On the Graph of Trees

C. K. Enyioha, D. C. Tarraf L. Li, and J. C. Doyle

Abstract— We consider an "n-graph of trees" whose nodes
are the set of trees of fixed order n, and in which two nodes are
adjacent if one tree can be derived from the other through a
single application of a local edge transformation rule. We derive
an exact formula for the length of the shortest path from any
node to any "canonical" node in the n-graph of trees. We use
this result to derive upper and lower bounds on the diameter
of the n-graph of trees. We then propose a coordinate system
that is convenient for studying the structure of the n-graph of
trees, and in which trees having the same degree sequence are
projected onto a single point.

I. INTRODUCTION

Networks abound in both natural and man-made systems,
with gene networks and social networks being examples
of the former and the internet being an example of the
latter. Graph theoretic models of networks are often used
for studying the topological properties of the network. In
this context, the networks of interest are typically complex,
both due to their large size and their evolving structure.
Network evolution has been an object of focus in many
fields, including the study of evolving random graphs in
statistical physics and mathematics [3], [1], [4], [2], graph
grammars in computer science and engineering [8], [6], and
graph transformations in mathematics and engineering [10],
[12], [5]. In particular, graphs that evolve according to certain
edge, node, and edge/node transformation rules have been
considered [10]. In this context, it has been recognized that
certain degree preserving (sometimes called "switching")
transformations can be used to define a metric on the space
of graphs reachable by these rules [11].

In this paper, we consider a problem in which we start
with a labeled tree with n nodes. We allow the tree to evolve
according to a local edge transformation rule referred to as a
"general flip" [7]. We then construct an "n-graph of trees",
whose nodes are the set of all trees of size n, and in which
two nodes are adjacent if one tree can be derived from the
other through a general flip. We derive an exact formula
for the length of the shortest path from any node to any of
the "canonical" nodes (which will be defined precisely in
the next section) in this n-graph of trees. We use this result

to derive upper and lower bounds on the diameter of the
n-graph of trees. We then propose a coordinate system in
which to study a projection of the n-graph of trees, and in
which nodes corresponding to same degree sequence trees
are projected onto a single point.

II. PRELIMINARIES

A. Notation

The notation used in this paper is standard: $\mathbb{N}$ refers to
the set of natural numbers. Given a set A, 2^A refers to the
set of all subsets of A. $A \setminus B$ is the set of all elements of A
that are not elements of B.

B. Brief Review of Graph Theoretic Concepts

For the sake of completeness, we begin by reviewing some
relevant elementary graph theoretic concepts. An undirected
graph is an ordered pair $G = (V, E)$ where V is a set of
nodes and E is a set of unordered pairs of nodes, referred to
as edges. The order and size of the graph are the cardinality
of V and E, respectively. A graph is said to be labelled if
each of its nodes is assigned a unique integer between 1 and
n, where n is the cardinality of V, referred to as the order
of the graph. Two nodes i and j are said to be adjacent if
they form an edge, that is $\{i, j\} \in E$. A self loop is an edge
that begins and ends at the same node.

A path between nodes i and j is a sequence of edges
of the form $\{i, l\}, \{l, k\}, ..., \{o, m\}, \{m, j\}$. The length of the
path is the number of edges in the sequence. A graph is
said to be connected if there exists a path between any two
nodes, and complete if there exists an edge between any two
distinct nodes. A cycle is a path that begins and ends at the
same node. A graph is simple if it has no self loops and
no repeated edges. The degree of a given node in a simple
graph is the number of nodes adjacent to it.

The shortest path between two nodes i and j is the (non-
unique) path of (unique) minimum length, which we denote
by l_{ij}. The diameter D of a connected graph is the length
of the longest shortest path between two nodes of the graph,
that is, $D = \max_{i,j} l_{ij}$.

A simple, undirected, connected (labeled) graph with no
cycles is called a (labeled) tree.

