

Kontextdatenprognose auf mobilen Geräten

Abschlussbericht zur Diplomarbeit



Matthias Meiners

03.12.2009

- ◆ Problemfeld
- ◆ Bestehende Lösungsmöglichkeiten
- ◆ Verfahren der strukturierten Kontextdatenprognose
- ◆ Untersuchung der Ergebnisse und Demo
- ◆ Fazit und Ausblick

- ◆ **Problemfeld**
 - ◆ Motivation und Ziel
 - ◆ Kontext
 - ◆ Prognose
 - ◆ Anforderungen
- ◆ Bestehende Lösungsmöglichkeiten
- ◆ Verfahren der strukturierten Kontextdatenprognose
- ◆ Untersuchung der Ergebnisse und Demo
- ◆ Fazit und Ausblick

- ◆ Ubiquitous Computing:
 - ◆ Allgegenwärtigkeit von Informationstechnik
 - ◆ Einbettung in die Umgebung
 - ◆ Unterstützung des Nutzers im Alltag



- ◆ Mobile Geräte:
 - ◆ Anpassung an Mobilität des Nutzers
 - ◆ Einschränkungen gegenüber stationären Geräten

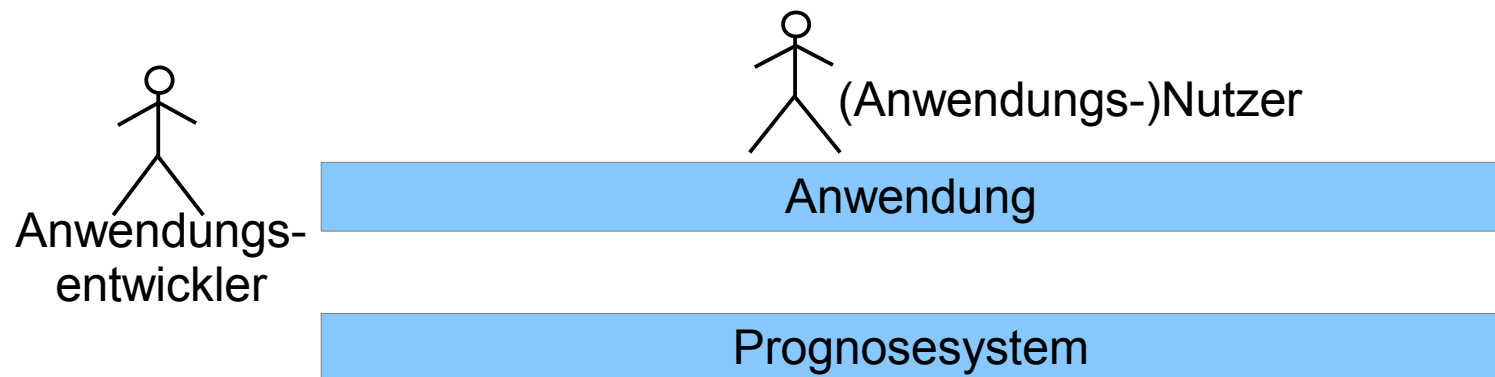
- ◆ Context-Awareness:
 - ◆ Bewusster Umgang mit der Umgebung
 - ◆ Proaktivität durch Prognose
 - ◆ Systemunterstützung benötigt

z.B. Prognose der Nutzung eines Mobiltelefons für das Energiemanagement (Bsp. 1)



Ziel der Diplomarbeit

- ◆ Entwicklung und Implementierung eines Verfahrens zur Prognose zukünftigen Kontextes:
 - ◆ Zur generischen Systemunterstützung für Anwendungen
 - ◆ Lauffähigkeit auf mobilen Geräten
 - ◆ Orientierung am Ubiquitous Computing



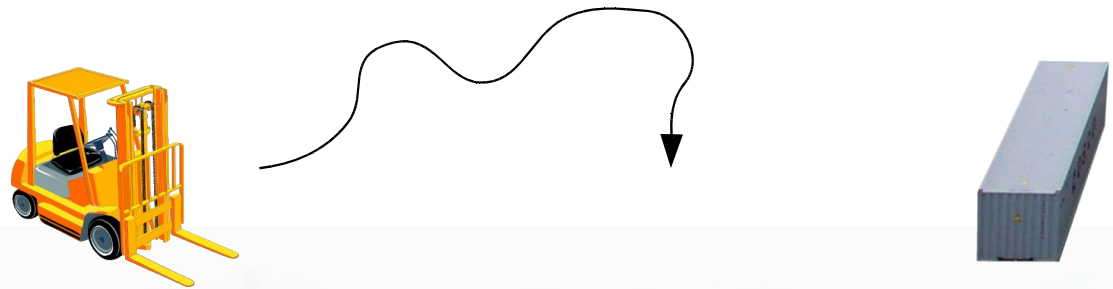
Motivation - Beispiel 2 (Dienstverfügbarkeit)

