

3 Forschungsprogramm

3.1 Forschungscluster im Überblick

Die Cluster Selbstorganisation (A), Modellgetriebene Softwareentwicklung (B), Dezentrale Informationssysteme (C), Workflow-Management (D) und Katastrophenmanagement (E) werden auch in der zweiten Förderphase beibehalten. Dabei soll eine noch engere Verzahnung untereinander angestrebt werden sowie die betreffenden Fragestellungen vertieft und erweitert werden.

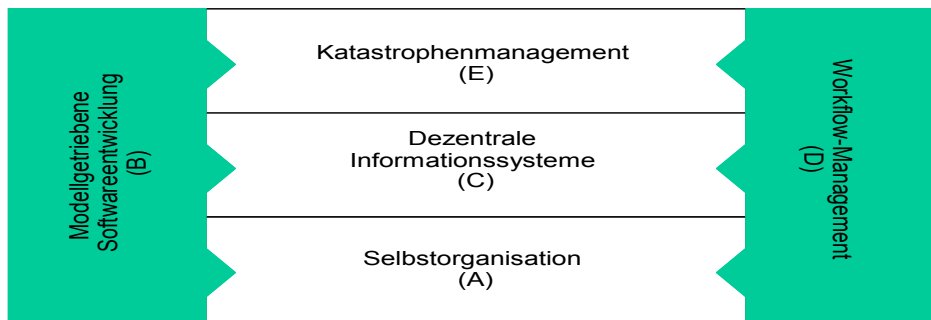


Abb. 2 Die Verzahnung der Forschungscluster

Personell verstärkt werden die Bereiche A (durch Frau Prof. Dr. Verena Hafner) und E (durch Frau Prof. Dr. Tobia Lakes). Während Frau Hafner im Bereich Robotik ausgewiesen ist und sich auf die Mobilität und Sensorik von Knoten in selbstorganisierenden Netzstrukturen konzentriert, arbeitet Frau Lakes als Geographin auf dem Gebiet der Analyse und Modellierung von Mensch-Umwelt-Interaktionen. Sie wird speziell den Umweltmodellierungsbereich verstärken. Damit vertieft das Graduiertenkolleg seinen schon bestehenden interdisziplinären Anspruch. Aus dem Kolleg ausscheiden wird Herr Dr. Holz, der vom HPI Potsdam in die Industrie nach Yokohama/Japan wechselte. Im Jahr 2012 wird Herr Prof. Malek aus Altersgründen ausscheiden.

Die gegebenenfalls notwendige Übernahme der Betreuung von Doktoranden von Herrn Malek wird durch die METRIK-Hochschullehrer Redlich und Reinefeld übernommen. Auch für den Fall, dass die derzeitige Juniorprofessorin Frau Lakes einem Ruf an einer anderen Universität annehmen sollte, ist die Übernahme der Betreuung durch das Institut für Geographie der Humboldt-Universität zu Berlin gesichert.

