

Einsatz künstlicher neuronaler Netze bei der kurzfristigen Lastprognose

Kai Hufendiek und Martin Kaltschmitt

*Universität Stuttgart, Institut für Energiewirtschaft und Rationelle Energieanwendung (IER),
Heßbrühlstraße 49a, D-70565 Stuttgart*

1 Einführung

Um die Kosten zur Deckung der Nachfrage nach leitungsgebundenen Energieträgern zu minimieren, wird in Energieversorgungs- und anderen Unternehmen die Energiebereitstellung ggf. einschließlich des Kraftwerkseinsatzes geplant. Aufgrund der anstehenden Liberalisierung des Strommarktes wird diese Planung für den wirtschaftlichen Erfolg eines Unternehmens zunehmend wichtiger, da zu erwarten ist, daß sich größere Variationen infolge erweiterter Verkaufs- und Bezugsstrukturen ergeben werden. Da außerdem damit zu rechnen ist, daß zur optimalen Nutzung dieses entstehenden Marktes Verträge durchweg mit kürzeren Laufzeiten abgeschlossen werden, um flexibel auf Marktänderungen reagieren zu können, steigt durch den Wegfall langfristiger Optimierungsmöglichkeiten die Bedeutung der kurz- und mittelfristigen Optimierung der Deckung der Energienachfrage [14].

Für diese wichtigen Planungsaufgaben müssen die Ganglinien der zu erwartenden Nachfrage nach Strom, Fernwärme und ggf. Gas bekannt sein. Da eine optimale Nutzung der Planungsergebnisse wiederum direkt von der Güte der Prognose abhängt, ist es wichtig, den zeitlichen Verlauf der nachgefragten Endenergieträger verlässlich vorherzusagen.

In der betrieblichen Praxis werden eine Vielzahl verschiedener Ansätze und Verfahren zur Lastprognose eingesetzt. Dabei ist derzeit das Verfahren der multiplen Regression am weitesten verbreitet [4]. In jüngster Zeit wird auch dem Einsatz künstlicher neuronaler Netze eine hohe Aufmerksamkeit geschenkt, was u. a. die große Zahl von Veröffentlichungen auf diesem Gebiet zeigt: In einem 1996 erschienenen Übersichtsartikel zu diesem Thema werden bereits über einhundert Literaturverweise aufgeführt [6].

Vor diesem Hintergrund ist es das Ziel der folgenden Ausführungen, die Möglichkeiten und Grenzen des Einsatzes künstlicher neuronaler Netze zur Lastprognose zu untersuchen. Dabei zeigt sich, daß - obwohl künstliche neuronale Netze teilweise bereits mit Erfolg zur Lastprognose eingesetzt werden - v. a. im Bereich der Entwicklung und Anpassung solcher Systeme noch einige Probleme bestehen, die sich bisher nur durch „Versuch und Irrtum“ lösen lassen. Für eine breite wirtschaftliche Anwendung individuell angepaßter Systeme ist jedoch eine zielgerichtete methodische Vorgehensweise nötig, für die abschließend ein in mehreren Schritten aufgebautes Schema vorgestellt wird.

2 Künstliche neuronale Netze

Künstliche neuronale Netze weisen als gemeinsames Merkmal auf, daß sie aus einer Anzahl einfacher, miteinander verbundener, funktionaler Einheiten, den Neuronen, aufgebaut sind (Abb. 1) [3].

Im Bereich der Lastprognose werden künstliche neuronale Netze häufig analog den weitverbreiteten Regressionsverfahren zur Modellierung des Zusammenhangs zwischen der Last und ihren Einflußgrößen eingesetzt. Daneben gibt es auch künstliche neuronale Netze, die der Klassifikation typischer Lastmuster dienen.

Theoretische Untersuchungen belegen, daß schon ein einfacher Typus der künstlichen neuronalen Netze beliebige nichtlineare, stetige Zusammenhänge beliebig genau approximieren kann, wenn das Netz nur genügend, aber endlich, groß ist [5], [7]. Darüber hinaus wurde der Nachweis erbracht [1], daß speziell bei mehrdimensionalen Eingangsgrößen solche künstlichen neuronalen Netze der klassischen multiplen Regression überlegen sind.

Da bei der Lastprognose in der Regel mehrere Einflußgrößen berücksichtigt werden und teilweise hoch nichtlineare Beziehungen vorliegen, wird erwartet, daß sich diese theoretisch belegten Vorteile der künstlichen neuronalen Netze auch bei der Lastprognose bemerkbar machen. Zudem bieten künstliche neuronale Netze weitere Vorteile (vgl. auch [9]), die eine Anwendung im Bereich der Lastprognose attraktiv erscheinen lassen, wie z. B. das robuste Verhalten gegenüber Ausreißern in den Datensätzen oder die unproblematische Handhabung durch am Markt verfügbare Simulatoren.