- ◆ Ziel: Kontextbasierte Kooperation in dienstorientierten, mobilen Umgebungen mit Ad-Hoc-Netzen (vgl. [Kun08], DEMAC-Projekt)
- ◆ d.h. Zuteilung einer Aufgabe zum geeignetsten Gerät, Aufgabenausführung durch lokalen oder entfernten Dienstaufruf
- ◆ Prognose der Verfügbarkeit eines Dienstes für ein Gerät zur vorausschauenden Aufteilung von Aufgaben auf Geräte

Motivation - Beispiel 2 (Dienstverfügbarkeit)

- ◆ Konkret z.B. Prognose der Verfügbarkeit eines Notfallmeldedienstes für die Aufgabe des Meldens eines defekten Kühlcontainers:

1.



2.



Motivation - Beispiel 2 (Dienstverfügbarkeit)

3.

Prognose
Verfügbarkeit
Notfallmeldedienst



4.



positive
Antwort
über
WLAN



5.



SOS
WLAN



Motivation - Beispiel 2 (Dienstverfügbarkeit)

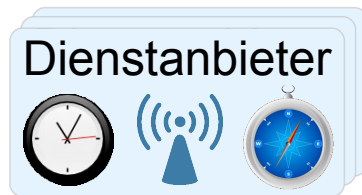
6.



7.



- ◆ Synonyme für Kontext: Umgebung, Situation
- ◆ Kontext in Beispiel 2 über Dienstverfügbarkeit:



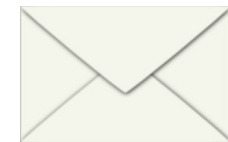
Position des
Dienstnutzers



Tageszeit



Möglichkeiten des
Dienstnutzers zur
Vernetzung



Notfallmeldedienst
Verfügbarkeit für
Dienstnutzer

Zusammenhänge

- ◆ Historische Daten: Variablen V_i mit Werten $v_{i,j}$ zu Zeitpunkten j

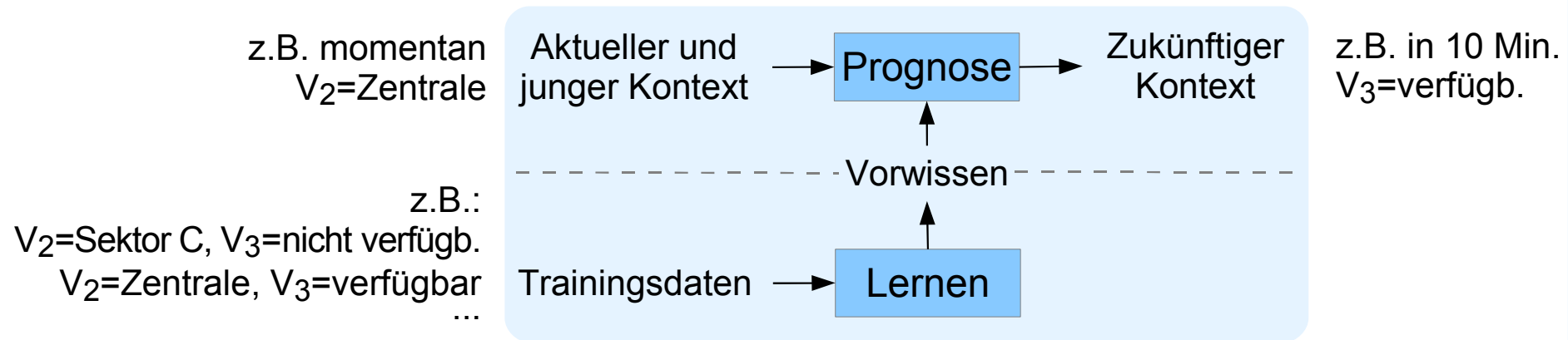
V₃	V _{3,1}	V _{3,2}	V _{3,3}	V _{3,4}	V _{3,5}
V₂	V _{2,1}	V _{2,2}	V _{2,3}	V _{2,4}	V _{2,5}
V₁	V _{1,1}	V _{1,2}	V _{1,3}	V _{1,4}	V _{1,5}