3.2 Profilierung des Clusters *Selbstorganisation (A)*

Hochschullehrer: Hafner, Malek, Redlich, Reinefeld

In der ersten Förderperiode wurden Verfahren zur Selbstorganisation von Netzwerken sowie Protokolle für Dienste auf unstrukturierten Overlay-Netzen untersucht, die auf den für METRIK relevanten Netzwerkstrukturen effizient arbeiten. Die Leistungsfähigkeit von drahtlosen Maschennetzwerken wird wesentlich durch Signal-Interferenzen begrenzt. Es war daher wichtig, das Entstehen von Interferenzen und deren Auswirkungen auf spezifische Leistungsparameter (Durchsatz, Latenz, Jitter) genauestens zu analysieren, um daraus neue (meist opportunistische) Kommunikationsprotokolle abzuleiten (*Jens Nachtigall*). Eine weitere, naheliegende Frage war die Skalierbarkeit von Systemen, deren Kommunikation über drahtlose Maschennetze realisiert wird. Grundsätzliche Grenzen homogener Netze sind etwa seit dem Jahr 2000 bekannt (Kumar, Gupta). Im Rahmen von METRIK wurde der Einsatz von steuerbaren Richtungsantennen (steerable beamforming antennas) zur Hierarchisierung des physikalischen Netzes und damit die Entlastung sogenannter Hotspots untersucht (*Sebastian Heglmeier*). Schließlich muss die Kommunikation gegen Angreifer abgesichert werden, was dadurch erschwert wird, dass den meisten selbstorganisierenden Systemen eine vertrauensbildende, zentrale Instanz, ähnlich einem Operator, fehlt. Für dieses Problem wurden neue Schlüsselverteilungsprotokolle entwickelt (*Henryk Plötz*). Weiterhin sind Methoden zur gegenseitigen Erkennung der Knoten auf Netzwerk- und Applikationsebene sowie zum selbständigen Aufbau einer Netzwerktopologie untersucht und weiterentwickelt worden (*Jens Nachtigall, Sebastian Heglmeier*). Zusätzlich wurden Verlässlichkeitseigenschaften wie Verfügbarkeit und Leistungsbereitschaft von statischen, verteilten Anwendungen modellbasiert untersucht, um Aussagen über Gütegarantien treffen zu können (*Andreas Dittrich*).

Methoden der Selbstorganisation, die zunächst für überwiegend statische Netzwerke ausgelegt waren, sollen in der zweiten Förderperiode für den Einsatz in hochgradig dynamischen Systemen adaptiert werden. Basierend auf den vorhandenen Fähigkeiten der beteiligten Knoten sollen Dienstleistungen wie z.B. Datenspeicherung, Sensordatenabfrage, Gruppenkommunikation, Berechnungsaufgaben und darauf aufbauende komplexere Applikationen angeboten werden. Die Dienste müssen sich automatisch an Änderungen der Umgebungsbedingungen (Netzwerkverfügbarkeit, Knotenlast, Batteriestand) sowie geänderte Nutzungsprofile anpassen können. Dazu sollen adaptive Peer-to-Peer-Netzwerke mit Gütegarantien untersucht werden. Diese Gütegarantien beziehen sich auf nicht-funktionale Eigenschaften wie Sicherheit, Zuverlässigkeit, maximale Kosten, Genauigkeit und zeitliches oder transaktionales Verhalten. Es sollen Komponenten für die Systemüberwachung, Fehlerdiagnose, Fehlervorhersage, kooperative Kommunikation und die dynamische Korrektur von Betriebsparametern (z.B. auf Basis von Regelkreisen) erforscht und entwickelt werden.

Themenvorschläge:

- Entwicklung verbesserter Datenübertragungsverfahren, z.B. durch Verbesserung bekannter opportunistischer Routingverfahren, die auch bei stark fluktuierenden Kommunikationskanälen noch leistungsfähig arbeiten,
- Ausnutzung der Mehr-Nutzer-Diversität (multi user diversity) zur Kooperation von mehreren Knoten beim physikalischen Senden und/oder Empfangen von Datenpaketen,
- Adaptive Verfahren zur Optimierung von Netzwerktopologie und Kommunikation,
- Sensorfusion und heterogene Sensornetze,
- Adaptive Überwachungs- und Kontrollmechanismen in strukturierten Peer-to-Peer-Netzwerken,
- Entwicklung von vorhersagegetriebenen Modellen zur Selbstheilung in Peer-to-Peer-Netzwerken,
- Kooperatives Sendeverhalten von Knoten als adaptives Verfahren zur Erhöhung von Sendereichweiten,
- Einhaltung von Gütegarantien durch koordinierte Kommunikation mehrerer Netzwerkknoten,
- Entwicklung verbesserter Routingverfahren durch Nutzung Gossip-basierter Kommunikation in strukturierten Overlay-Netzwerken,
- Anwendung von Konzepten aus der Regelungstechnik auf adaptive strukturierte Peer-to-Peer-Netze,
- Modellbasierte Analyse adaptiver strukturierter Peer-to-Peer-Netze.