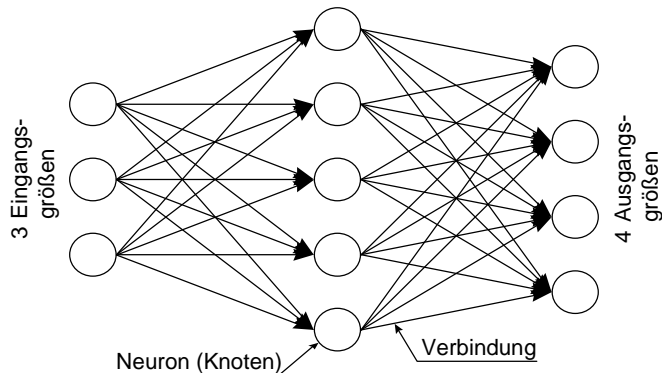


Abb. 1: Struktur eines künstlichen neuronalen Netzes (exemplarisch)

3 Lastprognosesysteme

Der Einsatz künstlicher neuronaler Netze zur Lastprognose ist im Rahmen sehr verschiedener Systemstrukturen denkbar. Dies zeigt auch die Vielzahl der in der Literatur vorhandenen Prognosesysteme (vgl. [10]).

Als Eingangsgrößen werden in der Regel exogene Einflußgrößen (z. B. Wettergrößen, Wochentage, Ferienzeiten) und im Falle von autoregressiven Zusammenhängen auch Lastdaten vergangener Zeiträume (z. B. die Lastwerte des vergangenen Tages, der letzten Stunden und/oder des gleichen Wochentages in der Vorwoche) verwendet. Als Ausgangsgröße wird die Lastprognose erzeugt.

Bei der Systemarchitektur reicht die Spannbreite vom Einsatz eines einzelnen künstlichen neuronalen Netzes bis hin zu weit mehr als zehn miteinander verbundenen bzw. parallel verknüpften Netzen (vgl. [6]). Prinzipiell lassen sich beliebig komplexe Architekturen erstellen, bei denen verschiedene künstliche neuronale Netze hintereinander geschaltet werden, um z. B. den Ausgang des ersten Netzes nochmals zu korrigieren oder die Ausgänge verschiedener parallel vorgeschalteter Netze, die unterschiedliche Lastmodelle darstellen, zusammenzufassen.

Auch hinsichtlich der zum Belernen der Netze verwendeten Datenbasis lassen sich die vorhandenen Prognosesysteme unterscheiden (vgl. [6]). Daraus folgen unterschiedliche Eigenschaften, wie z. B. ein stabiles Verhalten bei einer umfassenden festen Datenbasis von einem oder mehreren Jahren oder eine schnelle Anpassung an Veränderungen bei einer „wandernden“ Datenbasis, die beispielsweise nur den Zeitraum der letzten sechs Wochen vor der Prognose umfaßt. Kombinierte Systeme suchen darüber hinaus die Vorteile dieser beiden Ansätze zu verbinden (vgl. [10]).

Welche Systemarchitektur auf welcher Datenbasis letztlich zum Einsatz kommen sollte, hängt dabei von den spezifischen Anforderungen und Gegebenheiten des betrachteten Unternehmens ab.

4 Bestehende Probleme

Obwohl die dargestellten Vorteile - u. a. universelle nichtlineare Funktionsapproximation, genauere Approximation im Vergleich mit den üblichen Regressionsverfahren - einen breiten Einsatz künstlicher neuronaler Netze in der Praxis erwarten lassen, befinden sich derartige Prognosesysteme bisher nur in eingeschränktem Maße in der industriellen Anwendung. Dies ist nicht zuletzt auf einige bestehende Schwierigkeiten bei der Handhabung künstlicher neuronaler Netze zurückzuführen, in denen möglicherweise auch die Ursache für die in der Literatur recht unterschiedliche Beurteilung solcher Lastprognoseverfahren bei Vergleichsstudien (z. B. [2], [8], [11], [13]) zu suchen ist.

