

Abschlussbericht

gemäß BNBest-BMBF 98 (3.2)

Zuwendungsempfänger: FUB, Prof. Dr. Christof Schütte

Förderkennzeichen: 03SCM1B2

Laufzeit des Vorhabens: 01.10.2000 - 30.09.2003

Berichtszeitraum: 2001-2003

1. Kurzdarstellung

1.1. Aufgabenstellung

Im Laufe des Projekts wurden vier Hauptaufgaben bearbeitet:

- Verbesserung Parameteroptimierung für chemische und biokinetische Reaktionssysteme
- Entwicklung einer Methode zur Modelldiskriminierung für die Modellierungsphase bei schwacher Datenlage
- Implementierung der entwickelten Modelldiskriminierungsmethode
- Testen, Validieren und Anwenden der verbesserten Parameterschätzung und der neu entwickelten Modelldiskriminierungsmethode an real existierenden Reaktionssystemen

1.2. Planung und Ablauf des Vorhabens

Das Projekt läßt sich in folgende Abschnitte grob einteilen:

- Verbesserung des Konvergenzverhaltens des Optimierungsalgorithmus in der Parameterschätzung für existierende industrielle Anwendungen
- Entwicklung des Overlap-Konzepts zur Modelldiskriminierung
- Implementierung des Overlap-Algorithmus in die Softwarepakete PRESTO™ und PREDICI™

- Testen und Verifizieren des neuen Modelldiskriminierungsalgorithmus durch Vergleichsrechnungen
- Anwenden des neuen Modelldiskriminierungsalgorithmus für industrielle Anwendungen (Polymerreaktionen)

1.3. Voraussetzungen, unter denen das Projekt durchgeführt wurde sowie wissenschaftlicher und technischer Stand zu Beginn

In den letzten Jahren bekommen Modellierungs- und Simulationstechniken in der chemischen und pharmazeutischen Industrie ein immer stärkeres Gewicht. Grund hierfür ist unter anderem die verbesserte Kenntnis der Reaktionsteilprozesse und eine damit einhergehende höhere Komplexität der Modelle. Ziel der industriellen Forschung ist der Aufbau von Modellierungen komplexer Vorgänge mit ausreichender Verlässlichkeit, die eine Optimierung der Prozeßabläufe anhand der Modelle möglich macht. Dazu müssen die mathematischen und software-technischen Werkzeuge zur Modellierung, Simulation und Optimierung permanent weiterentwickelt werden. Gleichzeitig stellen sich aber neuartige Probleme, sobald der technologische Fortschritt die Grenzen der Gültigkeit der akzeptierten Modelle überschreitet. Dann entsteht häufig die Situation, daß auch Experten nicht entscheiden können, welches von einer Reihe möglicher Modelle für die in Frage stehende Problemklasse das verlässlichste ist und daher für die nächsten Modellierungsschritte als Ausgangspunkt dienen sollte. Hier sollte das Projekt ansetzen: es sollten neue mathematische und algorithmische Ansätze zur Modelldiskriminierung entwickelt werden, mit dem Ziel den Schritt der Auswahl eines Modells wenigstens teilweise zu automatisieren und qualitativ abzusichern.

Die beiden Softwarepaketen PRESTO™ und PREDICI™ Firma CiT GmbH unterstützen die Modellierung, Simulation und Optimierung komplexer chemische und biotechnologischer Systeme, inklusive aller Teilschritte von der Modellerstellung, über die Parameterschätzung, bis hin zur Vernetzung von Modellen für Teilprozesse. Diese Softwarepakete sind in allen großen Firmen der chemischen Industrie weltweit im Einsatz und sollten daher als Basis für die Entwicklung

dienen. Ziel war es, algorithmische Neuentwicklungen direkt hier einzubauen und unmittelbar im industriellen Alltag zu nutzen und zu validieren.

Ansätze wie diesen gibt es im Bereich der experimentellen und statistischen Modelldiskriminierung, aber nicht im Bereich der Modellierung mittels dynamischer Systeme. Die anfänglichen Hauptschwierigkeiten waren das Fehlen eines im industriellen Alltag tauglichen Diskriminierungskonzepts (woran soll man die Qualität eines Modells messen?), algorithmische Probleme schon mit den Standardmethoden im Bereich Modellanpassung (Konvergenzprobleme der Parameterschätzung für große dynamische Systeme mit vielen Parametern und relativ wenig Daten) und das Vorhandensein einer großen Lücke zwischen den in der universitären Forschung behandelten Fragen und der Realität der Anwender in der Industrie.

