リカレントニューラルネットワークによる重合反応器内状態変化の長期予測[†]

黒田千秋^{††}・引地昭二^{*}・小川浩平

東京工業大学工学部 化学工学専攻, 152-8550 目黒区大岡山 2-12-1

重合反応器内の状態変化を長期に渡って予測することを目的とし、ニューラルネットワーク(NN)を用いて行う方法について検討した。スチレン連続塊状重合プロセスにおける反応器出口温度の変化を予測対象とし、階層型 NN とリカレント NN (RNN) を用いた長期予測方法について、RNN の構造上の問題点を明らかにし、ネットワークの隠れ層による処理法に関する 2 種類の改良方法(隠れユニットでの余分の処理を付加した H-RNN と、階層型 NN の隠れユニット計算機構をモジュールとして付加した M-RNN)を提案し、各方法の予測性能を比較評価することができた。

RNN の改良法を使用することにより、階層型 NN と比較して初期段階の予測および極値を持つ変化の予測の性能を改善することができた。特に、初期段階の予測の精度を上げるために改良され、階層型 NN の隠れ層と RNN を融合した構造を持つ M-RNN の予測性能は、予測前半は勿論、長期予測の全体に渡り満足できるものであった。

緒 言

化学プロセスにおける安全運転や品質管理のために、異常や操作変更によるプロセスの状態変化をオペレータが早期に予測し対処することが必要である。しかし、反応や装置内流動の非線形な挙動、また測定や制御の際の遅れとノイズのために、簡略化されたモデル式の係数に全ての複雑要素を集中させた実験式だけでは、十分な予測を行うことが難しい。このため、非線形で複雑な現象を精密かつ定量的に捉えるために、ニューラルネットワーク(NN)を用いた情報処理手法の応用が盛んに試みられてきている。プロセスのモデル化(Cheng *et al.*, 1995; Su and McAvoy, 1992; Watanabe, 1994)、さらに制御(Bhat and McAvoy, 1990; Chovan *et al.*, 1996)、また異常診断(Watanabe *et al.*, 1989)などに応用されている。

本研究では、重合反応器の状態変化を長期に渡って予測することを目的とし、NN を用いて行う方法について検討した。NN には多くの種類があるが、応用例の多い一般的な階層型 NN と、動的挙動の予測に適していると言われるリカレントニューラルネットワーク(RNN)を用い、それらの予測性能の比較検討を行った。

階層型 NN は過去の入出力データに基づいたモデルを学習により獲得できるが、プロセス挙動の予測のような外挿的推算を行う性能は NN の汎化能力に依存したもので十分とは言えない。一方、RNN は時系列データを直接学習することによって時間的な変化を含めたモデルを獲得できるという利点を持つ。しかし、RNN の構造や学習アルゴリズム、また応用の仕方については、明確な見解があるとは言えない状況である。そこで本研究では、ネットワーク上の情報の流れを時間展開した構造を持つ一般的な RNN を示し、その構造的な問題点を明らかにし、改良方法について検討することを主な目的としている。

1. 階層型 NN, リカレント NN の構造と改良

1.1 階層型 NN

階層型 NN は入力層、隠れ層、出力層からなる構造を持ち、入力を隠れ層で前処理した結果をさらに出力層で処理するという 2 段階の処理により出力値を得るネットワークである。Fig. 1 に示すように制御器を含めたプロセスの挙動の予測に階層型 NN を用いる場合、制御対象である状態量の未来値を出力させるために、入力ユニットには設定値の現在値と、制御量および操作量の現在値、過去値(履歴)を与えることになる。

この階層型 NN を用いて、ある時点までのプロセスデータを基にそれ以降の長期予測を行う場合には、1 サンプルステップ先の予測値を出力した後、これを入力に戻して再び 2 サンプルステップ先の予測値を計算するように、入力時系列をずらしながらネットワークの計算ステップを繰り返すことで実現する。本研究では制御系(PI 制御系)を含めた重合反応器の挙動を長期予測するために、計算ステップを進める毎に予測された制御量(状態量)によって、入力ユニットに入力される操作量を変化させる必要がある。そこで、実際の制御系と等価な計算機構を用いて、制御系が出力するはずの未来の操作量を計算して入力ユニットに戻す機構を加えた。本論文では、階層型 NN に外部(External)情報のフィードバック機構(制御系と等価な計算機構)が付加されていることから、このネットワークを E-NN と呼ぶ。