Probleme im Zusammenhang mit zur Lastprognose eingesetzten künstlichen neuronalen Netzen (vgl. [10]) ergeben sich durch offene Fragen insbesondere in den folgenden Bereichen:

- Systemarchitektur
 - Anzahl der verwendeten Netze
 - Verknüpfung dieser Netze
- Datenauswahl/-aufbereitung
 - Identifikation geeigneter Einflußgrößen
 - Normierung/Codierung der Ein- und Ausgangsgrößen
 - Auswahl der Lerndaten, d. h. der Vergangenheitsdaten, die zur Modellanpassung verwendet werden
- Netzstruktur
 - Anzahl Neuronen und Verbindungen
 - Anzahl verdeckter Schichten
 - Verknüpfung der Neuronen
- Modellanpassung
 - Wahl eines geeigneten Lernverfahrens
 - Übertrainieren, d. h. das Rauschen in den Lerndaten wird angepaßt, die Generalisierung des zugrunde liegenden funktionalen Zusammenhangs verschlechtert sich
 - schlechte Modellanpassung durch Abbruch des Lernverfahrens in lokalem Minimum der Fehlerfunktion

Mit Erfolg realisierte Lastprognosesysteme mit künstlichen neuronalen Netzen zeigen allerdings (vgl. [6]), daß diese Probleme lösbar sind. Jedoch ist die Entwicklung solcher Systeme bislang durch ein stark experimentelles Vorgehen geprägt, das durch „Versuch und Irrtum“ aufwendig ein für das jeweilige Unternehmen angepaßtes Prognosesystem erstellt.

5 Entwicklungsmethodik

Sollen künstliche neuronale Netze, die aufgrund ihrer Eigenschaften eine Verbesserung der Prognosegüte erwarten lassen, eine weite Verbreitung bei der Lastprognose finden, müssen solche Systeme zielgerichtet an die spezifischen Anforderungen des jeweiligen Unternehmens angepaßt werden können. Dazu ist die Entwicklung einer methodischen Vorgehensweise

notwendig, die sich in die in Abb. 2 dargestellten Schritte gliedern läßt. Für die einzelnen Schritte sind dabei Verfahrensweisen zu entwickeln, die ein schematisiertes und damit zumindest in Teilen automatisierbares Vorgehen erlauben.

Ausgehend von den Daten der Last und der möglicher Einflußgrößen sowie weiteren beim betreffenden Unternehmen bekannten Informationen zum Lastverhalten, wird zunächst im Schritt „Datenanalyse“ nach typischen Lastmustern und Zusammenhängen zwischen Last und Einflußgrößen gesucht. Hierzu können herkömmliche statistische Verfahren wie die Clusteranalyse oder regressionsanalytische Ansätze zum Einsatz kommen. Da einer der wichtigsten Vorteile der künstlichen neuronalen Netze gegenüber der klassischen multiplen Regression in der Modellierung nichtlinearer Zusammenhänge besteht, müssen allerdings entsprechend erweiterte Ansätze bei der Lastanalyse verwendet werden, die die Identifikation auch nichtlinearer Einflüsse ermöglichen (z. B. das Konzept der lokalen linearen Korrelationen [12]).

Sind die für eine hinreichend genaue Modellierung notwendigen Einflußgrößen identifiziert und ist geklärt, für welche Teilbereiche (z. B. unterschiedliche Typtage), aufgrund des unterschiedlichen Verhaltens eine getrennte Modellierung erfolgen soll, müssen die Daten für die Modellerstellung, d. h. das Belernen des Netzes, aufbereitet werden (Schritt „Datenaufbereitung“). Hier sind teilweise noch zu entwickelnde heuristische Vorgehensweisen einzusetzen, die die unterschiedlich skalierten Einflußgrößen geeignet codieren und normieren. Ferner sind aus den vorhandenen Daten unabhängige Lern- und Validierungsdatensätze zu generieren, deren gleichmäßige Verteilung im Datenraum entscheidend für das Ergebnis der Modellanpassung ist.

Anschließend muß im Schritt „Netzarchitektur“ festgelegt werden, wie zum einen die ggf. unterschiedlichen Netze in der Systemarchitektur zu verknüpfen sind und zum anderen wie die einzelnen Netze selbst zu strukturieren sind. Hier ist ebenfalls die Entwicklung einer entsprechenden Heuristik sinnvoll, die z. B. Richtlinien für die Netzgröße in Abhängigkeit der Anzahl Lerndaten und dem darin enthaltenen Rauschen sowie Erkenntnisse zu den Effekten der Verwendung mehrerer verdeckter Schichten und unterschiedlicher Vernetzung beinhaltet.

