)$ is of the form (v_i, v_0, v_j) if $i, j \neq 0$, and is an edge of the form $v_i v_0$ or $v_0 v_j$ when $j = 0$ or $i = 0$, respectively. Since $Q^\circ \simeq K_n$, it follows that

$$Q = \{(v_i, v_0, v_j) \mid 0 \neq i < j \neq 0\} \cup \{(v_i, v_0) \mid i \neq 0\} \cup \{(v_0, v_j) \mid j \neq 0\}$$

Since there are $(n-1)$ paths connecting a vertex v_i to all other $(n-1)$ vertices, all these paths contain the edge $v_0 v_i$. Hence the congestion on the edge $v_o v_i$ is $(n-1)$, which means that $\tau_{S_n}(Q) = n-1$, so $e_n^*(S_n, K_n) = \log_2(n-1)$ Finally, $d_n^*(S_n, K_n) = e_n^*(S_n, K_n) + e_n^*(P_n, K_n) = \log_2(n-1)$. $\square$

`Proposition 2.3.` $d_n^*(S_n, P_n) = 1 + log_2(\lceil \frac{n}{2} \rceil)$

Proof. Let S_n consist of vertices $v_0, \ldots, v_{n-1}$, with edges (v_0, v_i), $i = 1, \ldots, n-1$, and take P_n to consist of vertices $v_1, \ldots, v_n$, with edges (v_i, v_{i+1}), $i = 1, \ldots, n-1$.

To compute $e_n^*(S_n, P_n)$, it suffices to find a set of *paths* $Q \subset \mathcal{P}_{S_n}$ such that $Q^\circ \simeq P_n$ and $\tau_{S_n}(Q)$ is minimal. Take

$$Q = \{(v_i, v_0, v_{i+1} \mid i = 1, \ldots n-2\} \cup \{(v_{n-1}, v_0)\}$$

Now it is easy to check that $\tau_{S_n}(Q) = 2$, and $Q^\circ \simeq P_n$, which implies that $e_n^*(S_n, P_n) \leqslant log_2 2 = 1$. On the other hand, $e_n^*(S_n, P_n)$ cannot be < 1 since this would mean $e_n^*(S_n, P_n) = 0$, and then lemma 1.7 would imply that P_n is a subgraph of S_n, a contradiction. Thus, $e_n^*(S_n, P_n) = 1$.

To compute $e_n^*(P_n, S_n)$, by Definition 1.4, it suffices to find a set of *paths* $R \subset \mathcal{P}_{P_n}$ such that $R^\circ \simeq S_n$ and $\tau_{P_n}(R)$ is minimal. We note that the *path* between v_i and v_j, $(i < j)$ is just the subchain of P_n induced by the vertices $v_i, v_{i+1}, \ldots, v_{j-1}, v_j$; we denote this subchains as $P_{i,j}$. First, if R is such a set of paths for which $R^\circ \simeq S_n$, it must be that

$$R = \{P_{i,m} \mid i = 1, \ldots, n, i \neq m\}$$

for some $m \in \{1, \ldots, n\}$. Observe that $\tau_{P_n}(R)$ is minimized when $m = \lceil \frac{n}{2} \rceil$, and that for the resulting set of paths R, $\tau_{P_n}(R) = \lfloor \frac{n}{2} \rfloor$. Thus, $e_n^*(P_n, S_n) = \log_2(\lfloor \frac{n}{2} \rfloor)$.

Finally, $d_n^*(S_n, P_n) = e_n^*(S_n, P_n) + e_n^*(P_n, S_n) = 1 + log_2(\lfloor \frac{n}{2} \rfloor)$. $\square$