また長期予測を行う際には、予測を繰り返すことにより E-NN の出力誤差が増幅され、予測性能が徐々に悪化することになる。これを避けるためには、E-NN に十分な量の広い範囲に分散した教師データを与えて、できるだけ想起が外挿ではなく内挿になるようにしなければならない。しかし、教師データの選択は時系列用入力ユニット数にも影響され、適切な教師データを選択することは難しい。

1.2 リカレント NN (RNN)

RNN は一般的には Fig. 2 に示すような構造を持った NN であり、入力層ユニット以外のすべてのユニットが相互に結合している。情報がフィードフォワードに伝わる結合とフィードバックする結合とが混在しているが、Fig. 3 のように全ての結合をフィ

[†] 1997 年 9 月 2 日受理

化学工学会第 30 回秋季大会(福岡, 1997 年)にて発表

^{††} ckuroda@chemeng.titech.ac.jp

^{*} 住友化学工業(株)

Fig. 1 A long-term predicting method using a three-layered NN

Fig. 2 A general structure of a recurrent neural network (RNN)

ードバック結合（時間遅れ結合）として時間展開することにより扱いが容易になる。ネットワークのユニットは階層型 NN と同様に入力ユニット、隠れユニット、出力ユニットに分類することができる。本研究では、この時間展開した RNN を一般的な RNN と考えて用いた。

以上の RNN の学習アルゴリズムには、ある時刻までの出力を計算してから各時刻での誤差に基づいて重みの修正を行う BP アルゴリズムを拡張した方法 (Epochwise BPTT, Epochwise Back Propagation Through Time) と、出力計算と重み修正を同時に行う方法 (RTRL, Real Time Recurrent Learning) があるが (Douya, 1991, Nisikawa and Kitamura, 1995), 本研究ではオンラインでの学習が可能な RTRL アルゴリズムを用いた。

長期予測を行う場合の入力値の与え方は、Fig. 1 に示した E-NN の場合と同様であり、前項で説明した通りである。すなわち E-NN と同様、実際の制御系と等価な計算機構を用いて、制御系が出力するはずの未来の操作量を計算して入力ユニットに戻すことになる。また、時間展開した RNN の各時刻で入力値と教師データを与えることにより想起と学習が行われる。ここで、対象プロセスにおけるデータのサンプリング時刻は RNN の計算時刻に同期させ、ネットワークの計算が 1 回行われるたびにプロセスの 1 サンプル時間後の予測が行われるものとした。この関係は Fig. 3 に示されているが、同図で t はサンプリングステップ、 τ はネッ

Fig. 3 Application of RNN to time series processing

Fig. 4 An improved method (H-RNN) on processing by hidden layer units of RNN

トワークの計算ステップを表わしている。

1.3 RNN の改良

Fig. 3 の RNN では全ての結合をフィードバック結合（時間遅れ結合）としているために、1 回の計算ステップでは階層型 NN のような入力を隠れ層で処理してから出力層で処理をするという 2 回の処理が行われていない。すなわち、計算ステップ $\tau=2$ における出力ユニットの計算は 1 ステップ前の $\tau=1$ でのユニットの値に基づいて計算されるが、 $\tau=1$ での隠れユニットの計算は $\tau=0$ でのユニットの値に基づいて計算され、また $\tau=1$ での入力は処理されることなく荷重されて出力ユニットに直接入力される。ここで、階層型 NN が隠れ層を加えることで非線形のパターン処理能力を向上させていることを考えると、上記の RNN において $\tau=1$ での入力を一端隠れ層で処理したものも出力ユニットに伝えた方が性能の向上が計れるものと考えられる。この処理手順は Fig. 4 に示す様に入力と教師データを与えるタイミングを変えることで実現できる。本論文では、RNN に隠れ (Hidden) ユニットの余分の処理を付加していることから、この改良方法を H-RNN と呼ぶ。しかし、この H-RNN を用いて RNN と同じ時間の予測を行うには 2 倍の計算ステップを必要とし、また学習時間も倍増する欠点がある。

Fig. 5 A combined method (M-RNN) of RNN with processing by hidden layer units of three-layered NN

一方, RNN で予測を開始する時($\tau=1$ の計算をする時), $\tau=1$ の入力情報が無いため, $\tau=0$ の隠れユニットの値を決めることができず, この値を任意に決めてしまうことの悪影響が長期予測結果に現われると思われる. そこで本研究では, Fig. 5 に示す様に階層型 NN を用いて始めの 1 サンプル時間後の予測を行い, その時の NN の隠れユニットの出力値を RNN の隠れユニットの初期値 ($\tau=0$ での値) とする改良方法を考えた. 本論文では, 階層型 NN を隠れユニットの計算を行うためのモジュール (Module) として導入することから, この改良方法を M-RNN と呼ぶ.

