Average distance and maximum induced forest

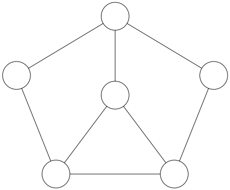
Prof. Pierre Hansen, HEC Montréal
Prof. Alain Hertz, École Polytechnique de Montréal
Rim Kilani, Ph.D. student
Prof. Odile Marcotte, Université du Québec à Montréal
David Schindl, Postdoctoral student

GERAD

Definitions

Connected graph : $G = (V, E), n = |V|$

$$\mu(G) = \text{average distance among all pairs of vertices of } G$$
$$= \frac{\sum_{v \neq w} d(v,w)}{\frac{n(n-1)}{2}}$$

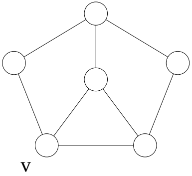


$$\mu(G) = \frac{22}{15} \simeq 1.47$$

Definitions

Connected graph : $G = (V, E), G' \subseteq G, v \in V$

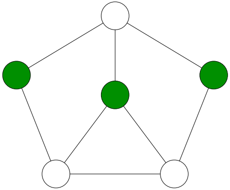
$$\begin{aligned} \sigma_{G'}(v) &= \text{sum of all distances from } v \text{ to all other vertices in } G' \\ &= \text{transmission of } v \text{ in } G' \\ &= \sum_{w \in V(G')} d(v, w) \end{aligned}$$



$\sigma_G(v) = 7$

Definitions

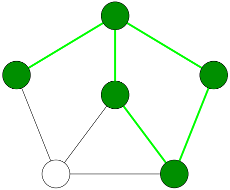
$$\alpha(G) = \text{size of a maximum stable set in } G$$
$$= \text{stability number of } G$$



$$\alpha(G) = 3$$

Definitions

$$\alpha_2(G) = \text{size of a maximum induced bipartite subgraph in } G$$
$$= \text{2-stability number of } G$$



$$\alpha_2(G) = 5$$

2 disjoint stable sets induce a bipartite graph $\Longrightarrow \frac{\alpha_2(G)}{2} \leq \alpha(G)$

Conjecture 747

From "Written on the Wall" (conjectures automatically generated with the help of the Graffiti system, S. Fajtlowicz) :

Summer 92.

747. *Let b be the order of a largest bipartite subgraph of a connected graph G. Then the average distance between distinct vertices of G is not more than $b/2$.*

If correct this conjecture would generalize conj. 2 that the average distance is not more than the independence number. This conjecture was proved by **Fan Chung.**

Conjecture 747

Summer 92.

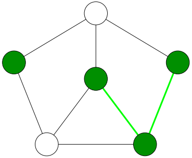
747. *Let b be the order of a largest bipartite subgraph of a connected graph G. Then the average distance between distinct vertices of G is not more than $b/2$.*

If correct this conjecture would generalize conj. 2 that the average distance is not more than the independence number. This conjecture was proved by **Fan Chung.**

In our terms : $\mu(G) \leq \frac{\alpha_2(G)}{2}$?

Fan Chung Theorem : $\mu(G) \leq \alpha(G)$

$$
\begin{aligned}
F(G) &= \text{size of a maximum induced forest in } G. \\
&= \text{forest number of } G \\
&= |V|\text{- minimum number of vertices needed to cover all cycles}
\end{aligned}
$$

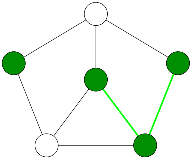


$$F(G) = 4$$

A forest is special bipartite graph $\Longrightarrow F(G) \leq \alpha_2(G)$

Computing $F(G)$ is NP-hard

$$F(G) = \text{size of a maximum induced forest in } G.$$
$$= \text{forest number of } G$$



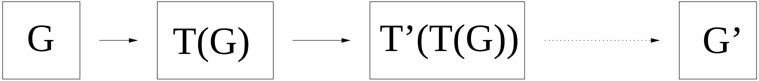
$$F(G) = 4$$

A forest is a special bipartite graph $\Longrightarrow F(G) \leq \alpha_2(G)$

Our theorem : $\mu(G) \leq \frac{F(G)}{2}$

Corollary : $\mu(G) \leq \frac{\alpha_2(G)}{2}$

Our constructive proof :



