# MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRL-RO-BI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

☐ Reprint (Orig + 2 copies)

☐ Technical Report (Orig + 2 copies)

☐ Manuscript (1 copy)

☒ Final Progress Report (Orig + 2 copies)

☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: DAAD 19-99-C-0031

REPORT TITLE: Final Progress Report on Robust and/or Adaptive Filtering by Neural Networks

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

James T. Lo
Principal Investigator

# REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188



| 1. AGENCY USE ONLY ( Leave Blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | September 30, 2002 | July 01, 1999 – June 30, 2002 |

4. TITLE AND SUBTITLE

Final Progress Report on Robust and/or Adaptive Filtering by Neural Networks

5. FUNDING NUMBERS

DAAD 19-99-C-0031

6. AUTHOR(S)

James T. Lo

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Maryland Technology Corporation
10210 Sunway Terrace, Ellicott City, MD 21042

8. PERFORMING ORGANIZATION REPORT NUMBER

TARO-02B

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The purpose of this project is to develop a general and systematic approach to robust and/or adaptive filtering in the presence of uncertain environmental parameters. Mathematical justification, intuitive understanding and numerical confirmation of risk-averting neural networks for general robust processing with various degrees of robustness have been achieved. Those of neural networks with long- and short-term memories for general adaptive processing have also been accomplished. A novel method of training neural networks that is effective in avoiding poor local minima has been discovered. This discovery is a major breakthrough in the development of neural computing. Robust neural filters have been mathematically justified and numerically tested. General adaptive filtering and general robust adaptive filtering turned out to be much more difficult than expected. Nevertheless, schemes for them by neural computing, which are mathematically natural and convincing, have finally been conceived

14. SUBJECT TERMS

Robust filter, adaptive filter, robust adaptive filter, synthetic approach, neural computing, neural network, recurrent neural network, risk-averting criterion, risk-averting training, robust identification, adaptive identification, series-parallel formulation, parallel formulation, degree of robustness, local minimum, dynamical system, signal, measurement, estimation

15. NUMBER OF PAGES

9

16. PRICE CODE

| 17. SECURITY CLASSIFICATION OR REPORT | 18. SECURITY CLASSIFICATION ON THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| **UNCLASSIFIED** | **UNCLASSIFIED** | **UNCLASSIFIED** | **UL** |

## GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used for announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements.***

**Block 1.** Agency Use Only (Leave blank)

**Block 2.** Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least year.

**Block 3.** Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

**Block 4.** Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, and volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

**Block 5.** Funding Numbers. To include contract and grant numbers; may include program element number(s) project number(s), task number(s), and work unit number(s). Use the following labels:

| | |
|---|---|
| **C** - Contract | **PR** - Project |
| **G** - Grant | **TA** - Task |
| **PE** - Program Element | **WU** - Work Unit Accession No. |

**Block 6.** Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

**Block 7.** Performing Organization Name(s) and Address(es). Self-explanatory.

**Block 8.** Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

**Block 9.** Sponsoring/Monitoring Agency Name(s) and Address(es) Self-explanatory.

**Block 10.** Sponsoring/Monitoring Agency Report Number. (*if known)*

**Block 11.** Supplementary Notes. Enter information not included elsewhere such as; prepared in cooperation with....; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

**Block 12a.** Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NORFORN, REL, ITAR).

**DOD** - See DoDD 4230.25, "Distribution Statements on Technical Documents."
**DOE** - See authorities.
**NASA** - See Handbook NHB 2200.2.
**NTIS** - Leave blank.

**Block 12b.** Distribution Code.

**DOD** - Leave Blank
**DOE** - Enter DOE distribution categories from the Standard Distribution for unclassified Scientific and Technical Reports
**NASA** - Leave Blank.
**NTIS** - Leave Blank.

**Block 13.** Abstract. Include a brief (*Maximum 200 words)* factual summary of the most significant information contained in the report.