以上, H-RNN と M-RNN の 2 種類の改良された RNN を提案し, その長期予測性能を E-NN 及び通常の RNN と比較しながら検討した.

2. 予測対象プロセスと状態量

本研究では多段攪拌槽型反応器によるスチレン連続塊状重合プロセスにおける反応器出口温度を予測対象とし, 2 段目反応器のシミュレータ (Hagimori *et al.*, 1991) を対象プロセスとして用いた. Fig. 6 に同反応器シミュレータの概念図を示す. 反応器の内部を 10 個の領域に分割し, 各領域での反応速度や粘度などを推算しつつ, 温度 T , 重合率 X , 重量平均分子量 M_w の経時変化を物質収支, 熱収支に基づき RKG 法により数値積分して求めている.

運転は熱的に不安定な状態付近で行われ, PI 制御により冷媒温度 T_j を操作して反応器出口温度 T_{out} を制御しており, 定常運転となれば原料供給流量 F_{in} , 攪拌回転数 N の条件などを設定すれば出口重合率 X_{out} も決まってくる. T_{out} の測定精度は 0.1 K で, ± 0.2 K の範囲のランダムノイズが測定値に含まれる. Fig. 7 に運転実験の一例を示す. 同図は出口設定温度 T_s を変更したときの T_{out} の経時変化であり, T_{out} が新たな設定温度に移行することに成功した一例である. しかし, 制御が追い付かず反応が暴走することもあり, 安定した運転のためには操作条件の変更に伴う挙動を予測することが重要である. 本研究では製品の品質に密接に関係し, オンライン計測の可能な反応器出口温度 T_{out} の予測

Fig. 6 An outline of simulation for a continuous bulk polymerization process

Fig. 7 An example of control of T_{out} by T_j

を試みる.

3. 学習と予測結果の考察

3.1 ネットワークの学習

4 種類の NN (E-NN, RNN, H-RNN, M-RNN) について, 入力 4 つの設定値 (出口設定温度 T_s , 出口設定重合率 X_s , 原料供給流量 F_{in} , 攪拌回転数 N) と反応器出口温度 T_{out} , 冷媒温度 T_j とし, 制御量である T_{out} と操作量である T_j は現在値だけではなく 1 ステップ前の過去値も用いて計 8 個の入力情報とした. 長期予測に際し, 4 つの設定値として一旦は予測を開始する時点の値を使うが, 予測期間中のある時点で設定値が変更される運転条件の場合には, T_{out} や T_j と同様に時間的に変化するパターン値となる. 出力は T_{out} の 1 ステップ先の未来値 1 個とした. 隠れユニット数の適正値については確証は得られないため, E-NN の学習状態を基準にして本研究の条件範囲では十分であると考えられる 8 個に固定した. 本研究における以上の諸量の変化範囲を Table 1 に示したが, これらの値を用いて NN の入力値, 出力値が 0 から 1 の範囲に収まるように規格化した.

Table 1 Ranges of variables in normalizing inputs and outputs for NN.

Variables	Ranges
Feed rate, F_{in} [m ³ /s]	$6.0 \times 10^{-4} \sim 1.0 \times 10^{-3}$
Mixing speed, N [1/s]	0~1.0
Setting temperature of T_{out} , T_s [K]	390~425
Setting monomer inversion rate, X_s [—]	0~1
Temperature of coolant, T_i [K]	340~405
Temperature at the outlet, T_{out} [K]	390~425

Fig. 8 An example of prediction of T_{out} (teacher data) using four networks

NNの学習に用いる時系列データとしては、定常運転の途中で上記の4つの設定値を変更した6種類の運転例から、連続する各40サンプリングデータを切り出し、計240データを用いた。RNN, H-RNN, M-RNNの学習では学習開始時刻($\tau=0$)から6サンプリングステップ先まで予測させつつRTRLアルゴリズムにより結合重みを修正し、学習開始時刻を変えながら同じ修正計算を繰り返し行った。なお、後述するNNの想起による予測においては、18サンプリングステップ先まで長期予測を行っているが、初期予測の性能が長期予測の性能を支配することから、予測初期の学習に重点をおいてステップ数は上述のように短めの6に設定した。