3.3 Profilierung des Clusters Modellgetriebene Softwareentwicklung (B)

Hochschullehrer: Fischer, Reisig, Schlingloff

In der ersten Förderperiode konnten durch Adaption etablierter Compiler-, Simulations-, Codierungs-, Decodierungs- und GIS-Technologien entscheidende Grundlagen geschaffen werden, die die automatische Erzeugung anwendungsspezifischer Software-Komponenten für selbstorganisierende Sensorsysteme aus abstrakten Struktur- und Verhaltensmodellen (UML/SDL/ASN.1) erlauben. Darüber hinaus konnte gezeigt werden, dass sich modellbasierte Entwicklungsinfrastrukturen auch zur Installation und Administration von Sensornetzwerken einsetzen lassen (*Frank Kühnlenz*). Komponentenbasierte Technologien zum Deployment von Software-Komponenten selbstorganisierender Systeme ergänzen die Infrastrukturen mit hohem methodischen und praktischen Wert (*Siamak Haschemi*). Resultate dieser Forschungen flossen sowohl in die Entwicklung eines prototypischen Erdbebenfrühwarnsystems im Cluster E (siehe 3.6) als auch in die Umsetzung einer Computing-Infrastruktur zur Entwicklung optischer Nanostrukturen ein (*Arif Wider*).

Die Entwicklung domänenspezifischer Modellierungssprachen inklusive benötigter Standardwerkzeuge (wie Editoren, Transformatoren, Testwerkzeuge, Interpreter, Debugger) im Geo-Umfeld wurde durch Etablierung, Weiterentwicklung und Anwendung metamodellbasierter Technologien vorbereitet (*Markus Scheidgen, Daniel Sadilek, Guido Wachsmuth, Hajo Eichler, Michael Soden*). Eine erste erfolgreiche Anwendung dieser Technologien eröffnete in Cluster E die Entwicklung geeigneter Umgebungsmodellierungssprachen im Geo-Kontext.

Weitere künftig in den Clustern C, D und E (siehe 3.4, 3.5 und 3.6) identifizierte bzw. zu konzipierende Modellierungssprachen und Werkzeuge sollen helfen, Konzepte und Technologien im Metamodellierungsbereich zu konsolidieren und ferner dazu beitragen, konkrete spezifische Modellierungsanforderungen zu erfüllen.

Zur Untersuchung von Abstraktionen einzelner Knoten wurden in der ersten Förderperiode Verifikations- und Testmethoden für selbstorganisierende Systeme entwickelt (*Jan Calta, Stephan Weißleder*). So lag z.B. das Hauptaugenmerk im Testbereich auf der Erzeugung linearer Testfälle. Künftig ist es nötig, die Knoten nicht nur einzeln, sondern im Verbund als Gesamtsystem zu testen. Das bedeutet, dass auch die Testfälle verteilt und parallel ausgeführt werden müssen.

Allgemeine Dienste (Services), realisiert auf Basis einer Middleware für selbstorganisierende Systeme, verwirklichen ein bewährtes Konzept, das darin besteht, bestimmte Leistungen nach außen über eine einheitliche Schnittstelle unabhängig von ihrer Realisierung anzubieten. Darunter fallen die in Cluster E (siehe 3.6) für das Erdbebenfrühwarnsystem entwickelten Netzstatus-, Synchronisations- und Benachrichtigungs-Services. Nachdem in der ersten Förderperiode die Fehlerdiagnose und -korrektur dienstbasierter Anwendungen (*Niels Lohmann*), die Austauschbarkeit von Diensten (*Christian Stahl, Jarungjit Parnjai*) und die Verarbeitung von Daten in Diensten (*Christoph Wagner*) durch Einsatz von Petri-Netzen untersucht worden sind, sollen in der zweiten Förderperiode die methodische Komposition und Verifikation von Services und Korrektheitsnachweise unter Verwendung Szenario-basierter Modelle zum Gegenstand der Forschung gemacht werden.