Each transformation T has the following two properties :

- $\mu(T(G)) \geq \mu(G)$
- $F(T(G)) = F(G)$

The graph G' obtained at the end is very specific and verifies

$$\mu(G') \leq \frac{F(G')}{2}$$

$$\Rightarrow \mu(G) \leq \mu(G') \leq \frac{F(G')}{2} = \frac{F(G)}{2}$$

Structure :

Input : G connected
Output : G' such that $\mu(G) \leq \mu(G')$, $F(G') = F(G)$ and $|V(G')| = |V(G)|$

1: Set $G' = G$
2: If T_1 is applicable to G', set $G' = T_1(G')$ and go to 2
3: If T_2 is applicable to G', set $G' = T_2(G')$ and go to 2
4: If T_3 is applicable to G', set $G' = T_3(G')$ and go to 2
5: If T_4 is applicable to G', set $G' = T_4(G')$ and go to 2
6: If T_5 is applicable to G', set $G' = T_5(G')$ and go to 2
7: Set $G' = T_6(G')$ as many times as T_5 has been applied.

For T_1, T_2, T_3, T_4 and T_6, we have :

- $\mu(T_i(G)) > \mu(G)$
- $F(T_i(G)) = F(G)$

Moreover,

- $\mu(T_5(G)) \geq \mu(G)$ (there is a possible equality)
- $F(T_5(G)) = F(G)$

Since T_6 is applied as often as T_5, we conclude that :

If G' is not equal to G, then

$$\mu(G) < \mu(G') \leq \frac{F(G')}{2} = \frac{F(G)}{2}$$

Claim : the graph obtained at the end of the algorithm is a **balanced double kite**.

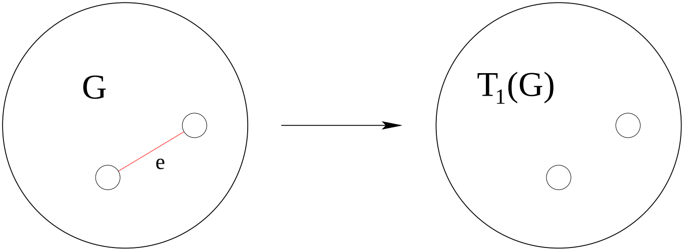
Both cliques have the same (or almost the same) number of vertices.

T_1 : edge removal

An edge e such that $F(G - e) > F(G)$ is called **critical**.

An edge e such that G is connected and $G - e$ is disconnected is called a **bridge**.

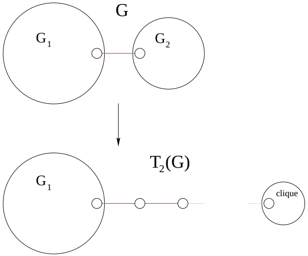
Requirement : $\exists$ a non-critical edge e which is not a bridge.



Property : if every edge is critical, then G is 2-(vertex-)connected.

T_2 : kite creation

Requirement : G contains a bridge (edge e such that $G - e$ is not connected).

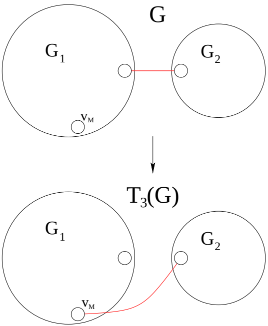


- Length of the path chosen so that $F(T_2(G)) = F(G)$.
- Size of the clique chosen so that $|V(T_2(G))| = |V(G)|$.

Limitation : only applicable to the smallest side of the bridge.

T_3 : bridge shifting

Requirement : G contains a bridge.

