FINAL REPORT

Image Representation, Clustering, and Search in Proximity Graphs and Pathfinder Networks

ONR Grant Number N00014-92-J-1509

Donald W. Dearholt

Department of
COMPUTER SCIENCE

MISSISSIPPI STATE UNIVERSITY
SERVICE
LEARNING
RESEARCH
1878

College of Engineering
MISSISSIPPI STATE UNIVERSITY

REPORT DOCUMENTATION PAGE

FORM APPROVED
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 15, 1996	Final Research Report 04/01/92 - 09/30/95

4. TITLE AND SUBTITLE OF REPORT
Image Representation, Clustering, and Search in Proximity Graphs and Pathfinder Networks

5. FUNDING NUMBERS
ONR N00014-92-J-1509

6. AUTHOR(S)
Donald W. Dearholt

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Mississippi State University
ATTN: Sponsored Programs Administration
PO Box 6156
Mississippi State, MS 39762

8. PERFORMING ORGANIZATION REPORT NUMBER:

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
ATTN: Dr. Marc Lipman
800 North Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER:

11. SUPPLEMENTARY NOTES:

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This research extended Pathfinder networks and proximity graphs to new domains, and resulted in new proximity graphs. Pathfinder networks for both attributes and concepts can now be generated from the data used to construct a concept lattice (R. Wille), which can be represented by an overlay of the networks on the lattice. Growing sphere graphs (GSGs) model energy dispersion, and can generate sphere of influence graphs (SIGs), the union of mintrees (the sparsest Pathfinder network), or more general graphs, depending upon constraints. K-local image graphs (KLIGs) store information about the neighborhood surrounding each node in an arbitrary dynamic network. KLIGs provide a mechanism for planning, so that routing under conditions of failed nodes or edges can be near optimum. All minimum-cost paths between any pair of nodes in a KLIG consist of edges of the Pathfinder network PFN(r=1, q=n-1). Pathfinder networks and other proximity graphs can now be generated dynamically by counting co-occurrences of events of interest, thus ensuring the incorporation of both clustering information and optimal paths through the graphs. A new paradigm for controlling weapons delivery vehicles (DVs) has been developed. Called *Procrustes* (it ensures that adequate resources are available), it provides a new way of viewing targets and DVs which utilizes a systematic way of substituting for failed DVs. It is a robust way of ensuring the success of such missions.

14. SUBJECT TERMS
Proximity graphs, Pathfinder networks, concept lattices, dynamic systems, energy dispersion, reliable message routing, robust weapons delivery

15. NUMBER OF PAGES:

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT:	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

FINAL REPORT

Image Representation, Clustering, and Search in Proximity Graphs and Pathfinder Networks

ONR Grant Number N00014-92-J-1509

Donald W. Dearholt

Department of Computer Science
Mississippi State University
MS 39762
(601) 325-2756

April 1, 1992 to September 30, 1995

I. RESEARCH ACCOMPLISHMENTS: OVERVIEW

This work is a continuation of research begun more than fifteen years ago. It has emphasized discovering the properties of Pathfinder associative networks and proximity graphs, extending this theory to new domains, and finding new applications. Pathfinder networks have been used in a variety of applications; a list of those who have either used them or contributed to their development is included as section VI. in this report. Among the applications already developed are the organization of data in a human-computer interface (such as the information in an online help system) and the organization and retrieval of information in a robotics vision database. This research is also relevant to modeling such apparently diverse phenomena as consensus of a group of experts, approximating the external contours of objects viewed with a video camera, a cluster-learning paradigm to emulate many aspects of human learning, and modeling aspects of dynamic systems (either discrete or sampled continuous systems). The definition of two new types of graphs has resulted in (a) the capability of modeling aspects of the propagation of energy, such as radio communications, and (b) the facilitation of local planning with respect to low-cost paths in arbitrary graphs. Both types of graphs are applicable to the theoretical properties of a robust weapons management and delivery system (called *Procrustes*) now under development, and major features of Procrustes are outlined.

Some details of results in each area studied under this grant are summarized in the paragraphs following; for convenience, theoretical results and applications are discussed separately.

A. Concept Lattices

1. Theoretical: Associations have been established between the concept lattices developed by R. Wille and Pathfinder networks. In particular, Pathfinder networks for both the set of attributes and the set of concepts can be generated from the data used to construct a con-

cept lattice, which can be represented by an overlay of the Pathfinder networks on the lattice.