Themenvorschläge:

- Generierung verschiedener netzwerk- und anwendungsspezifischer Simulatoren in Kopplung mit Umweltsimulatoren,
- Erweiterung des Eclipse-basierten EMF-Ansatzes um CMOF-Modellierungskonzepte der OMG zur Erreichung eines stärkeren Abstraktionsgrades bei der Beschreibung von Metamodellen als Teil von Sprachdefinitionen,

- Konzeption einer automatenbasierten, domänenspezifischen Sprache zum Entwurf strukturierter Overlay-Sensornetze,
- Verteilte Testumgebung für Sensornetzwerke,
- Modellbasierte Entwicklung von Multi-Core und Many-Core Simulationswerkzeugen,
- Spezifikation und formale Ableitung von Selbstorganisationsprinzipien,
- Methodische Komposition von Services.

3.4 Profilierung des Clusters Dezentrale Informationssysteme (C)

Hochschullehrer: Freytag, Leser, Reinefeld

Die Forschungen in diesem Bereich hatten in der ersten Förderperiode des Graduiertenkollegs die Entwicklung von Techniken zum Ziel, mit denen die Komplexität der Informationsverarbeitung in einem selbstorganisierenden Netz vor höheren Anwendungsschichten verborgen werden kann. Netze sollen als Ganzes nach außen bestimmte Dienste anbieten können, deren Realisierung transparent bleibt.

Beispielhafte Dienste sind Knotenlokalisierung, Benutzung des Netzes als hochverteiltes Speichermedium oder die Sammlung von Messwerten; potentielle Schwierigkeiten, die bei der Bearbeitung entsprechender Anfragen an das Netz auftreten und so weit wie möglich transparent behandelt werden sollten, sind Knotenausfälle, Überlastung von Knoten in Engpässen und die wechselnde Qualität der einzelnen (drahtlosen) Verbindungen. Zu diesen Problemstellungen hat das Kolleg eine Reihe von Lösungen entwickelt. Wir nennen hier insbesondere das SenseDB-System (*Timo Gläßer*) zur Abarbeitung deklarativer Anfragen in drahtlosen Netzen, die optimale Zuordnung von Workflow-Aufgaben auf Knoten eines Netzwerkes einschließlich der Fehlerbehandlung (Artin Avanes), sowie die fehlertolerante Replikplatzierung in weitverteilten Netzwerken (*Joanna Geibig*).

Viele der bisher entwickelten Methoden wurden zunächst mit vereinfachten Annahmen realisiert; in der kommenden Förderperiode sollen diese Ansätze unter weniger restriktiven Bedingungen weiterentwickelt werden. Beispielsweise setzt das SenseDB-System für die Optimierung ein weitgehend statisches Netz voraus. Generell wurde in den bisherigen Arbeiten der Aspekt der Adaptivität nur ansatzweise berücksichtigt. Daher sollen in der zweiten Förderperiode die Fragestellungen des Clusters C prinzipiell beibehalten werden, aber unter besonderer Berücksichtigung der Adaptivität und dem damit verbundenen Trade-Off zwischen Redundanz (in Kommunikation, Speicherorten etc.), Robustheit und Effizienz geschärft werden. Mit den bisherigen Ansätzen werden zwar Anforderungen für das Katastrophenmanagement grundsätzlich behandelt; dennoch ist geplant, spezifische Erweiterungen an bisherigen Anfragesprachen vorzunehmen bzw. deklarative Beschreibungen domänenspezifischer Prozeduren auf ihre Einsatzmöglichkeiten hin zu untersuchen.