過度の学習による汎化性能の低下を避けるため、学習誤差が T_{out} の測定誤差(± 0.2 K)の範囲に到達した時点で学習を打ち切った。

3.2 ネットワークの想起による予測

設定値変更の運転データの検討により、変更に伴う T_{out} の変化は約18サンプリングステップ(90 min)以内で、その大勢を捉えられることが分かった。そこで本研究では、長期予測の時間を18サンプリングステップ先までに設定した。

学習に用いたデータを再び想起させて予測を試みた例をFig. 8に示す。一方、汎化性能を検討するために、未学習のデータに対して予測を試みた例をFig. 9に示す。いずれの場合も全体を通してM-RNNによる予測が最も良好と思われ、特に T_{out} の変化の極小点付近の傾向を良好に予測している。極値を持つような非線形特性に関する汎化性能では、4種類のNNの中でM-RNNが最も優れているものと考えられる。

3.3 予測結果の評価

本研究では1章で述べた様に、学習済みのネットワークによる

Fig. 9 An example of prediction of T_{out} (un-learned data) using four networks

Fig. 10 Comparison of the predictive performance among four networks

想起計算を繰り返し、予測ステップを更新して長期の予測を行う。そこで、学習のための教師データとは異なる汎化試験用の時系列データ(6種類の運転例から連続する各40サンプリングデータを切り出したもの)を用いて予測確認試験を行い、予測ステップ毎の自乗誤差を確認試験回数で平均した E_{ms} を用いて各NNの予測性能を評価した。一連の運転データにおいても、予測開始時刻が異なれば予測性能に差が現われるため、異なる試験用データとして扱った。Fig. 10には各NNについて予測開始後の各サンプリングステップにおける E_{ms} の結果を示す。

E-NNにおいては、初期の予測性能は良好であるが、予測ステップを進めるに従い性能が悪化してゆく。しかし後半で再び性能の改善が見られ、他のネットワークと異なる傾向を示している。これは3.1項で述べた学習方法の違いによるものと思われる。すなわち、E-NNの学習では1サンプリングステップ先の予測値の誤差だけに基いて結合重み修正が繰り返されるのに対して、RNN, H-RNN, M-RNNの学習では6サンプリングステップ先までの予測時系列の全ての誤差について結合重み修正が繰り返されている。一連の時系列を学習させていないE-NNによる長期の予測性能は、試験用データの各時点での変化の複雑さに依存する

と思われる。本研究の場合、予測の中期に極値を持つ試験用データを多く用いたため、中期での変化が複雑で予測が難しく性能が悪くなっているものと考えられる。一方、RNN、H-RNN、M-RNNの予測性能の全体的傾向は、前半が良好で予測ステップを進めるに従い徐々に性能が落ちて行くという結果となった。

RNNの予測性能はE-NNと比較し若干の改善しか見られないが、予測ステップを進めても E_{ms} の変化が小さく安定した性能が得られるという点で優れていると思われる。

H-RNNの予測性能は1, 2ステップの初期予測において特に優れているが、予測ステップを進めると性能が急速に悪化する。入力隠れ層での計算処理を付加したことで初期の予測性能は向上したものの、ネットワークの学習処理が複雑なため、汎化能力を向上させるにはより多くの教師データを用いた十分な学習を行う必要があったと考えられる。

M-RNNの予測性能は予測前半において特に良好で、後半で悪化するものの、これは他のネットワークと同様な傾向であり、全体的には満足できる結果を示した。

以上の結果より、さらに予測性能を向上させるには、最初の1, 2ステップ分の予測をH-RNNとM-RNNの合体した構造のNNで行い、学習量を減らすことを考慮して、その後はM-RNNで予測をしていくのが良いと考えられる。

結 言

多段撹拌槽型反応器によるスチレン連続塊状重合プロセスにおける反応器出口温度の変化を予測対象とし、階層型NN(E-NN)とリカレントNN(RNN)を用いた長期予測の方法について検討を行った。特に、RNNの構造上の問題点を明らかにし、その改良方法(H-RNN, M-RNN)を提案し、予測性能について検討した。その結果、以下の結論を得た。