**Block 14.** Subject Terms. Keywords or phrases identifying major subject in the report.

**Block 15.** Number of Pages. Enter the total number of pages.

**Block 16.** Price Code. Enter appropriate price code (NTIS *only*).

**Block 17. - 19.** Security Classifications. Self-explanatory. Enter U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

**Block 20.** Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (Unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

# Final Progress Report

on

## Robust and/or Adaptive Filtering by Neural Networks

**submitted to**

**Army Research Office**

**Contract DAAD19-99-C-0031**

**07/01/99 - 06/30/02**

## 1. Foreword

The project has been a great success:

- A major breakthrough in neural network training has been achieved. The novel risk-averting training method is effective in avoiding poor local minima of the training criterion.
- Mathematical justification and numerical confirmation of neural networks with long- and short-term memories for general adaptive processing have been accomplished.
- Mathematical justification, intuitive understanding and numerical confirmation of risk-averting neural network for general robust processing with various degrees of robustness have been obtained.
- Robust neural filters have been mathematically justified and numerically tested.
- General adaptive filtering and general robust adaptive filtering turned out to be much more difficult than expected. Nevertheless, schemes for them by neural computing, which are mathematically natural and convincing, have finally been conceived.

## 2. Statement of the problem studied

If the signal or measurement process involved in a filtering problem has an uncertain environmental parameter that is either observable or unobservable but constant long enough for adaptation, an adaptive filter is required. If the environmental parameter is unobservable and changes too fast for adaptation, then the filter needs to be robust toward the environmental parameter to avoid excessively large or disastrous filtering errors. The main objective of the project is to develop systematic and general methods of designing practical robust and/or adaptive filters that have virtually optimal online performances. Such filters are much needed in many applications of great practical importance.

Novel neural networks with long- and short-term memories and novel robust neural networks were proposed as building blocks of the adaptive and/or robust filters respectively. Fundamental issues on training these neural networks and their effectiveness to approximate dynamical systems have been studied in detail during the contract period and breakthrough results have been obtained. Based on these results, schemes of robust and/or adaptive filtering have been conceived and under mathematical analysis and numerical testing.

## 3. Summary of the most important results

Results, that support the conclusions stated in the Foreword above, are summarized in the following for each publication in which they appear.

- Mathematical Justification of Risk-Sensitive Neural Filtering, *Proceedings of the 2000 Conference on Information Sciences and Systems*, pp. WA1-7-WA1-11, March 2000, Princeton, New Jersey.

For general signal and measurement processes under mild regularity conditions, the optimal filtering performance with respect to a general risk-sensitive criterion can be approximated to any accuracy by a recurrent neural network trained with a risk-averting criterion. This provides a mathematical justification of robust neural filtering.

- An Adaptive Method of Training Multilayer Perceptrons, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 2013-2018, Washington, D.C.

This paper proposes a method of training MLPs that selects a training criterion most suitable for the function to be approximated; the measurement noises and the sampling distribution in training data, and that, perhaps most important, avoids poor local minima.

Numerical examples demonstrate that the proposed training method is able to include fine features and under-represented segments of the function being approximated in the multilayer perceptron under training. Most important perhaps, the method has good ability to avoid getting trapped in a poor local minimum.

- Training Multilayer Perceptrons in the Presence of Measurement Outliers, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 2030-2035, Washington, D.C.

Outlying measurement noises in the training data may distort an MLP trained with the ordinary quadratic criterion on the data. Much research has been done to reduce such effects of the outlying measurement noise. Robust estimation criteria such as logistic, Huber's, Tukey's biweight and Talwar's criteria from statistics have been used for training neural networks. However, there are serious difficulties in the selection of suitable initial MLP weights and the selection of the scale estimator value.

The purpose of this paper is to propose an alternative training method without such difficulties. Numerical examples show that when there are outlying measurement noises in the training data and the function under approximation is reasonably smooth and does not have an under-represented segment, the proposed adaptive method of training MLPs in the presence of outlying measurement noises works effectively.

- Adaptive versus Accommodative Neural Networks for Adaptive System Identification, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 1279-1284, Washington, D.C.