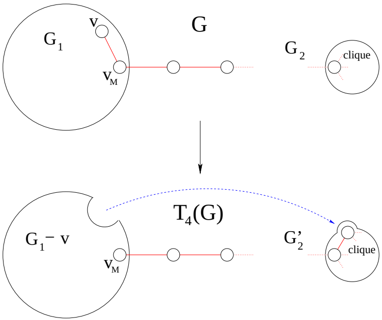


$$v_M \text{ : vertex of } G_1 \text{ such that } \sigma_{G_1}(v_M) = \max_{v \in G_1} \sigma_{G_1}(v).$$

T_4 : vertex shifting

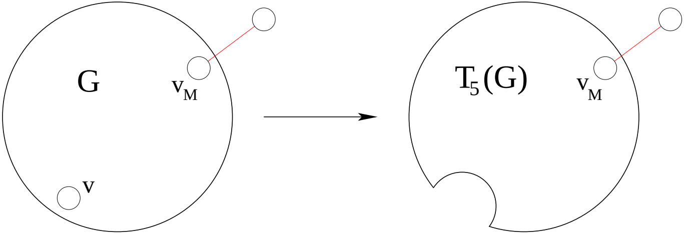
Requirements :

- G contains a bridge with end point v_M in G_1
- G_2 is a kite
- $|V(G_1)| > |V(G_2)|$
- v_M has at least three neighbors



T_5 : vertex removal

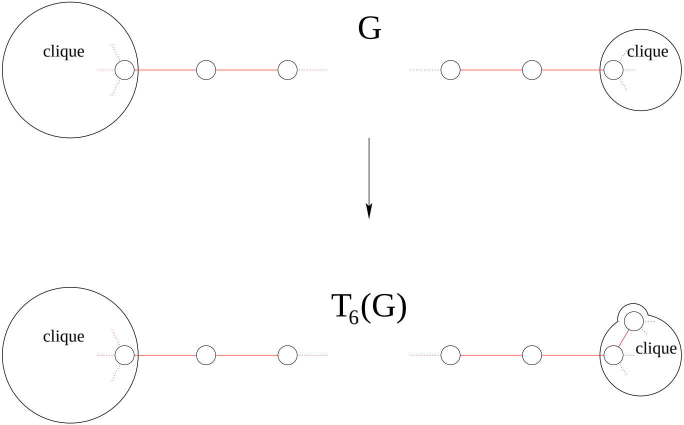
Requirement : there is at most one non-critical edge of G, and it is a pending edge.



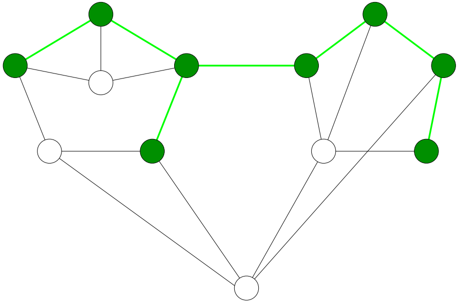
v : "central" vertex of G.

T_6 : double kite increasing

Requirement : G is a double kite.



