**Final Report for Period:** 09/2009 - 08/2010
**Submitted on:** 11/30/2010
**Principal Investigator:** Huo, Xiaoming .
**Award ID:** 0700152
**Organization:** GA Tech Res Corp - GIT
**Submitted By:**
Huo, Xiaoming - Principal Investigator
**Title:**
Fundamentals and Applications of Connect-the-Dots Methods

## Project Participants

**Senior Personnel**

**Name:** Huo, Xiaoming
**Worked for more than 160 Hours:** Yes
**Contribution to Project:**

**Name:** Tovey, Craig
**Worked for more than 160 Hours:** Yes
**Contribution to Project:**

**Post-doc**

**Graduate Student**

**Name:** Chen, Jie
**Worked for more than 160 Hours:** No
**Contribution to Project:**
GRA for one semester from this project.

**Name:** Smith, Andrew
**Worked for more than 160 Hours:** No
**Contribution to Project:**
GRA for one semester in Spring 2008.

**Name:** Kim, Heeyoung
**Worked for more than 160 Hours:** Yes
**Contribution to Project:**

**Name:** Shepardson, Dylan
**Worked for more than 160 Hours:** Yes
**Contribution to Project:**

**Undergraduate Student**

**Technician, Programmer**

**Other Participant**

**Research Experience for Undergraduates**

## Organizational Partners

## Other Collaborators or Contacts

## Activities and Findings

**Research and Education Activities:**

We studied the dynamic programming approach for the connect-the-dots (CTD) problems. Numerical experiments were carried out. Some research findings have been written in a research report that is released online. Four graduate students have been involved in various stage of this project. Since the beginning of this project, nine published journal papers and five submitted journal papers are produced. Not all of them are on the CTD problems. However all involves dynamic programming and optimization. All of these papers are coauthored with PhD students. Research activities also include releasing a software package online.

**Findings:**

On the CTD problems, we use the multiscale idea to develop some fast algorithms for the CTD problems. Some proposed algorithms have the best order of complexity that we know of. Efficient implementation in computer has been done. The technical report (TR) that was released online includes most of the results we have obtained so far. The URL is provided. We summarize our findings below.

1. We describe the dynamic programming as a generic approach to solve various connect-the-dots problems. These problems may involve functional classes such as increasing functions, unimodal functions, Holder functions, convex functions, and so on. We point out some connections between the CTD problems associated with different functional classes. For example, we found that the unimodal functional class is as hard to process as the increasing functional class, which consequently is as hard to process as the Lipschitz functional class.
2. We further demonstrate that the direct application of the dynamic programming is not very efficient. Instead, one can utilized the multiscale (AKA divide and conquer) approach to reduce the required computational complexity significantly. For example, a direct application of dynamic programming in the increasing function gives an order O(n^2) algorithm. The multiscale strategy can reduce the complexity to O(n log n).
3. A new dynamic programming algorithm for the Holder-2 graphs is developed: the algorithm is nontrivial, and to our knowledge is in its first appearance. Our computational experiments indicate that the algorithm is very efficient and potentially has low order of average complexity. We have not been able to prove the theoretical complexity for that algorithm. This is a work ongoing.
4. The problem of directional data, which is named the connect-the-darts problem, is explicitly solved for the first time in the aforementioned TR. Several interesting conjectures are drawn, motivated by simulations. This problem has interesting connection with neural science. We discussed the connection in the TR.
5. We also use the above numerical algorithms to study the possible theoretical/analytical distributions on the solution. For

example, we studied the empirical number of the dots linked by an increasing function within a uniformly distributed point cloud. The asymptotic distribution in this case is known. Our simulations confirm these distributional analyses. We made some new conjectures corresponding to our own simulation studies.
This project has spawned the following outputs. They are not exactly the aforementioned CTD problems, however the methodology that applied shares many commonalities.
We have studied the relation between the computational problem in CTD and the detectability problem. Some findings are reported in a paper published in Statistic Sinica---a top 10 journal in statistics.
We have studied how to utilize the multiscale approaches to expedite the dynamic programming algorithms in solving the shortest path problems. In theoretical analysis, we showed that the multiscale strategy can reduce the computational order of complexity. In simulations, we observed that the improvement is even better than prediction based on the theoretical analysis. Two submitted papers are under revision.
We have also studied various statistical and computational problems, which either utilized dynamic programming, or optimization. It has resulted in several papers published or under revision.

**Training and Development:**

Students gain research experience by working with us and writing papers together.