1.4. Zusammenarbeit mit anderen Stellen

Während des Projekts fanden mehrere Treffen zwischen allen Projektpartnern an verschiedenen Ort der Kooperationspartner (Berlin, Ludwigshafen, Rastede) statt, an denen die Ergebnisse und der weitere Projektverlauf festgeschrieben wurden. Des Weiteren fanden paarweise Treffen der Projektpartner an verschiedenen Ort statt. Ziel der Projektpartner war es, im Ringtausch zwischen Neuentwicklung eines Diskriminierungskonzepts und seiner algorithmischen Realisierung (FU/CiT GmbH), der Implementierung als Testversionen und im Rahmen von PRESTO™ und PREDICI™ (CiT GmbH/FU), der Validierung an Testbeispielen (FU/CiT GmbH/BASF) und der Anwendung in der industriellen Praxis im Vergleich mit realen Datensätzen (BASF) Schritt für Schritt zu einem industriell tauglichen Konzept mit Verfügbarkeit in einem industriellen Standard genügenden Softwaresystem zu kommen. Die durch die Projektmittel finanzierten Mitarbeiter waren an allen diesen Aspekten beteiligt, haben aber vor allem an der Entwicklung des neuen Diskriminierungskonzepts (genannt Overlap-Konzept, siehe unten für eine Darstellung) und seiner algorithmischen Umsetzung und Implementierung mitgearbeitet. Die von den gewerblichen Partnern übernommenen Teilschritte können wie folgt zusammengefasst werden.

BASF AG:

- Bereitstellung der experimentellen Daten zu Test- und Validierungszwecken des Algorithmus für die CiT GmbH und die FUB
- Bereitstellung einer detailliert dokumentierten Polymerisationreaktion zwecks Anwendung des neuen Overlapkonzepts zur Modelldiskriminierung
- Testen und Validieren der verbessertes Optimierungsalgorithmus in der Parameterschätzung
- Anwendung des Overlapkonzepts zur Modelldiskriminierung an BASF-spezifischen Modellklassen

CiT GmbH:

- Implementierung und Testen des verbesserten Optimierungsalgorithmus in der Parameterschätzung in die existierenden Plattformen PRESTO™ und PREDICI™
- Testen und Validieren des verbessertes Optimierungsalgorithmus in der Parameterschätzung an den von der BASF bereitgestellten industriellen Beispielklassen
- Implementierung und Testen des neu entwickelten Overlapkonzepts zur Modelldiskriminierung in die existierenden Plattformen PRESTO™ und PREDICI™

Darüber hinaus entwickelte sich zu Ende des Projekts eine enge Zusammenarbeit mit Dr. Illia Horenko (SfB 450) und Dr. Martin Weiser (ZIB) auf dem Gebiet der Propagation von Modellvariabilitäten. Durch die Adaption des von ihnen entwickelten TRAIL-Algorithmus konnte die Aussagekraft des Overlapkonzepts für einige Beispielklassen deutlich verbessert werden.

2. Eingehende Darstellung

2.1. Erzielte Ergebnisse

Overlapkonzept. Ein neues Konzept zur Modelldiskriminierung, das Overlapkonzept, wurde entwickelt. Im Gegensatz zu bisher etablierten Verfahren werden

zur Ermittlung der Übereinstimmungsgüte von Modell und Daten einerseits die Schwankung in den Daten (Datenvariabilität) und andererseits die Variabilität im betrachteten Modell, resultierend aufgrund von Schwankungen in den Parametern (Parametervariabilität), direkt zum Abgleich gebracht (overlap). Ziel ist es, ein Ranking einer Reihe möglicher Modelle hinsichtlich der *Wahrscheinlichkeit, daß das betrachtete Modell die Daten samt ihrer Schwankungen erklären kann*, zu erstellen. Dieser aus der Statistik bekannte Grundansatz führt bei der Anwendung auf Modelle in Form von dynamischen Systemen auf eine Reihe neuartiger Probleme und unterscheidet sich auch konzeptionell von bekannten Ansätzen [1].