2. Applications

This complex data structure possesses all the advantages of both the lattice (complete representation of the data) and of the Pathfinder networks (clustering and structural information as well as low-cost paths). Consequently, it is useful in special database applications requiring information retrieval based upon either specific keys (concepts or attributes) or upon clustering and higher-level abstractions utilizing the clustering.

B. Growing Sphere Graphs

1. Theoretical: Growing sphere graphs (GSGs) are designed to model dynamic aspects of energy dispersion, and can generate either sphere of influence graphs (SIGs) or the union of minimum-cost spanning trees (which is also the sparsest Pathfinder network), depending upon the constraints applied. In addition, GSGs can generate many other graphs, more general than either SIGs or Pathfinder networks.

2. Applications

a. GSGs can model aspects of radio communications or wireless networks associated with computer data communications. Once energy levels and receiver and antenna characteristics are known, the graphs can be designed to indicate which transmitters and receivers can communicate either one-way, or both ways. GSGs support modeling dynamic scenarios in which transmitters are turned on or off, or transmitter power is changed.

b. Under simple assumptions, GSGs can be used to describe the kill radii and kill volumes of individual targets in battlefield scenarios. Furthermore, the clustering properties and hull of such a GSG provides information of use in planning the strike force designed to neutralize the cluster(s) of targets in the Procrustes system, described in E. below.

C. K-Local Image Graphs

1. Theoretical: Defined for an arbitrary dynamic network, k-local image graphs (KLIGs) store information about the neighborhood surrounding each node. Every node in a KLIG is assumed to have sufficient memory and information to store the graph structure surrounding that node, to include each node within k edges, and all interconnecting edges for that set of nodes. KLIGs provide a mechanism for planning associated with each node, so that planning can precede action. It has been shown that all minimum-cost paths between any pair of nodes in a KLIG (based upon the edge costs at that time) consist of edges of the Pathfinder network PFN(r=1, q=n-1), where n is the number of nodes in the KLIG.

2. Applications

a. KLIGs are being investigated to plan paths dynamically through an arbitrary communications network which is composed of nodes and links which can fail. It is known that Pathfinder networks computed on the arbitrary network provide an abundance of information on alternative, low-cost paths through the network. The development and testing of suitable algorithms for exploiting this information efficiently is needed to apply the theory to real networks. A demonstration/simulation is being designed now to serve as a testbed for algorithms utilizing this paradigm.

b. KLIGs will be useful for distributing the control of delivery vehicles (DVs) in the Procrustes weapons management/delivery system, described below. For the control of large numbers of DVs assigned to a large number of targets, the division of targets and DVs into clusters and the use of KLIGs to maintain relatively local control of the DVs seems a distinct advantage, capable of minimizing system vulnerability and maximizing effectiveness.

D. Modeling Dynamic Systems

1. Theoretical: The theory of Pathfinder networks and other proximity graphs has been further extended into dynamic systems by the development of a constructive approach to generating such graphs, based upon monitoring and counting co-occurrences of events of interest. Pathfinder associative networks constructed on the resulting dissimilarity matrix provide a dynamic model of some aspects of learning, and a fixed memory capacity is usually assumed. Such perspectives on dynamic systems incorporate both clustering information and paths through the graphs, so that algorithms utilizing both abstractions and navigation are supported.

2. Applications

Since co-occurrence of "events of interest" can be construed to be a variety of things, there is, correspondingly, a wide variety of potential applications. Two possibilities, selected for their differences, are described.

a. In natural language text, some words or groups of words can be selected as key terms. Whenever pairs of these co-occur within some syntactic boundary, such as a clause, a co-occurrence counter can be incremented by one. In this manner, a co-occurrence matrix for the set of key terms can be constructed. This matrix is inverted by subtracting each entry from a constant larger than any co-occurrence count in the matrix, and the resulting dissimilarity matrix is used to generate a Pathfinder network; there may be a window of text length limited to a certain number of clauses to correspond to memory limitations. Such networks possess the clustering properties associated with Pathfinder networks, in which key terms co-occurring frequently are in the same cluster. Bridges to other clusters are formed by those pairs which co-occur less frequently.

b. In combat, the positions of aircraft (or other weapons systems) can be monitored, and a dissimilarity matrix can be computed from the euclidean distance between each pair of aircraft. Co-occurrences are defined in terms of lunes; the lunes, in turn, are defined in terms of the locations of pairs of the aircraft. Either a Pathfinder network or a sphere-of-influence graph can be constructed from the resulting information at appropriate intervals of time, providing another perspective on the motion of the aircraft, and of their vulnerability *vis-a-vis* the clustering that results.