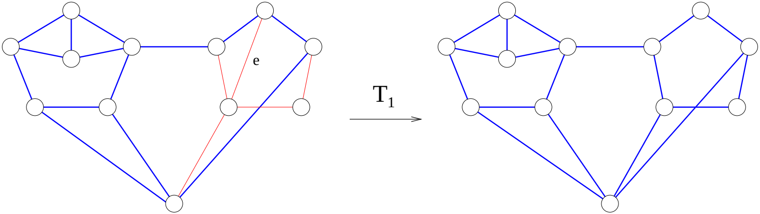
An example



$$\mu(G) = 2.14$$

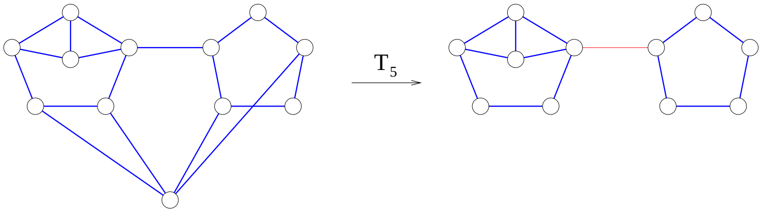
$$F(G) = 8$$

An example



$$\mu(G) = 2.14 \qquad \mu(G') = 2.15$$
$$F(G) = 8 \qquad F(G') = 8$$

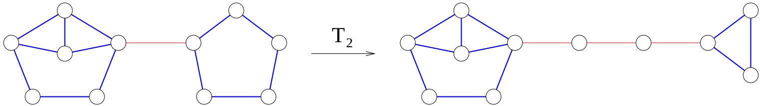
An example



$$\mu(G') = 2.15$$
$$F(G') = 8$$

$$\mu(G') = 2.51$$
$$F(G') = 8$$

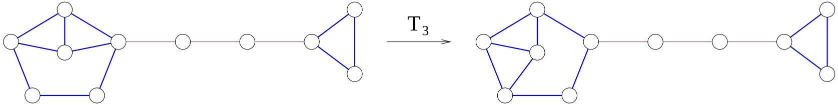
An example



$$\mu(G') = 2.51 \qquad F(G') = 8$$

$$\mu(G') = 2.87 \qquad F(G') = 8$$

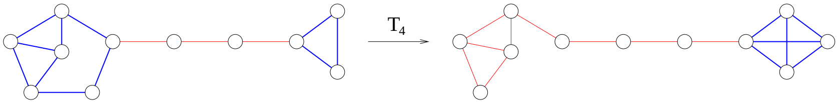
An example



$$\mu(G') = 2.87$$
$$F(G') = 8$$

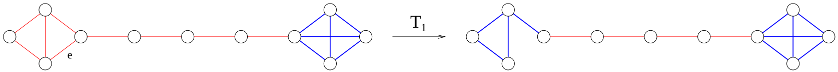
$$\mu(G') = 2.96$$
$$F(G') = 8$$

An example



$$\mu(G') = 2.96 \qquad \mu(G') = 3.24$$
$$F(G') = 8 \qquad F(G') = 8$$

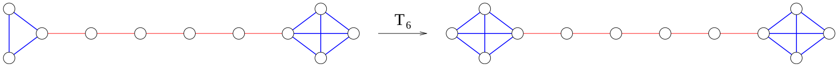
An example



$$\mu(G') = 3.24 \qquad F(G') = 8$$

$$\mu(G') = 3.38 \qquad F(G') = 8$$

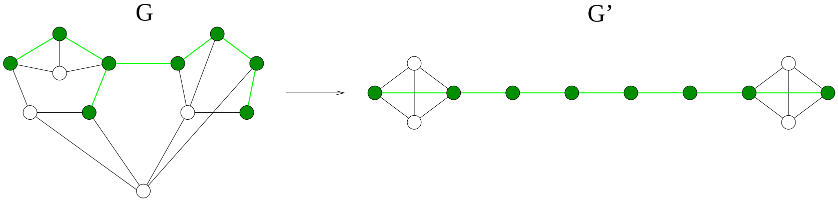
An example



$$\mu(G') = 3.38$$
$$F(G') = 8$$

$$\mu(G') = 3.49$$
$$F(G') = 8$$

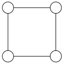
An example



$$\mu(G) \leq \mu(G') = 3.49 \leq 4 = \frac{F(G')}{2} = \frac{F(G)}{2}$$

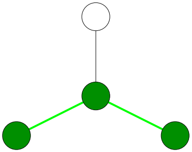
Extremal graphs

The graphs $G = (V, E)$ with $|V| = n$ vertices such that $F(G) = k$ and maximum $\mu(G)$ are

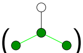
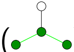
- the balanced double kite if $k \geq 4$
- the clique K_n if $k \leq 2$
- the balanced double kite or some other graphs like  if $k = 3$

A stronger conjecture

$LF(G)$=size of a maximum induced linear forest in G.



$$LF(G) = 3$$

$\mu($  $) = 1.5 \leq 1.5 = \dfrac{LF(\text{ } }$  $\text{ })}{2}$

Conjecture

$$???\ \mu(G) \leq \frac{LF(G)}{2}\ ???$$

Questions ?