Zur Umsetzung des Konzepts muß auf der experimentellen Seite bei der Datenerfassung die Variabilität (Streuung, Störung, etc.) der Daten mit erhoben oder modelliert werden; gleichzeitig muß auf der Modellseite die Sensitivität des Modells bezüglich Parameterstörung charakterisierbar sein und die Variabilität des Modells für gegebene Parametervariabilität berechenbar sein. Diese Anforderung erzwingt

1. den Entwurf eines neuen Gauß-Newton Algorithmus zur Maximierung des Overlaps für ein gegebenes Modell bezüglich optimaler Wahl der Parameter. Dieser Algorithmus zeigt (wie üblicherweise ähnliche Newton-artige Algorithmen zur Optimierung mit statistisch korrelierten Parametern und daraus resultierender schlechter Kondition der Jacobi-Matrizen) unter bestimmten Voraussetzungen Konvergenzprobleme. Hier konnte partiell Abhilfe geschaffen werden, siehe unten.
2. den Entwurf eines neuartigen Algorithmus zur Propagation der Modellvariabilität. Hier stellte sich im Laufe der Arbeiten heraus, daß die Propagation im Sinne der linearen Sensitivitätstheorie i.A. nicht ausreicht, da die zu berücksichtigenden Parameterstörungen zu groß sind. Daher mußte ein Ansatz und effizienter Algorithmus zur nichtlinearen Sensitivitätsanalyse entwickelt und implementiert werden.

3. die Integration der algorithmischen Neuentwicklungen in die bestehende Software-Plattform, ihre Validierung und Anwendung im industriellen Alltag.

Diese Teilschritte konnte allesamt mit Erfolg abgeschlossen werden [1] und bieten über das Projektende hinaus Anlass zu weiteren Forschungen, insbesondere auch in Kooperation der Projektpartner.

Parameterschätzung. Für die Parameteroptimierung innerhalb des Overlap-konzepts wurden, wie bereits erwähnt, Konvergenzprobleme beobachtet. Die korrelierten Parameter in der Parameterschätzung durch das Gauß-Newton-Verfahren (Parameter und ihre Unsicherheiten sind korreliert) führen zu einem Konditionsproblem der im Gauß-Newton-Schritt verwendeten Jacobimatrix. Nach einer Korrelationsanalyse können mittels Unterraumtechniken das ursprünglich schlecht konditionierte Gleichungssystem gelöst werden.

Für die Unterraumprojektion können vom Experimentator Schwellwerte vorgegeben werden. Nach umfangreichen Testläufen an realen Daten konnten empirisch Schwellwerte ermittelt werden, die sich in der Praxis bewährt haben und dadurch eine signifikante Verbesserung im Konvergenzverhalten des Algorithmus bewirkt haben.

Nichtlineare Sensitivität. Wie bereits berichtet, reicht die Propagation der Modellvariabilitäten im Sinne einer linearen, jedoch zeiteffizienten Sensitivitätstheorie im Allgemeinen nicht aus. Andererseits ist das Abtasten des Parameterraums mit anschließender Berechnung der Einzeltrajektoren bei hochdimensionalen Problemen im Sinne einer „exakten“ Propagation rechenzeittechnisch nicht realisierbar. Deshalb wurde zur Verbesserung der Propagation von Modellvariabilitäten der von Dr. Illia Horenko (SfB 450) und Dr. Martin Weiser (ZIB) entwickelte TRAIL-Algorithmus (Trapezoid rule for adaptive integration of liouville dynamics) adaptiert und angewendet [2], [4].

In der linearen Sensitivitätstheorie bleiben gaußverteilte Parametersensitivitäten gaußverteilt. Die Modellvariabilitätsverteilung wird vom TRAIL-Algorithmus adaptive in Gaußpakete zerlegt, um sie dann die einzelnen Pakete stückweise linear zu propagieren und sie anschließend wieder zusammensetzen. Der

Linearisierungsfehler von nichtlinearen Dynamiken wird adaptiv in Ort und Zeit minimiert (siehe auch: Illia Horenko: Mathematical modelling and numerical simulation of quantum effect in molecular dynamics. Dissertation FU Berlin. 2003) Der Rechenaufwand dieser Methode ist höher als für die lineare Propagation, aber signifikant geringer als für die „exakte“ Propagation.