C. Definitions

Given a labeled tree G with node set $V = \{1, 2, , \ldots, n\}$,
the degree sequence of G is the sequence of node degrees
$(d_1, d_2, \ldots, d_n)$. In particular, we will refer to trees of
order n whose degree sequences are permutations of $(n -$

C. K. Enyioha is a graduate student in the Electrical and
Systems Engineering Department at the University of Pennsylvania
(cenyioha@seas.upenn.edu).

D. C. Tarraf (corresponding author) is with the Electrical and
Computer Engineering Department at the Johns Hopkins University
(dtarraf@jhu.edu).

L. Li is a visitor at the Division of Control and Dynamical Systems at
the California Institute of Technology (lun@cds.caltech.edu).

J. C. Doyle is with the Division of Control and Dynamical Systems at
the California Institute of Technology (doyle@cds.caltech.edu).

$1, 1, 1, \ldots, 1$) as *canonical trees*, and we will refer to the degree $n - 1$ node in a canonical tree as the *star node*.

We use the notation $\mathbb{T}(n)$ to refer to the set of all labeled trees with n nodes. The cardinality of $\mathbb{T}(n)$ is known to be n^{n-2} according to Cayley's formula [9].

Definition 1: Consider the map

$$f : \mathbb{T}(n) \to 2^{\mathbb{T}(n)}$$

such that $G' \in f(G)$ if there exists $\{i, j\}, \{i, k\}$ in E such that

$$E' = \left(E \setminus \{\{i, j\}\} \right) \cup \{\{j, k\}\}$$

If $G' \in f(G)$, we say that G' can be derived from G by a general flip.

This map essentially defines a class of graph transformations that replace one of a pair of adjacent edges with common node i by another edge as shown in Figure 1. Note that it follows directly from the definition that
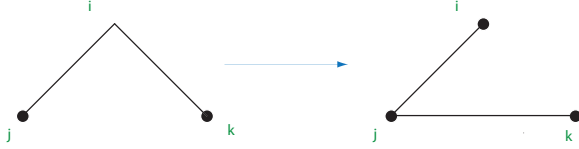
$$G \in f(G') \Leftrightarrow G' \in f(G)$$

Fig. 1. **An example of a tree derived from another by a general flip**

This class of graph transformations was shown to be sound and general in [7]: In other words, a general flip transforms a labeled connected graph into a another labeled connected graph of same order and size. Moreover, it is possible to transform any labeled graph of order n and size m to any other labeled graph of the same order and size by a sequence of general flips. It can be similarly shown that this class of transformations is sound and general when the graphs in questions are trees with n nodes. We are now ready to define the graph of all trees of size n, referred to as the n-graph of trees.

Definition 2: The n-graph of trees, denoted $\mathcal{G}(n)$, is a connected, undirected graph $\mathcal{G}(n) = (V, E)$ with nodes:

$$V = \mathbb{T}(n)$$

and edges:

$$\{G, G'\} \in E \Leftrightarrow G \in f(G') \Leftrightarrow G' \in f(G)$$

The n-graph of trees is thus a connected undirected graph whose nodes correspond to the labeled trees of order n and in which two nodes are adjacent if and only if each of them can be derived from the other by a general flip.

III. STRUCTURE OF THE n-GRAPH OF TREES

We begin by deriving an exact expression for the length of the shortest path between an arbitrary node (corresponding to a tree G) and a "canonical" node (corresponding to a canonical tree G') in the n-graph of trees.

Lemma 1: Consider $G \in \mathbb{T}(n)$, and let $G' \in \mathbb{T}(n)$ be a canonical tree with star node j. The length of the shortest path between G and G' in $\mathcal{G}(n)$ equals the total number of nodes in G not adjacent to node j, not counting node j.

Proof: Let m be the total number of nodes of G that are distinct from node j and that do not share an edge with it. In order to transform G into G', at least one edge needs to be "flipped" for each of the m nodes. Thus m is a lower bound on the length of any path between G and G' in $\mathcal{G}(n)$.