Subject Classification: 05C99, 05C12

Proposition 2.4. $d_n^*(P_n, C_n) = 1$

Proof. Since P_n is a subgraph of C_n, lemma 1.7 implies $e_n^*(C_n, P_n) = 0$. If P_n consists of vertices $v_1, \ldots, v_n$, with edges (v_i, v_{i+1}), $i = 1, \ldots, n-1$, take $Q \subseteq \mathcal{P}_{P_n}$ to be $\{(v_i, v_{i+1}) \mid i = 1, \ldots, n-1\} \cup \{(v_1, v_2, \ldots, v_n)\}$. Then it is easy to check that $Q^\circ \simeq C_n$ and $\tau_{P_n}(Q) \leqslant 2$, so $e_n^*(P_n, C_n) \leqslant 1$. On the other hand, $e_n^*(P_n, C_n)$ cannot be < 1 since this would mean $e_n^*(P_n, C_n) = 0$, and then lemma 1.7 would imply that C_n is a subgraph of P_n, a contradiction. Thus, $e_n^*(P_n, C_n) = 1$, and so $d_n^*(P_n, C_n) = e_n^*(P_n, C_n) + e_n^*(C_n, P_n) = 1$. $\square$

Proposition 2.5. $d_n^*(S_n, C_n) = 1 + \log_2(\lceil \frac{n-1}{2} \rceil)$

Proof. Let S_n consist of vertices $v_0, \ldots, v_{n-1}$ with edges (v_0, v_i), $i = 1, \ldots, n-1$ and take C_n to consist of vertices $v_1, \ldots, v_n$, with edges (v_1, v_n) and (v_i, v_j), $|i - j| = 1$ for $i, j \in \{1 \ldots, n\}$.

To compute $e_n^*(S_n, C_n)$, by Definition 1.4, it suffices for find a set of *paths* $Q \subset \mathcal{P}_{S_n}$ such that $Q^\circ \simeq C_n$ and $\tau_{S_n}(Q)$ is minimal. Take

$$Q = \{(v_i, v_0, v_{i+1}) \mid i = 1, \ldots n-1\} \cup \{(v_{n-1}, v_0), (v_0, v_1)\}$$

Now it is easy to check that $\tau_{S_n}(Q) = 2$, and $Q^\circ \simeq C_n$, which implies that $e_n^*(S_n, C_n) \leqslant log_2 2 = 1$. On the other hand, $e_n^*(S_n, C_n)$ cannot be < 1 since this would mean $e_n^*(S_n, C_n) = 0$, and then lemma 1.7 would imply that C_n is a subgraph of S_n, a contradiction. Thus $e_n^*(S_n, C_n) = 1$.

To compute $e_n^*(C_n, S_n)$, it suffices for find a set of *paths* $Q \subset \mathcal{P}_{C_n}$ such that $Q^\circ \simeq S_n$ and $\tau_{C_n}(Q)$ is minimal. Take

$$\begin{aligned} Q = &\{(v_1, v_2, \ldots, v_{i-1}, v_i) \mid i = 2, 3, \ldots, \lceil n/2 \rceil\} \\ &\cup \{(v_1, v_{n-1}, v_{n-2}, \ldots, v_{j+1}, v_j) \mid j = \lceil n/2 \rceil + 1, \ldots, n-1\} \end{aligned}$$

Now it is easy to check that $\tau_{C_n}(Q) = \lceil (n-1)/2 \rceil$, and $Q^\circ \simeq S_n$, which implies that $e_n^*(C_n, S_n) \leqslant \log_2(\lceil \frac{n-1}{2} \rceil)$. On the other hand, $e_n^*(C_n, S_n)$ must be at least this large, by the following pigeonhole argument: Suppose $i : S_n \to Q^\circ$ is an isomorphism. Let $z \stackrel{\text{def}}{=} i(v_0)$; then $z = v_k \in C_n$ for some $k \in \{1, \ldots, n\}$. But then $m(Q, e1) + m(Q, e2) = n - 1$, where $e_1 = (v_{(k-1) \bmod n}, v_k)$ and $e_2 = (v_k, v_{(k+1) \bmod n})$. It follows that $\tau_{C_n}(Q) \geqslant \lceil \frac{n-1}{2} \rceil$. Hence $e_n^*(C_n, S_n) = \log_2(\lceil \frac{n-1}{2} \rceil)$.