E. The Procrustes Weapons-Management and Delivery System

1. Theoretical: A long-range project based upon theoretical work resulting from this research is the design of a system for controlling DVs assigned to neutralize targets. Called *Procrustes* (because it ensures that adequate resources are available, when possible, to neutralize prioritized targets), it provides a new way of viewing targets and DVs which makes use of aspects of modeling dynamic systems using co-occurrence, consensus, and an extension of Pathfinder networks, reported two years ago. When combined with the progress in GSGs, KLIGs, and modeling dynamic systems, it provides a robust way of assuring the success of such missions. Work on refining the mathematics needed to assure that

all parts of the paradigm are theoretically sound has begun. A properly designed simulation or development for deployment will require substantial effort, and a draft of the steps believed to be needed was outlined in a prior document.

2. Applications

a. The principles of Procrustes can be incorporated in a virtual environment for training and testing of present battle protocols against such a system. This would make evident problems in our present defensive strategies, if such problems exist. This would also support testing of different algorithms utilizing the principles of Procrustes against each other, to determine which is most robust under a variety of circumstances.

b. The most significant (long range) application is the incorporation of the Procrustes paradigm into the battle protocols of our armed forces, once design and testing are complete. This requires substantial development and testing, but has truly significant benefits; it assures that highest priority targets are nearly certain to be destroyed, and that coverage of all targets remains essentially complete, using the substitution strategy of Procrustes whenever any DV fails.

The results of the research sponsored by this grant have been communicated in a variety of channels, summarized below.

II. ORGANIZATION OF CONFERENCE SESSIONS AND WORKSHOPS

Pathfinder Networks and Proximity Graphs, Invited Session at the *Classification Society of North America* meeting. In progress for the annual meeting. Amherst, Massachusetts, June 14-15, 1996.

Pathfinder Networks, Invited Session at the *Classification Society of North America* meeting. Organized a session on Pathfinder networks, and obtained three other speakers for the session. Pittsburgh, Pennsylvania, June 25-26, 1993.

Organized and hosted *The Second Workshop on Proximity Graphs* (Marc Lipman was co-organizer). Sponsored by the ONR, the Department of Computer Science at Mississippi State University, the NSF Engineering Research Center at MSU, and the Dow Chemical Company, held at Mississippi State University, February 10-12, 1993.

Organized and hosted *The Third Workshop on Proximity Graphs* (Marc Lipman was co-organizer). Sponsored by the ONR, the Department of Computer Science at Mississippi State University, the NSF Engineering Research Center at MSU, and The Exxon Corporation, held at Mississippi State University, December 1 - 3, 1994.

III. TECHNICAL REPORTS

Dearholt, D. W., and M. Lipman, eds. (1996). "Proceedings of the Third Workshop on Proximity Graphs". Technical report in progress. Department of Computer Science, Mississippi State University, Mississippi 39762.

Dearholt, D. W., and M. Lipman, eds. (1993). "Proceedings of the Second Workshop on Proximity Graphs". *Technical Report Number MSU-930210.* Department of Computer

Science, Mississippi State University, Mississippi 39762, September, 1993.

IV. CONFERENCE AND WORKSHOP PRESENTATIONS

Modeling Dynamic Systems Using Co-occurrence and Associative Networks, presented at the Tenth International Conference on Mathematical and Computer Modelling and Scientific Computing, Boston, Massachusetts, July 5, 1995 (to be published in the conference proceedings).

Growing-Sphere Graphs as a Generalization of Sphere of Influence Graphs and the Union of Mintrees, presented at the annual meeting of the Classification Society of North America, Denver, Colorado, June 11, 1995.

K-Local Image Graphs and Cost-Effective Paths in Dynamic Networks, presented at the Third Workshop on Proximity Graphs, Mississippi State University, December 3, 1994.

Growing-Sphere Graphs as a Generalization of Sphere of Influence Graphs, presented at the Third Workshop on Proximity Graphs, Mississippi State University, December 2, 1994.

An Introduction to Pathfinder Networks, Proximity Graphs, and Applications, plenary address at the Second International Conference on Ordinal Data Analysis, University of Massachusetts, October 16, 1993.

Modeling Dynamic Phenomena Using Co-occurrence, Clustering, and Pathfinder Networks, presented at the annual meeting of the *Classification Society of North America*, Pittsburgh, Pennsylvania, June 26, 1993.