Themenvorschläge:

- Multiple Query Optimierung in Sensornetzwerken,
- Adaptive Anfrageoptimierung in Sensornetzwerken,
- Adaptive Rekonfiguration unter Berücksichtigung vorgegebener Anfragen,
- Schichtenbezogene Spezifikation von Anfrageausführungsansätzen und ihre schichtenübergreifende Optimierung (Cross-Layer-Optimierung),
- Einbeziehung domänenspezifischer Funktionen in deklarative und adaptive Anfragesprachen,

3.5 Profilierung des Clusters *Workflow-Management (D)*

Hochschullehrer: Fischer, Freytag, Reisig, Schlingloff

Ausgehend von der klassischen Vorstellung, dass ein Workflow-Management-System idealerweise die partiell geordnete Abfolge elementarer Aktivitäten eines komplexen Workflows einer Organisation initiiert und steuert, ergeben sich in konkreten Anwendungsszenarien eine Reihe spezifischer Probleme, die sowohl das Katastrophenmanagement im Kontext des Clusters E (siehe 3.6) als auch die Workflow-Ausführung in unsicheren Netzen im Kontext des Clusters A (siehe 3.2) betreffen. In der ersten Förderperiode entstanden Beiträge zur Szenario-basierten Modellierung und zur Korrektheit von Workflows im Kontext mit anderen Workflows und Services (*Dirk Fahland*), zum Umgang mit spontan auftretenden Fehlern im Ablauf eines Workflows (*Artin Avanes*) sowie zur Charakterisierung von Mengen von Workflows und Services mit speziellen Eigenschaften (*Peter Massuthe*). Mit der Definition verhaltensbewahrender Transformationen industrieller Workflow- und Service-Modellierungssprachen (UML-AD, BPEL und BPMN) in formale Modelle (Petri-Netze) wurden erste Grundlagen für eine Anwendung der im Forschungscluster B entwickelten Techniken gelegt (*Dirk Fahland, Niels Lohmann, Peter Massuthe*). Eine Reihe dieser Ergebnisse entstand in Kooperation mit der TU Eindhoven im Rahmen des von METRIK mitgetragenen B.E.S.T.-Projektes (siehe 6.4).

Aufbauend auf diesen Arbeiten sollen in der zweiten Förderperiode Themen bearbeitet werden, in denen zum einen die formale Modellierung der oft nur informell beschriebenen Dienstvorschriften und Arbeitsabläufe behandelt wird, und zum anderen Abweichungen und Ausnahmen der normalen Szenarien betrachtet werden. Besonders interessant sind Fragen nach automatisch generierbaren *Adaptern*, die Unstimmigkeiten bei der Kommunikation von Workflows ausgleichen. Hierzu erwarten wir entscheidende Impulse von der B.E.S.T.-Kooperation. Fragestellungen des Clusters C können mit diesen Ansätzen angegangen werden. Hinzu kommen Arbeiten, die die Simulation von Workflow-Szenarien auf der Basis metamodellbasierter Beschreibungen erlauben, wobei sowohl Daten- als auch Zeitabhängigkeiten in den Zustandsänderungen berücksichtigt werden.

Themenvorschläge:

- Formalisierung von Anforderungen an Workflows,
- Erweiterung Szenario-basierter Modellierungstechniken zu einer Methode für die systematische Entwicklung korrekter, lose gekoppelter Systeme auf der Basis von High-Level-Petri-Netz-Modellen mit (temporal-)logischer Eigenschaftsbeschreibung unter Berücksichtigung von Zeit und Daten im Modell,
- Behandlung komplexer Fehlersituationen bei Workflow-Ausführungen in unsicheren Netzen,
- Deklarative Beschreibung von Aktionen des Katastrophenmanagements über Anfragen hinaus und ihre Umsetzung durch erweiterte Datenbanktechniken,
- Metamodellbasierte Definition von BPMN bei simulativer Modellausführung mit Umwelteinflüssen,
- Transparente Behandlung von Netzwerk- bzw. Knotenfehlern bei der Ausführung von Workflows,
- Vorausschauende Rekonfiguration von Workflow-Ausführungen bei zu erwartenden Fehlern.