Now root G at node j, and let $L(i)$ denote the depth of node i, that is the length of the path from node i to the root. Consider a sequence of (general flip) transformations in which for every node i with $L(i) = 2$ connected to node k with $L(k) = 1$, edge $\{i, k\}$ is transformed into edge $\{i, j\}$. This sequence of transformations ends when there are no remaining nodes at depth 2, or equivalently, when G has been transformed to G'. Moreover, each transformation in the sequence decreases the depth of node i and all its descendants by 1; equivalently, each such transformations decreases the number of nodes not adjacent to the root node j by 1. Thus, a total of m such transformations are needed for a tree with m nodes not adjacent to j, and this corresponds to the shortest path. ∎
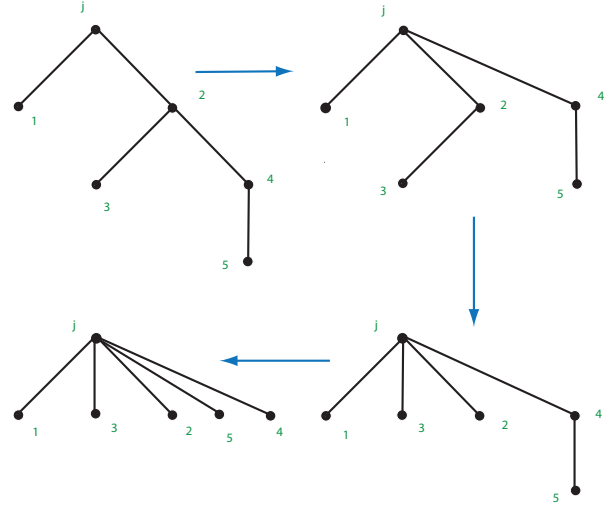

Fig. 2. **An example of a graph transformed intto a canonical graph via a minimal sequence of general flips**

Corollary 1: The length of the shortest path between any two canonical nodes in $\mathcal{G}(n)$ is $n - 2$.

Lemma 1 suggests that the n-graph of trees can be projected onto an n dimensional coordinate system in which the ith coordinate of the projection of each node is the length of the shortest path in $\mathcal{G}(n)$ from that tree to the canonical tree with star node i.

Lemma 2: Let $(x_1, x_2, ..., x_n)$ be the coordinates of a node in $\mathcal{G}(n)$ representing tree G with degree sequence $(d_1, d_2, ..., d_n)$. We have:

$$x_i = (n - 1) - d_i, \ 1 \leq i \leq n$$

Proof: Consider the tree with degree sequence $(d_1, d_2, ..., d_n)$. The tree has a total of $n - 1$ edges,

of which d_i correspond to nodes that are adjacent to node i, leaving $(n-1) - d_i$ nodes not adjacent to node i, not counting node i itself. It follows from the construction of the coordinate system and Lemma 1 that $x_i = (n-1) - d_i$. ∎

The projection of $\mathcal{G}(n)$ onto the proposed coordinate system is thus the set of integer points on the hyperplane defined by:

$$\sum_{i=1}^{n} x_i = (n-2)(n-1)$$

and satisfying $0 \leq x_i \leq n - 1$, for all i (see Figure 3). Nodes representing trees with identical degree sequences are projected onto the same point in this coordinate system.
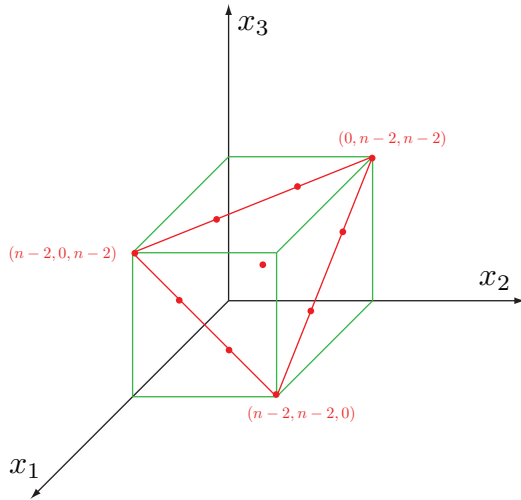

Fig. 3. **Projection of** $\mathcal{G}(n)$ *(red dots represent the projected nodes)*

In the following Theorem, we disregard the trivial case ($n = 2$) and we assume that $n \geq 3$.