It has been proven that MLPs and RMLPs (recurrent MLPs) with LASTMs are universal series-parallel and parallel identifiers of dynamical systems with an environmental parameter, under mild regularity conditions. This paper demonstrates the numerical feasibility of these mathematically proven results with numerical examples. In the same examples, the accommodative neural networks are also obtained that have the same accuracy as do the networks with LASTMs on the same training data. Generalization abilities of the two types of network are then compared. The comparison results show that networks with LASTMs are superior to the accommodative networks with respect to these two criteria.

- Robust Identification of Dynamical Systems by Neurocomputing, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 1285-1290, Washington, D.C.

If a dynamical system has a fine feature or an operating condition under-represented in the input/output data used for its identification, the ordinary quadratic identification criterion often leads to very large or disastrous identification errors. H-infinity, risk-sensitive and minimax criteria have been used to obtain robust identifiers mostly for linear systems. A neurocomputing approach to robust identification with respect to general risk-averting criteria was proposed and mathematically justified.

This paper studies the numerical feasibility of this approach, and provides a method of training neural networks into robust identifiers of dynamical systems. The new training method is tested on both the series-parallel and parallel identifications of the modified benchmark systems, using input/output data sets with and without noises. The performances of the resultant robust neural identifiers are compared with those of the least-squares neural identifiers trained with the ordinary quadratic criterion. The robust identifiers consistently outperform the least-squares identifiers in the examples.

- Virtually Convex Criteria for Training Neural Networks, *Proceedings of the 2001 Conference on Artificial Neural Networks in Engineering*, Nov. 4-7, 2001, St. Louis, Missouri.

This paper shows that the risk-averting error criterion is "virtually convex:" As the risk-sensitivity index of this criterion increases, the region on which the criterion is convex increases monotonically to the entire weight vector space except the intersection of a finite number of manifolds determined by the training data. As the convexity region expands, "worm holes" are created so that a local search optimization procedure can travel through them to a lower local minimum. It is also proved that the risk-averting error criterion approaches a minimax criterion as the risk-sensitivity index increases.

- Risk-Averting Criteria for Training Neural Networks, *Proceedings of the Eighth International Conference on Neural Information Processing*, pp. 476-481, Nov. 14-18, 2001, Shanghai, China.

Essentially the same as the above item.

- Minimization through Convexitization in Training Neural Networks, *Proceedings of the 2002 International Joint Conference on Neural Networks*, pp. 1558-1563, Honolulu, Hawaii, May 2002.

Essentially the same as the above item.

- Adaptive Multilayer Perceptrons with Long- and Short-Term Memories, *IEEE Transactions on Neural Networks*, vol. 13, no. 1, pp. 22-33, January 2001.

MLPs with LASTMs (long-and short-term memories) are proposed for adaptive processing. The activation functions of the output neurons of such a network are linear, and thus the weights in the last layer affect the outputs of the network linearly and are called linear weights. These linear weights constitute the short-term memory and other weights the long-term memory.

It is proven that virtually any function with an input variable and an environmental parameter can be approximated to any accuracy by an MLP with LASTMs whose long-term memory is independent of the environmental parameter. This independency of the environmental parameter allows the long-term memory to be determined in an a priori training and allows the online adjustment of only the short-term memory for adapting to the environmental parameter. The benefits of using an MLP with LASTMs include less online computation, no poor local extrema to fall into, and much more timely and better adaptation. Numerical examples illustrate that these benefits are realized satisfactorily.

- Robust Approximation of Uncertain Functions where Adaptation is Impossible, *Proceedings of the 2002 International Joint Conference on Neural Networks*, pp. 1889-1894, Honolulu, Hawaii, May 2002 (with Devasis Bassu).

This paper is concerned with robust approximation of functions with an environmental parameter that changes so fast that adaptation to it is impossible. Approximation with respect to the ordinary least-squares criterion provides a good overall approximation but at the cost of large approximation errors for some values of the independent variables. An alternative training method using the risk-averting training criterion is proposed that provides robust function approximation. The method adaptively adjusts the sensitivity index of the risk-averting criterion to tune to the effects of the uncertain environmental parameter, when the measurement noises are negligible or unbiased. Numerical examples are presented illustrating the efficacy of the proposed adaptive risk-averting training method for producing function approximates with different degrees of robustness.