Das Lernverhalten des Netzes wird außer durch die Datenaufbereitung und Netzarchitektur auch durch die Wahl des Lernverfahrens, die Initialisierung, d. h. die Vorgabe eines Startwertes für den Lernvorgang, und die Wahl der „Lernparameter“ des Lernverfahrens beeinflusst. Um sicherzustellen, daß sich das Verfahren (Schritt „Lernen“) nicht in einem lokalen Optimum verfängt, sollte hier ein Lernverfahren eingesetzt werden, das kleine lokale Minima überwinden kann; zudem sollte nacheinander von unterschiedlichen Punkten (Initialisierungen) aus gestartet sowie evtl. mit verschiedenen Lernparametern gelernt werden.

Anhand des Validierungsergebnisses (u. a. Fehler auf Validierungsdatensatz) wird im Schritt „Validierung“ entschieden, ob durch ein zielgerichtetes Anpassen der vorhergehenden Schritte weitere Verbesserungen erreicht werden können. Bei Bedarf sind dabei nicht nur die Schritte „Netzarchitektur“ und „Lernen“ erneut zu durchlaufen, sondern auch der Schritt

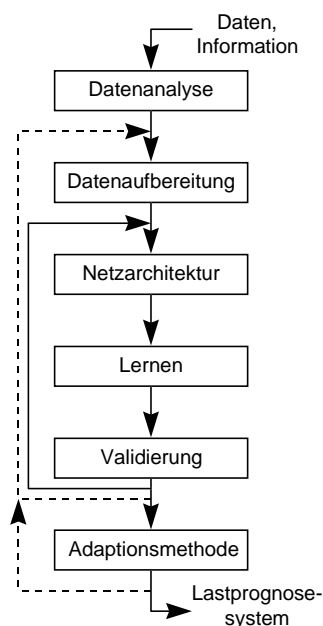


Abb. 2: Entwicklungsschema

„Datenaufbereitung“, wenn das Validierungsergebnis es nahelegt, die Einflußgrößen bzw. deren Aufbereitung zu verbessern. Für dieses iterative Vorgehen bietet sich der Einsatz von automatischen Verfahren an, die mit Hilfe genetischer Algorithmen das Ergebnis zielgerichtet verbessern.

Ist auf diese Weise ein geeignetes Modell für die Lastprognose erstellt, muß im abschließenden Schritt „Adaptionsmethode“ eine Strategie entwickelt werden, die eine Adaption des Modells an die aktuelle Entwicklung des Lastverhaltens ermöglicht. Um ein erneutes Durchlaufen der Iterationsschleife mit einer Modifikation des Lastmodells in der Entwicklungsphase zu vermeiden, sollte die Möglichkeit der Adaption an neue Daten innerhalb sinnvoller Zeiträume bereits beim Abarbeiten der vorhergehenden Schritte berücksichtigt werden.

Gelingt es, diese prinzipielle methodische Vorgehensweise entsprechend auszugestalten und ggf. zu automatisieren, ist zu erwarten, daß künstliche neuronale Netze aufgrund ihrer theoretisch belegten Vorteile gegenüber herkömmlichen Regressionsverfahren vermehrt zur Lastprognose eingesetzt werden und für den Anwender mit einem wirtschaftlichen Vorteil verbunden sind.