Theorem 1: The diameter of $\mathcal{G}(n)$ is bounded above by $2n - 5$ and below by $n - 2$.

Proof: The lower bound follows directly from Corollary 1 since the diameter cannot be less than the length of the shortest path between any two nodes of $\mathcal{G}(n)$. The upper bound follows directly from Lemmas 1 and 2. There exists a path between any two nodes of $\mathcal{G}(n)$ going through any of the n canonical nodes. The shortest such path includes the canonical node with star node i, where i is such that $d_i \geq 2$ in the tree corresponding to either of the nodes. The distance from that node to the ith canonical node is then bounded by $n - 3$, while the distance from the other node to the same canonical node is bounded by $n - 2$. The length of the shortest path is then bounded above by $2n - 5$. ∎

The upper bound given in Theorem 1, which constitutes an improvement on the upper bound established in [7] for this special class of graph of graphs, is quite remarkable considering that the order of $\mathcal{G}(n)$ is n^{n-2} and that $\mathcal{G}(n)$ is far from being complete. Moreover, it can be shown by direct inspection of the n-graph of trees that this bound is tight for $n \leq 5$. It is not known at this point whether the

bound remains tight for larger values of n, which is one question that will be further explored in future work.

IV. FUTURE WORK

Our ultimate goal remains to understand the topology of the n-graph of trees. Of immediate interest is further refinement of the bound on the diameter: In particular, we know the bound to be tight for $n \leq 5$. Larger values of n will be explored. Of longer term interest is further study of the structure of the projected graph in the proposed coordinate system with the hope of gaining more insight into its topology.

Another direction of future work will be to explore the "n-graph of graphs" in which the nodes are identified with labeled graphs of order n and size m. Since the results presented here hinge on the fact that a tree can be rooted, it remains unclear whether they can be generalized to arbitrary graphs.

V. ACKNOWLEDGMENTS

REFERENCES

[1] R. Albert and A.-L. Barabási, Topology of Evolving Networks: Local Events and Universality, *Physical Review Letters*, vol. 85, no. 24, 2000, pp 5234- 5237.
[2] R. Albert and A.-L. Barabási, Statistical Mechanics of Complex Networks, *Reviews of Modern Physics*, vol. 74, January 2002, pp 47-97.
[3] S. N. Dorogovstev and J. F. F. Mendes, Evolution of Networks, *Advances in Physics*, vol. 51, no. 4, 2002, pp 1079-1187.
[4] P. Erdös and A. Rényi, On the Evolution of Random Graphs, *Publications* , vol. 5, 1960, pp 17-61.
[5] J.M. Hendrickx, B. Fidan, C. Yu, B.D.O. Anderson, and V.D. Blondel, Formation Reorganization by Primitive Operations on Directed Graphs, IEEE Transactions on Automatic Control, vol. 53, no. 4, 2008, pp 968-979.
[6] E. Klavins, R. Ghrist and D. Lipsky, A Grammatical Approach to Self-Organizing Robotic Systems, *IEEE Transactions on Automatic Control*, vol. 51, no. 6, 2006, pp 949-962.
[7] L. Li, *Topologies of Complex Networks: Functions and Structures*, Ph.D. Thesis, California Institute of Technology, 2007.
[8] G. Rozenberg (editor), *Handbook of Graph Grammars and Computing by Graph Transformations, Volume 1: Foundations*, World Scientific Publishing Company, Inc., River Edge, NJ; 1997.
[9] S. Skiena, *Implementing Discrete Mathematics: Combinatorics and Graph theory with Mathematica*. Addison-Wesley, Reading, MA; 1990.
[10] R. Taylor, Switchings Constrained to 2-Connectivity in Simple Graphs, *SIAM Journal on Algebraic and Discrete Methods*, vol. 3, no. 1, March 1982, pp 114-121.
[11] Y. Tsukui, Transformations of Cubic Graphs, *Journal of the Franklin Institute*, vol. 333, no. 4, July 1996, pp 565-575.
[12] Y. Tsukui, Transformations of Edge-Colored Cubic Graphs, *Discrete Mathematics*, vol. 184, no. 1-3, April 1998, pp 183-194.