3.6 Profilierung des Clusters Umweltmodellierung, Katastrophen- und Risikomanagement (E)

Hochschullehrer: Dransch, Fischer, Hafner, Lakes, Leser

Im Graduiertenkolleg wurden Methoden zur Gefahren- und Umweltmodellierung untersucht und verbessert. Dabei konnten Technologien zur modellbasierten Softwareentwicklung aus Cluster B (*Markus Scheidgen, Frank Kühnlenz*) hinsichtlich gestellter Anforderungen aus der Umweltmodellierung adaptiert und ein exemplarischer Ansatz zur Modellierung und Simulation Hybrider Zellulärer Automaten entwickelt werden (*Falko Theisselmann*). Mit letzterem wurden sowohl Feuerausbreitungen modelliert als auch erste Ansätze zur Ausbreitung von Erdbebenwellen abgeleitet. Mit diesen entwickelten Methoden ließen sich sowohl Wiederverwendbarkeit und Kopplung von Umweltsimulationsmodellen erheblich verbessern als auch Agentenmodellierungen von Katastrophenmanagementszenarien realisieren.

Auf einer gänzlich anderen Ebene der Risikobewertung bewegte sich das von mehreren Doktoranden (insb. *Bastian Quilitz*) entwickelte HOUDINI-System, mit dem verteilte und heterogene Informationen aus einer Vielzahl von Datenquellen gesammelt und mit Techniken des Semantic Web zu einer homogenen Informationsschicht integriert werden können. Eine besondere Rolle spielte dabei User Generated Content (UGC) von Webseiten wie Flickr. UGC stand auch im Mittelpunkt einer Arbeit zur schnellen Schadensabschätzung bei Flutereignissen. Überprüft wurde die generelle Eignung der Daten. Außerdem ist ein automatisiertes Verfahren zur Qualitätsbewertung aktueller Daten entwickelt worden (*Kathrin Poser*).

Ein wesentlicher Aspekt der Risikokommunikation ist die Informationsvisualisierung. Nur eine adäquate Visualisierungsform kann entscheidungs- und handlungswirksam sein; vor allem im zeitkritischen Katastrophenfall mit heterogenen, unsicheren und z.T. widersprüchlichen Informationen ist dies von außerordentlicher Wichtigkeit. Daher wurde untersucht, wie hochaggregierte Informationsdarstellungen, sogenannte chorematische Darstellungen, automatisiert erzeugt werden können (*Andreas Reimer*).

Die bisher entwickelten Ansätze brachten Ergebnisse für wichtige Teilaspekte der genannten, komplexen Themenfelder. In der zweiten Phase wird deshalb die Palette methodischer Lösungen weiter ausgebaut. Im Bereich der Gefahren- und Umweltmodellierung sollen die Defizite bei der raum- und zeitskalenübergreifenden Modellierung identifiziert und behoben werden. Um den Nutzen von UGC zu erhöhen, werden Methoden der Extraktion von Raum- und Zeitinformation aus unstrukturierten Informationsquellen verbessert. Insbesondere soll die Qualität der extrahierten Informationen bewertbar sein. Schließlich sollen Visualisierungsstrategien entwickelt werden, die die Unsicherheit wie auch die Multidimensionalität von Information effektiv für verschiedene Akteure kommunizieren können.

Themenvorschläge:

- Skalenübergreifende Umweltmodellierung,
- Kopplungen von Modellen unterschiedlicher Klassen (Siedlungsszenarium, Katastrophe, Auswirkung, Risiko), deren simulative Analyse und Bewertung,
- Räumliche Kognition in der Modellierung dynamischer Mensch-Umwelt-Systeme,
- Extraktion von Raum- und Zeitinformationen aus unstrukturierten Text- und Bild-Datenquellen,
- Integration unterschiedlicher Sensordaten in mobilen Sensorknoten,
- Automatische Kontexterkenkung und Entscheidungsfindung,
- Bewertung der Datenqualität von UGC,

- Visualisierung von Datenqualität und unsicherer Information,
- Informationsaggregation durch Visualisierung,
- Visual Analytics Methoden für Sensordaten
- Datenintegration für Katastrophenleitstände,
- TextMining von Katastrophennachrichten.