In conclusion, $d_n^*(S_n, P_n) = e_n^*(S_n, P_n) + e_n^*(P_n, S_n) = 1 + \log_2(\lceil \frac{n-1}{2} \rceil)$. $\square$

Subject Classification: 05C99, 05C12

Proposition 2.6. $d_n^*(C_n, K_n) = \begin{cases} \log_2\left(\frac{n^2-1}{8}\right) & \textit{if } n \textit{ is odd.} \\ \log_2\left(\frac{n(n-2)}{8} + \lfloor\frac{n}{4}\rfloor + 1\right) & \textit{if } n \textit{ is even.} \end{cases}$

Proof. Take C_n to consist of vertices $v_0, \dots, v_{n-1}$, with edges (v_i, v_j), $|i-j| = 1$ and (v_{n-1}, v_0) for $i, j \in \{0, 1 \dots, n-1\}$. Since C_n is a subgraph of K_n, lemma 1.7 implies $e_n^*(K_n, C_n) = 0$. To compute $e_n^*(C_n, K_n)$, by Definition 1.4, it suffices to find a set of *paths* $Q \subset \mathcal{P}_{C_n}$ such that $Q^\circ \simeq K_n$ and $\tau_{C_n}(Q)$ is minimal.

Case I: n is odd. Take

$$Q = \{(v_i, v_{(i+1) \bmod n}, \dots, v_{(i+d) \bmod n}) \mid i = 0, \dots, n-1. \ d = 1, \dots, \left\lfloor\frac{n}{2}\right\rfloor\}$$

Then $Q^\circ \simeq K_n$ and for any edge $e \in E[C_n]$, $m(Q, e) = \sum_{i=1}^{\lfloor\frac{n}{2}\rfloor} i = \frac{n^2-1}{8}$, so $\tau_{C_n}(Q) = \frac{n^2-1}{8}$. Since $Q^\circ \simeq K_n$,

it follows $e_n^*(C_n, K_n) \leqslant \log_2\left(\frac{n^2-1}{8}\right) \cdots\cdots\cdots\cdots (i)$

Further, since in C_n there are n pairs of points at distances $1, 2, \cdots, \frac{n-1}{2}$ respectively, it follows that the number of edges of C_n used to form paths in Q° such that $Q^\circ \simeq K_n$ is

$$n(1 + 2 + \cdots + \frac{n-1}{2}) = \frac{n(n^2-1)}{8}$$

Since C_n has exactly n edges, it follows by pigeon hole principal that at least one of the edges should be used at least $\frac{(n^2-1)}{8}$ times. Therefore $e_n^*(C_n, K_n) \geqslant \log_2\left(\frac{n^2-1}{8}\right)$, and so from (i) above we get $e_n^*(C_n, K_n) = \log_2\left(\frac{n^2-1}{8}\right)$. Since $e_n^*(K_n, C_n) = 0$ it follows $d_n^*(C_n, K_n) = e_n^*(K_n, C_n) + e_n^*(C_n, K_n) = \log_2\left(\frac{n^2-1}{8}\right)$.

Case II: n is even. Take $Q = Q_1 \cup Q_2 \cup Q_3$, where

$$Q_1 = \{(v_i, v_{(i+1) \bmod n}, \dots, v_{(i+d) \bmod n}) \mid i = 0, \dots, n-1; \ d = 1, \dots, (\frac{n}{2} - 1)\}$$

$$Q_2 = \{(v_i, v_{(i+1) \bmod n}, \dots, v_{(i+\frac{n}{2}) \bmod n}) \mid 0 \leqslant i \leqslant \frac{n}{2}, i \text{ even } \}$$

$$Q_3 = \{(v_i, v_{(i-1) \bmod n}, v_{(i-2) \bmod n}, \dots, v_{(i-\frac{n}{2}) \bmod n}) \mid 0 \leqslant i \leqslant \frac{n}{2}, i \text{ odd } \}$$