- Robust Identification of Uncertain Dynamical Systems where Adaptation is Impossible, *Proceedings of the 2002 International Joint Conference on Neural Networks*, pp. 1956-1961, Honolulu, Hawaii, May 2002 (with Devasis Bassu).

Depending on the applications, different degrees of robustness are required for system identification in the presence of an environmental parameter that is unobservable and changes so fast that adaptation is impossible. H-infinity and minimax criteria are too pessimistic for most applications. This paper proposes the risk-averting error criterion and shows that training with respect to it yields robust system identifiers with various degrees of robustness. Numerical results illustrate the efficacy of the proposed method and the effects of different degrees of robustness.

- Existence and Uniqueness of Risk-Sensitive Estimates, scheduled to appear in *IEEE Transactions on Automatic Control*, November 2002 (with Thomas Wanner).

Existence and uniqueness of conditional expectations and thus minimum-variance estimates are guaranteed by the Radon-Nikodym theorem in measure theory. Existence and uniqueness issues of robust estimates with respect to the risk-sensitive error with various risk-sensitivity indices are addressed in this paper. The existence of a unique risk-sensitive estimate is proven provided that the sensitivity index is positive and the power of the absolute deviations involved is greater than 1. For the remaining cases, a general existence result is not available. We do however prove the existence in certain special cases. Moreover, we present examples with uncountably many optimal risk-sensitive estimates, i.e., exhibiting an extremely high level of nonuniqueness.

- Recurrent Multilayer Perceptrons for Discrete-Time Dynamic System Identification, under revision.

The ability of two types of recurrent MLP (multilayer perceptron), namely the MLP with interconnected neurons (MLPWIN) and the MLP with output feedbacks (MLPWOF), for approximating discrete-time dynamic systems is studied in the context of system identification. They are both proven to approximate smooth dynamic system to any accuracy in the series-parallel identification formulation and, for a finite time period and under some more regularity conditions on the system, in the parallel formulation as well.

## 3. List of publications and reports

(a) Papers published in peer-reviewed journals

1. Adaptive Multilayer Perceptrons with Long- and Short-Term Memories, *IEEE Transactions on Neural Networks*, vol. 13, no. 1, pp. 22-33, January 2002 (with Devasis Bassu).
2. Existence and Uniqueness of Risk-Sensitive Estimates, scheduled to appear in *IEEE Transactions on Automatic Control*, November 2002 (with Thomas Wanner).

(b) Papers published in peer-reviewed conference proceedings

1. Mathematical Justification of Risk-Sensitive Neural Filtering, *Proceedings of the 2000 Conference on Information Sciences and Systems*, pp. WA1-7-WA1-11, March 2000, Princeton, New Jersey.
2. An Adaptive Method of Training Multilayer Perceptrons, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 2013-2018, Washington, D.C. (with Devasis Bassu).
3. Training Multilayer Perceptrons in the Presence of Measurement Outliers, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 2030-2035, Washington, D.C. (with Devasis Bassu).
4. Adaptive versus Accommodative Neural Networks for Adaptive System Identification, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 1279-1284, Washington, D.C. (with Devasis Bassu).

5. Robust Identification of Dynamical Systems by neurocomputing, *Proceedings of the 2001 International Joint Conference on Neural Networks*, pp. 1285-1290, Washington, D.C. (with Devasis Bassu).
6. Virtually Convex Criteria for Training Neural Networks, *Proceedings of the 2001 Conference on Artificial Neural Networks in Engineering*, St. Louis, Missouri, November 2001.
7. Avoiding Poor Local Minima in Training Multilayer Perceptrons, *Proceedings of the Eighth International Conference on Neural Information Processing*, pp. 1327-1332, Shanghai, China, November 2001 (with Devasis Bassu).
8. Risk-Averting Criteria for Training Neural Networks, *Proceedings of the Eighth International Conference on Neural Information Processing*, pp. 476-481, Shanghai, China.
9. Minimization through Convexitization in Training Neural Networks, Proceedings of the 2002 International Joint Conference on Neural Networks, pp. 1558-1563, Honolulu, Hawaii, May 2002.
10. Robust Approximation of Uncertain Functions where Adaptation is Impossible, Proceedings of the 2002 International Joint Conference on Neural Networks, pp. 1889-1894, Honolulu, Hawaii, May 2002 (with Devasis Bassu).
11. Robust Identification of Uncertain Dynamical Systems where Adaptation is Impossible, Proceedings of the 2002 International Joint Conference on Neural Networks, pp. 1956-1961, Honolulu, Hawaii, May 2002 (with Devasis Bassu).