**Outreach Activities:**

## Journal Publications

Huo, XM; Ni, XL, "When do stepwise algorithms meet subset selection criteria?", ANNALS OF STATISTICS, p. 870, vol. 35, (2007). Published, 10.1214/009053606000000133

Ni, XL; Huo, XM, "Statistical interpretation of the importance of phase information in signal and image reconstruction", STATISTICS & PROBABILITY LETTERS, p. 447, vol. 77, (2007). Published, 10.1016/j.spl.2006.08.02

Chen, J; Deng, SJ; Huo, XM, "Electricity price curve modeling and forecasting by manifold learning", IEEE TRANSACTIONS ON POWER SYSTEMS, p. 877, vol. 23, (2008). Published, 10.1109/TPWRS.2008.92609

Ni, XS; Huo, XM, "Another look at Huber's estimator: A new minimax estimator in regression with stochastically bounded noise", JOURNAL OF STATISTICAL PLANNING AND INFERENCE, p. 503, vol. 139, (2009). Published, 10.1016/j.jspi.2008.03.04

Huo, XM; Smith, AK, "Matrix perturbation analysis of local tangent space alignment", LINEAR ALGEBRA AND ITS APPLICATIONS, p. 732, vol. 430, (2009). Published, 10.1016/j.laa.2008.09.01

Huo, XM; Ni, XL, "DETECTABILITY OF CONVEX-SHAPED OBJECTS IN DIGITAL IMAGES, ITS FUNDAMENTAL LIMIT AND MULTISCALE ANALYSIS", STATISTICA SINICA, p. 1439, vol. 19, (2009). Published,

Huo, XM; Chen, J, "Complexity of penalized likelihood estimation", JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION, p. 747, vol. 80, (2010). Published, 10.1080/00949650902773354

J. Chen and X. Huo, "A Hessian regularized nonlinear time series model.", Journal of Computational and Graphical Statistics,, p. 694, vol. 18(3), (2009). Published,

S. B. Kim , X. Huo , and K.-L. Tsui, "A finite-sample simulation study of cross validation in tree-based models", Information Technology and Management, p. 223, vol. 10(4), (2009). Published,

X. Huo, H.-Y. Kim, and J. Shi, "An interval bounded classifier for functional data", Annals of Operations Research, p. , vol. , (2010). Submitted,

X. Huo, H. Kim, M. Shilling, and H. D. Tran, "A Lipschitz regularity based statistical model, with applications in coordinate metrology: confidence bands and optimal sampling strategy", Technometrics, p. , vol. , (2010). Submitted,

X. Huo and H. Kim, "Spatially adaptive and asymptotically optimal smoothing splines", Biometrika, p. , vol. , (2010). Submitted,

Y. Lu, X. Huo, and P. Tsiotras, "Beamlet-based graph structure for path planning using multiscale information", IEEE Transactions on Automatic Control, p. , vol. , (2010). Submitted,

Y. Lu, X. Huo, O. Arslan, and P. Tsiotras, "A incremental, multi-scale search algorithm for dynamic path planning with low worst case complexity", IEEE Transactions on Systems, Man and Cybernetics, Part B, p. , vol. , (2010). Submitted,

## Books or Other One-time Publications

A. K. Smith, X. Huo, and H. Zha
, "Convergence and rate of
convergence of a manifold-
based dimension reduction
", (2008). Conference proceedings, Published
Bibliography: NIPS

J. Feng, L. Song, X. Huo, X.
Yang, and W. Zhang
, "Image denoising using local
tangent space alignment
", (2010). Conference proceedings, Published
Bibliography: Visual Communications and Image Processing (VCIP)

Y. Lu, X. Huo, and P. Tsiotras, "Beamlet-like data processing
for accelerated path-planning
using multiscale information
of the environment
", (2010). Conference proceedings, Published
Bibliography: 49th IEEE Conference on Decision and Control

## Web/Internet Site

**URL(s):**
http://www2.isye.gatech.edu/statistics/papers/07-07.pdf
**Description:**
Technical report online:
Huo, Xiaoming, Tovey, Craig, Donoho, David L. and Arias-Castro, Ery DYNAMIC PROGRAMMING METHODS FOR CONNECT THE DOTS" IN SCATTERED POINT CLOUDS.

This paper reports recent works on the algorithms that are related to CTD.

## Other Specific Products

**Product Type:**

**Software (or netware)**

**Product Description:**

We developed a comprehensive software package that is related to the CTD problems. We have released it online.

**Sharing Information:**

A version has been released online at http://www2.isye.gatech.edu/~xiaoming/CTDLab/.

## Contributions

**Contributions within Discipline:**

Many of those dynamic programming algorithms that are developed in the online technical report are new. The insights in extensive simulations may lead to new discovery in theory. The resulting publications contribute to various problems that involve optimization and dynamic programming.

**Contributions to Other Disciplines:**

The CTD problems have connections with probability theory, geometric discrepancy theory, filament detection, vision research, batched disk scheduling, and airplane boarding, as described in our proposal. Our results have potential impacts there.

**Contributions to Human Resource Development:**

Four PhD students were supported in various stages of their graduate study. They have gained experience via coauthoring papers with the PIs.

**Contributions to Resources for Research and Education:**

The online software will be useful to others who are interested in these problems that we have studied.

**Contributions Beyond Science and Engineering:**

## Conference Proceedings

## Categories for which nothing is reported:

Organizational Partners

Activities and Findings: Any Outreach Activities

Contributions: To Any Beyond Science and Engineering

Any Conference