Introduction to Pathfinder Networks and Proximity Graphs, presented at *The Second Workshop on Proximity Graphs*, Mississippi State University, February 11, 1993.

Dynamic Clustering, Co-occurrence, and Proximity Graphs, presented at *The Second Workshop on Proximity Graphs*, Mississippi State University, February 11, 1993.

Clustering, Image Representation, and Search in Proximity Graphs and Pathfinder Networks, invited presentation at the annual meeting of the Classification Society of North America, Michigan State University, June 13, 1992.

V. COLLOQUIA

An Introduction to Pathfinder Networks, Proximity Graphs, and Applications, Northern Arizona University, Flagstaff, Arizona, May 3, 1995.

Modeling Dynamic Phenomena Using Co-occurrence, Clustering, and Pathfinder Networks, Stennis Space Center, Mississippi, September 13, 1993.

Introduction to Pathfinder Networks and Their Uses in Clustering and Modeling Learning, Los Alamos National Laboratory, New Mexico, June 1, 1993.

Properties and Applications of Pathfinder-Based Associative Networks, the Upjohn Pharmaceutical Company, Kalamazoo, Michigan, June 10, 1992.

VI. INTERACTIONS INVOLVING RESEARCH RESULTS

The researchers listed below are those who are known to have used the Pathfinder paradigm or contributed to the development of the theory or applications. This development began in 1980, while the principal investigator (Don Dearholt) was at New Mexico State University. Frank Durso, Roger Schvaneveldt, and Don Dearholt were the collaborators who established the Pathfinder paradigm, and Ken Paap and Jim McDonald joined us soon after. As the only computer scientist in the group, Dearholt developed most of the theoretical properties of Pathfinder networks and significant parts of the generation algorithm. Several researchers listed below contributed chapters to a book, *Pathfinder Associative Networks: Studies in Knowledge Acquisition*, edited by Roger Schvaneveldt, and published by Ablex in 1990. Others attended and contributed to one of the three workshops on proximity graphs, held at New Mexico State University in 1989, and at Mississippi State University in 1993 and 1994. The first workshop was sponsored by the National Science Foundation and the Office of Naval Research, and resulted in the Proceedings of the First Workshop on Proximity Graphs, edited by Don Dearholt and Frank Harary. Subsequent workshops have been sponsored by the ONR, the NSF Engineering Research Center at Mississippi State University, the Dow Chemical Company, and The Exxon Corporation; the proceedings of the latter two workshops have been edited by Don Dearholt and Marc Lipman.

Research Unit:	**Researcher(s):**
Bell Laboratories Murray Hill, NJ	Richard Becker
The Boeing Company:	
Boeing Computer Services	Timothy J. Breen
Boeing Advanced Technology Center	Chris Esposito
Cambridge University MRC Applied Psychology Unit, England	John G. Gammack
Duke University Department of Psychology	David C. Rubin
Hartwick College Department of Psychology	Lisa A. Onorato
IBM:	
Santa Theresa Labs	Matt Anderson
Thomas J. Watson Research Center	Wendy A. Kellogg
Westlake	Cathy Cavendish, David Ehnebuske
Montana State University Department of Computer Science	Robert Cimikowski

NASA Langley, Virginia	Renate Roske-Hofstrand
New Mexico State University Computing Research Laboratory and Department of Psychology	Roger Schvaneveldt, Ken Paap, Deborah McDonald, Jim McDonald
Pan American University Department of Computer Science	Richard Fowler, Wendy Lawrence
Rice University Department of Psychology	Nancy J. Cooke
Sandia National Laboratories Albuquerque, NM	Daniel Davenport
University of Georgia Department of Educational Psychology	Paula Schwanenflugel
University of New Mexico Department of Psychology	Timothy Goldsmith, Peder Johnson
University of Oklahoma Department of Psychology	Cathy Coggins, Tom Dayton, Francis J. Durso, Jack D. Shepard
University of Pittsburgh	Stephen Hirtle
National University of Singapore Department of Information Systems and Computer Science	K. T. Lua
University of Toronto Department of Computer Science	Tony A. Plate
Xerox Corporation Webster, NY	Govinda Kurup

VII. SUMMARY

The foundations of the projects outlined in section I. above have been developed under support from the ONR Discrete Mathematics Program over the past few years, and this support is gratefully acknowledged. This research has been significant in enabling the cutting-edge projects described, which represent new approaches to addressing specific, predetermined applications of probable importance to the Navy.