(c) Manuscripts submitted, but not published

1. Recurrent Multilayer Perceptrons for Discrete-Time Dynamic System Identification, under revision.

## 3. List of participating scientific personnel

1. James T. Lo, P.I.
2. Devasis Bassu, not employed by the contract.
3. Justin Nave, not employed by the contract.

## 4. Report of inventions

1. Risk-Averting Method of Training Neural Networks and Estimating Regression Models, U.S. Patent Application No. 10/193,984, filed 07/13/2002.

## 5. Bibliography

Some typical references are listed below. More can be tracked down from these references and the publications listed in Item 3 above.


1. Simon Haykin, *Neural Networks – A Comprehensive Foundation*, Second edition, Prentice Hall, Upper Saddle River, NJ, 1999.
2. David A. White and Donald A. Sofge, Editors, *Handbook of Intelligent Control*, Van Nostrand Reinhold, New York,1992.
3. Irving W. Sandberg, James T. Lo, Craig L. Fancourt, Jose C. Principe, Shigeru Katagiri, and Simon Haykin, Editors, Nonlinear Dynamical Systems, Feedforward Neural Networks Perceptives, Wiley Interscience, New York, 2001.
4. Edward Waltz and James Llinas, *Multisensor Data Fusion*, Artech House, Boston, 1990.
5. Samuel S. Blackman, *Multiple-Target Tracking with Radar Application*, Artech House, Boston, 1986.

6. James T. Lo, Mathematical Justification of Multilayer Perceptrons with Long- and Short- Term Memories, *Intelligent Engineering Systems through Artificial Neural Networks*, Vol. 8, pp. 23-29, ASME Press, New York, 1998.

7. James T. Lo, Universal Neuroapproximation of Dynamic Systems for Robust Identification, *Proceedings of the 1998 International Joint Conference on Neural Networks*, pp. 2429-2434, Anchorage, Alaska, 1998.

8. James T. Lo, Multilayer Perceptrons and Radial Basis Functions are Universal Robust Approximators, *Proceedings of the 1998 International Joint Conference on Neural Networks,* pp. 1311-1314, Anchorage, Alaska, 1998.

9. James T. Lo, Neural Filtering - Has the extended Kalman filter been superseded? presented at the 1998 Workshop on Nonlinear Time Series for Learning, Prediction and Control, Technion, Israel, 1998.

10. James T. Lo and Devasis Bassu, Adaptive Multilayer Perceptrons, *Proceedings of the 1999 International Joint Conference on Neural Networks,* Washington, D.C. 1999.

11. James T. Lo, Initializing Multilayer Perceptrons with Interconnected Neurons, *Proceedings of the 1999 International Joint Conference on Neural Networks,* Washington, D.C. 1999.

12. James T. Lo and Devasis, Mathematical Justification of Recurrent Neural Networks with Long- and Short-Term Memories, *Proceedings of the 1999 International Joint Conference on Neural Networks*, Washington, D.C. 1999.

13. Devasis Bassu, James T. Lo and Justin Nave, Training Recursive Neural Networks With Noisy Input Measurements, *Proceedings of the 1999 International Joint Conference on Neural Networks*, Washington, D.C. 1999.

14. James T. Lo, Statistical Method of Pruning Neural Networks, Washington, *Proceedings of the 1999 International Joint Conference on Neural Networks*, Washington, D.C. 1999.