1) RNNの改良法を使用することにより、E-NNと比較して初期予測および極値を持つ変化の予測の性能を改善することができた。

2) RNNの予測性能はE-NNと比較して若干の改善しか見られないが、予測誤差が各予測ステップでほとんど変化せず、安定した性能が得られるという点で優れている。

3) H-RNNの予測性能は初期予測において特に優れているが、予測ステップを進めると性能が急速に悪化する。長期予測を

行う場合にはネットワークの結合数の制限など、ネット構造の簡略化が必要であると考えられる。

4) M-RNNの予測性能は予測前半において特に良好で、全体的に満足できる結果を示した。

5) 学習量を特に増やさずに、さらに予測性能を向上させるには、初期予測をH-RNNとM-RNNの合体した構造のNNで行い、途中でM-RNNに切り替えるのが良いと考えられる。

Nomenclature

E_{ms}	= mean square error in predictions of T_{out}	[—]
T_j	= temperature of coolant	[K]
T_{out}	= temperature at the outlet	[K]
T_s	= setting temperature of T_{out}	[K]
t	= sampling step of T_{out}	[—]
τ	= calculating step in networks	[—]

Literature cited

- Bhat N. and T. J. McAvoy; "Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems," *Computers Chem. Engng*, **14**, 573-583 (1990)
- Cheng, Y., T. W. Karjala and D. M. Himmelblau; "Identification of Nonlinear Dynamic Processes with Unknown and Variable Dead Time Using an Internal Recurrent Neural Network," *Ind. Eng. Chem. Res.*, **34**, 1735-1742 (1995)
- Chovan, T., T. Catfolis and K. Meert; "Neural Network Architecture for Process Control Based on the RTRL Algorithm," *AIChE J.*, **42**, 493-502 (1996)
- Douya, K.; "Learning Algorithm for Recurrent Networks," *J. Soc. Instrument and Control Engineers*, **30**, 296-301 (1991)
- Hagimori, S., C. Kuroda and M. Ishida; "Fuzzy Control for Continuous Bulk Polymerization Process for Polystyrene," *Kagaku Kogaku Ronbunshu*, **17**, 127-134 (1991)
- Nishikawa, Y. and S. Kitamura; *Neural Networks as Applied to Measurement and Control*, p. 37-49, Inc., Asakura Shoten, Tokyo, Japan (1995)
- Su, H. T., T. J. McAvoy and P. Werbos; "Long-Term Predictions of Chemical Processes Using Recurrent Neural Networks: A Parallel Training Approach," *Ind. Eng. Chem. Res.*, **31**, 1338-1352 (1992)
- Watanabe K., I. Matsuura, M. Abe, M. Kubota and D. M. Himmelblau; "Incipient Fault Diagnosis of Chemical Processes via Artificial Neural Networks," *AIChE J.*, **35**, 1803-1812 (1989)
- Watanabe, N.; "Dynamic Modeling of Continuous Polymerization Process by Neural Nets," *Chemical Engineering*, **39**, 41-45 (1994)

Long-Term Predictions using Recurrent Neural Networks for State Changes in Polymerization Reactors

CHIAKI KURODA, SHOUJI HIKICHI* and KOUHEI OGAWA

Graduate School of Chemical Engineering, Tokyo Institute of Technology, Tokyo 152-8550

Key words: process information, recurrent neural network, long-term prediction, polymerization reactor.

Long-term predicting methods using neural networks (NN) are discussed for state changes in polymerization reactors. The temperature at the outlet of a continuous bulk polystyrene polymerization reactor is the present target of predictions using a layered neural network and some recurrent neural networks (RNN). Some structural problems in a general RNN are indicated, and two improvements (H-RNN with additional processing in hidden layer units, M-RNN with an additional calculating module of a hidden layer in a layered NN) are proposed on data processing and arranging by hidden layer units in RNN. As to the above networks, each predictive performance can be comparatively evaluated using mean square error.

The predictive performance of H-RNN and M-RNN is superior to that of a layered NN in the initial stage of predictions, or in the state change with maximum or minimum points. In particular, long-term predictive performance is widely satisfied by M-RNN where the combined structure of RNN with hidden units of layered NN is built to improve accuracy in initial stage of predictions.

~~~~~
* Sumitomo Chemical Co., Ltd.