Beispiele. Für bereitgestellte Beispiele aus dem Bereich Biokinetik konnte die sowohl die Verbesserung der Parameterschätzung als auch das Overlapkonzept gezeigt werden. Dabei wurde deutlich, daß das Overlapkonzept andersartige Aussagen trifft, als die bisher etablierten Verfahren [1].

Untersuchungen für eine Polymerisationsreaktion werden zur Zeit vorbereitet.

Software. Sowohl das verbesserte Gauß-Newton-Verfahren als auch das Overlapkonzept wurden in die bestehenden Softwarepakete Predici™ und Presto™ implementiert. Seit der Testphase befinden sich beide Programmpakete bei der BASF im Einsatz und werden immer stärker im Prozeß der Modellierung, Simulation und Modelldiskriminierung verwendet. Durch den permanenten Austausch zwischen den Projektpartnern konnte die hohe Praktikabilität beider Programmpakete aufrecht erhalten werden.

2.2. Voraussichtlicher Nutzen

Für die einzelnen Partner ergibt sich folgender Nutzen:

FUB:

- Neues Konzept in der Modelldiskriminierung
- Ausbaumöglichkeiten in Richtung Pharmakokinetik und Spektroskopie
- Ein- und Anbettung anderer Forschungsgebiete der Gruppe

CiT GmbH:

- Ausbau der bestehenden Softwareplattformen in Richtung Modellierungs- und Simulationstoolbox

- Festigung der Ausnahmestellung der Programmpakete Presto™ und Predici™ im international Wettbewerb

BASF AG:

- Neues Modelldiskriminierungstool in der frühen Phase der Modellierung und Simulation von chemischen und biokinetischen Reaktionen in einer etablierten Softwareumgebung

2.3. Fortschritte bei anderen Stellen

In Zusammenarbeit mit der Nachwuchsgruppe „Stochastic Modelling in Pharmacokinetics“ von Dr. Wilhelm Huisinga am DFG-Forschungszentrum „Mathematik für Schlüsseltechnologie“ und Dr. Horenko wurde das Konzept der Overlap-Modelldiskriminierung an Beispielen aus der Pharmokokinetik erfolgreich angewendet [6]. Diese Kooperation wird zur Zeit fortgeführt.

Des weiteren ergibt sich eine Kooperation mit dem SFB 450 im Bereich Modelldiskriminierung in der Spektroskopie. Diese Kooperation wird zur Zeit geplant.

Während des Projekts gab es Kontakte zum SFB 540: „Model-based Experimental Analysis of Kinetic Phenomena in Fluid Multi-phase Reactive Systems“ in Aachen.

2.4. Erfolgte und geplante Veröffentlichungen

[1] Sönke Lorenz, Elmar Diederichs, Regina Telgmann, Christof Schütte: *Discrimination of dynamical system for biological and chemical systems*, Preprint, eingereicht bei SIAD.

[2] Illia Horenko, Sönke Lorenz, Wilhlem Huisinga: *Adaptive simulation of chemical reactions under parameter uncertainty*, in Vorbereitung, verfügbar Herbst 2004.

[3] Sönke Lorenz, Regina Telgmann, Christof Schütte: *Application of the overlap concept for polymerization reactions*, in Vorbereitung, verfügbar Ende 2004.

[4] Sönke Lorenz: *The overlap approach for discriminating between biological and chemical systems*. Dissertation in Arbeit. Verfügbar Ende 2004

- [5] Regina Telgmann: Numerical methods in model discrimination. Dissertation in Arbeit. Verfügbar Ende 2004
- [6] Wilhelm Huisinga, Illia Horenko, Sönke Lorenz: A new approach for incorporating parameter variability into model discrimination. Poster präsentiert auf der Cost B15 Final conference: In silico driven drug discovery and development. The integration of modelling and simulation throughout the discovery and development process. Genf 26.-28. Mai 2004
- [7] Christof Schütte, Elmar Diederichs, Sönke Lorenz, Regina Telgmann: Experimentally controlled discrimination of models, parameter identification and overlap estimation". Poster präsentiert auf dem Statusseminar des BMBF-Programms: Neue Mathematische Verfahren in Industrie und Dienstleistungen

2.5. Sonstiges

Durch das Projekt wurde ein Forschungsstipendium für Sönke Lorenz durch die FUB von der BASF AG eingeworben.