Literatur

- [1] Barron, A.: Universal approximation bounds for superpositions of a sigmoidal function. In: IEEE Trans. on Information Theory, Vol. 39, No. 3, 1993; pp. 930–945.
- [2] Brace, M.C.; Schmidt, J.; Hadlin, M.: Comparison of the forecasting accuracy of neural networks with other established techniques. In: Proc. of Applications of Neural Networks to Power Systems (ANNPS '91), IEEE, Piscataway, NJ, U.S.A, 1991; pp. 31-35.
- [3] Cichocki, A.; Unbehauen, R.: Neural Networks for Optimization and Signal Processing. John Wiley & Sons and B. G. Teubner, Chichester and Stuttgart, 1993.
- [4] CIGRE Working Group WG 39.03: Present practices on load forecasting and load management: A survey. In: Electra, No. 145, Paris, December, 1992; pp. 69–89.
- [5] Cybenko, G.: Approximation by superpositions of a sigmoidal function. In: Mathematics of Control, Signals and Systems, Springer, New York, Vol. 2, 1989; pp. 303–314.
- [6] Czernichow, T.; Piras, A.; Imhof, K.; Caire, P.; Jaccard, Y.; Dorizzi, B; Germond, A.: Short term electrical load forecasting with artificial neural networks. In: Engineering Intelligent Systems, Vol. 4, No. 2, 1996; pp. 85–99.
- [7] Hornik, K.; Stinchcombe, M.; White, H.: Multilayer feedforward networks are universal approximators. In: Neural Networks, Vol. 2, No. 5, 1989; pp. 359–366.
- [8] Hüsselmann, C.; Kraemer, W.; Klaus, M.: Gasverbrauchsprognose und Energiemanagement mit Neuronalen Netzen. In: VDI-Berichte Nr. 1226: Fortschrittlicher Rechnereinsatz in der kommunalen Energieversorgung. Tagung Nürnberg, 1.12.1995, VDI-Verlag, Düsseldorf, 1995; S. 139-154.
- [9] Hüvel, B.; Wülbeck, H.-F.: Lastprognose für den Strom- und Wärmebedarf bei den Stadtwerken Karlsruhe. VDI-GET-Tagung „Betriebsmanagement-Systeme in der Energiewirtschaft“, 12./13.3., Darmstadt, 1996.
- [10] Hufendiek, K.; Albiger, J.: Künstliche neuronale Netze bei der kurzfristigen Lastprognose für die Kraftwerkseinsatzoptimierung. VDI-GET-Tagung „Optimierung in der Energiewirtschaft II“, 9.10.97, Leverkusen. In: VDI-Berichte Nr. 1352, VDI-Verlag, Düsseldorf, 1997; S. 325–341.
- [11] Papalexopoulos, A.D.; Hao, S.; Peng, T.M.: An implementation of a neural network based load forecasting model for the EMS. In: IEEE Transactions on Power Systems, Vol. 9, No. 4, 1994; pp. 1956-1962.
- [12] Piras, A.: A multiresponse structural connectionist model for short term electrical load forecasting. Thèse No. 1546, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Schweiz, 1996.
- [13] Srinivasan, D.; Liew, A.C.; Chen, J.S.P.: A novel approach to electrical load forecasting based on a neural network. In: Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), Singapore, 18-21/11/1991, IEEE, Piscataway, NJ, U.S.A, 1991; pp. 1172-1177.
- [14] Verstege, J.; Werner, T.G.; Schäfer, K.F.: Liberalisierung der Energieversorgung - Auswirkung auf Planungs- und Optimierungsaufgaben. VDI-GET-Tagung „Optimierung in der Energieversorgung II“, 9./10.9.97 Leverkusen. In: VDI-Berichte Nr. 1352, VDI-Verlag, Düsseldorf, 1997; S. 9–24.

Einsatz künstlicher neuronaler Netze bei der kurzfristigen Lastprognose

Application of Artificial Neural Networks for Short-term Load Forecasting

Kai Hufendiek und Martin Kaltschmitt

*Universität Stuttgart, Institut für Energiewirtschaft und Rationelle Energieanwendung (IER),
Heßbrühlstraße 49a, D-70565 Stuttgart*

Kurzfassung

Um die erweiterten Möglichkeiten des Stromhandels, die sich durch die geplante Liberalisierung des Strommarktes ergeben, optimal nutzen zu können, muß die Planung zur Deckung der Stromnachfrage in Energieversorgungs- und anderen Unternehmen auf einer verlässlichen Lastprognose beruhen.

Künstliche neuronale Netze, über deren Möglichkeiten bei der Lastprognose ein kurzer Überblick gegeben wird, weisen in diesem Zusammenhang, u. a. gegenüber der klassischen multiplen Regression, Vorteile auf. Anhand typischer Merkmale werden die Lastprognosesysteme mit künstlichen neuronalen Netzen, die teilweise bereits mit Erfolg eingesetzt werden, kurz charakterisiert. Darüber hinaus werden noch vorhandene Probleme im Umgang mit dieser Methode aufgezeigt, die vor allem darin bestehen, daß die Entwicklung solcher Systeme bisher weitgehend auf Versuch und Irrtum basiert. Daher wird abschließend eine entsprechende Entwicklungsmethodik vorgestellt und diskutiert, die zwar im Detail noch auszugeskizziert ist, auf die aber für eine breite wirtschaftliche Anwendung individuell angepaßter Systeme nicht verzichtet werden kann.

Abstract

To yield the optimum benefit from the various possibilities for trading electricity in the intended liberalised electricity market, the planning process of utilities or other companies for meeting the actual demand must be based on a reliable load forecast.

In this context, a brief survey of artificial neural networks applied to load forecasting is presented. The advantages, compared especially to conventional multiple regression, of these networks are discussed. The typical features of load forecasting systems based on artificial neural networks, which have already been successfully applied in some cases, will be characterised concisely. Furthermore, current problems regarding the use of such networks are discussed. These mainly result from the fact, that the development of such systems is based, to a great extent, on trial and error. Finally, an appropriate design scheme is introduced, which is essential for the broad based commercial application of individually adapted systems, although the details still need to be developed.