Subject Classification: 05C99, 05C12

Then $Q^\circ \simeq K_n$ and for any edge $e \in E[C_n]$, $m(Q,e) = m(Q_1,e)+m(Q_2,e)+m(Q_3,e)$, where $m(Q_1,e) = \sum_{i=1}^{\frac{n}{2}-1} i = \frac{n(n-2)}{8}$ and $m(Q_2,e)+m(Q_3,e) \leqslant \left\lfloor \frac{n}{4} \right\rfloor + 1$. So $\tau_{C_n}(Q) \leqslant \frac{n(n-2)}{8} + \left\lfloor \frac{n}{4} \right\rfloor + 1$, and therefore $e_n^*(C_n, K_n) \leqslant \log_2 \left(\frac{n(n-2)}{8} + \left\lfloor \frac{n}{4} \right\rfloor + 1 \right) \cdots\cdots\cdots\cdots\cdots\cdots\cdots (i)$

Further when n is even there are n pairs of points at distances $1, 2, \cdots, \frac{n-2}{2}$ respectively and $\frac{n}{2}$ points at distance $\frac{n}{2}$.

Therefore in order to connect all pairs we need paths of the total length at least

$$n(1+2+\cdots+\frac{n-2}{2}) + \frac{n^2}{4} = n\left(\frac{n(n-2)}{8} + \frac{n}{4}\right)$$

Since there are only n edges in C_n, at least one of the edges should be used at least $(\frac{n(n-2)}{8} + \left\lfloor \frac{n}{4} \right\rfloor + 1)$ times. This shows that $e_n^*(C_n, K_n) \geqslant \log_2 \left(\frac{n(n-2)}{8} + \left\lfloor \frac{n}{4} \right\rfloor + 1 \right) \cdots\cdots\cdots\cdots\cdots\cdots\cdots (ii)$

From (i) and (ii) we get $e_n^*(C_n, K_n) = \log_2 \left(\frac{n(n-2)}{8} + \left\lfloor \frac{n}{4} \right\rfloor + 1 \right)$ and hence $d_n^*(C_n, K_n) = \log_2 \left(\frac{n(n-2)}{8} + \left\lfloor \frac{n}{4} \right\rfloor + 1 \right)$. □

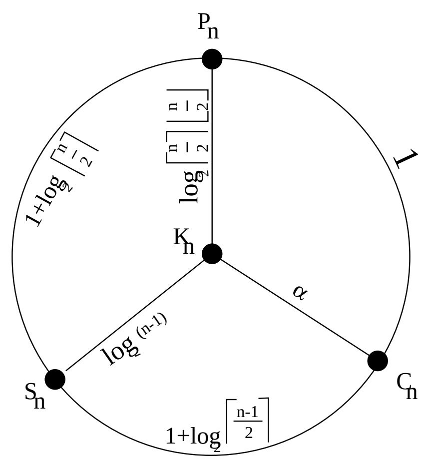

Figure 1: The distances between P_n, C_n, S_n, and K_n in $(\mathcal{G}_n, d_n^*)$.

`Remark 2.7.` The calculations of the previous results 2.1-2.6 are summarized in figure 1 using an undirected graph whose vertices are the points

Subject Classification: 05C99, 05C12

$P_n, C_n, S_n, K_n \in \mathcal{G}_n$ and whose edges are weighted by the distance between its endpoints in $(\mathcal{G}_n, d_n^*)$. In the above figure α is given by Proposition 2.6.

Open Questions

(1) If s_G for any $G \in \mathcal{G}_n$ denotes a selected subset of k vertices of a graph G, then is it possible to find an optimal map ϕ from $V(G)$ into $V(H)$ such that $\phi(s_G) = s_H$?

(2) If f_1 and f_2 are optimal maps for (H_1, K_1) and (H_2, K_2) in $\mathcal{G}_n$ and $\mathcal{G}_m$ respectively, then is $f_1 \times f_2$ an optimal map for $(H_1 \times H_2, K_1 \times K_2)$ in $\mathcal{G}_{mn}$?

(3) Given a graph G and a natural number n, can we find a graph H possessing a special property (e.g. Hamilton, Euler) with $d_n^*(G, H) \leq log_2 n$?

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 Professor Mikhail Ostrovskii for helpful discussions with proof of Proposition 2.6.

Subject Classification: 05C99, 05C12