→ Zeit

z.B. V_1 : Tageszeit, V_2 : Position, V_3 : Dienst-Verfügbarkeit

- ◆ Kernziel: Prognose eines zukünftigen Variablenwertes
- ◆ Grundlage: Zusammenhänge wie „Dienst in Hafenzentrale verfügbar“
- ◆ Zusammenhänge oft mit Unsicherheit behaftet

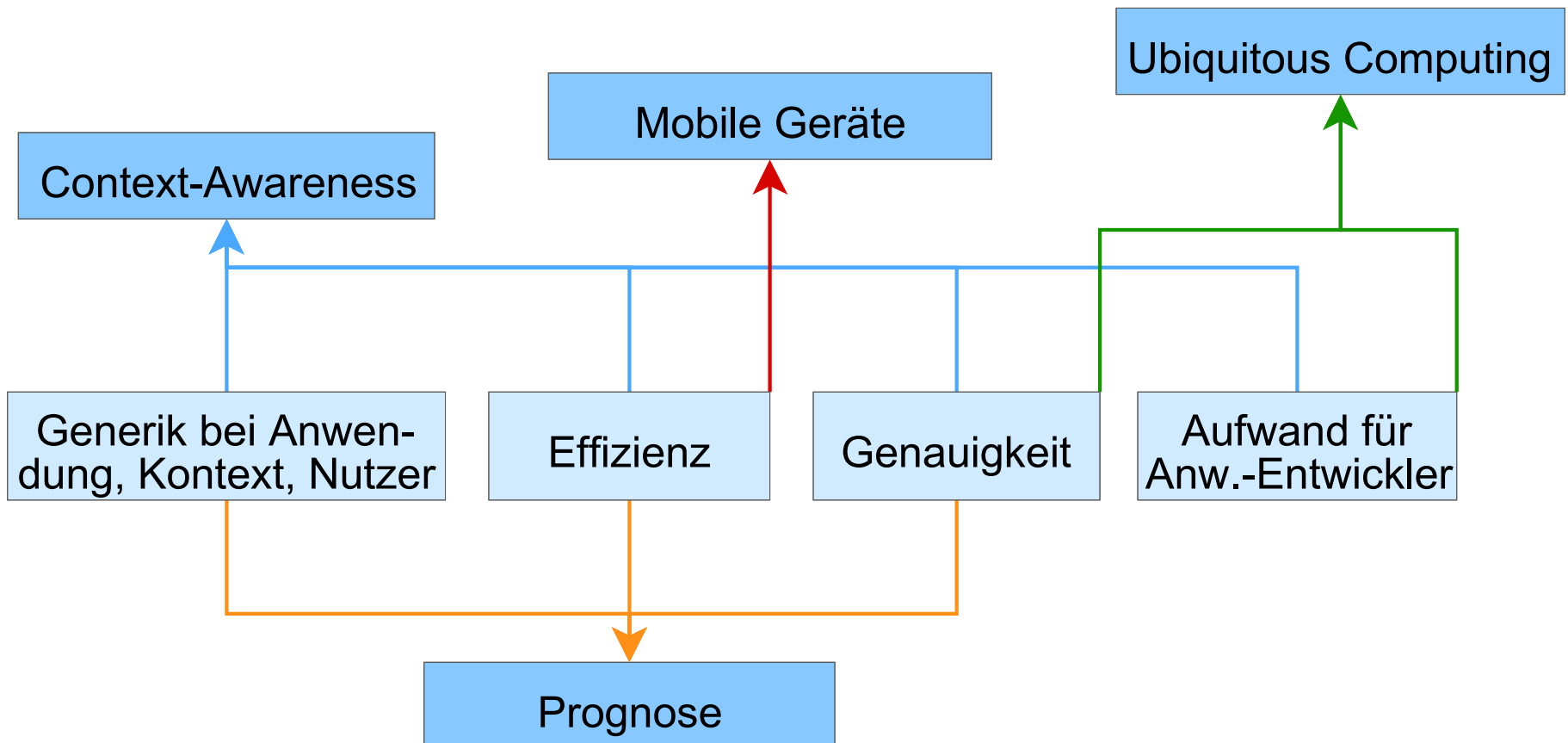
- ◆ Vorwissen (= Abbild von Zusammenhängen) als Voraussetzung für Prognose kann durch Lernen gewonnen werden:



- ◆ (Induktives) Lernen:
 - ◆ ≈ Verallgemeinerung von Einzelfällen (Trainingsdaten) zu Vorwissen für Prognosen
 - ◆ Extraktion des Vorwissens aus den historischen Daten
 - ◆ Hauptnutzen von Lernen: Anpassung an individuellen Nutzer zur Laufzeit

Anforderungen

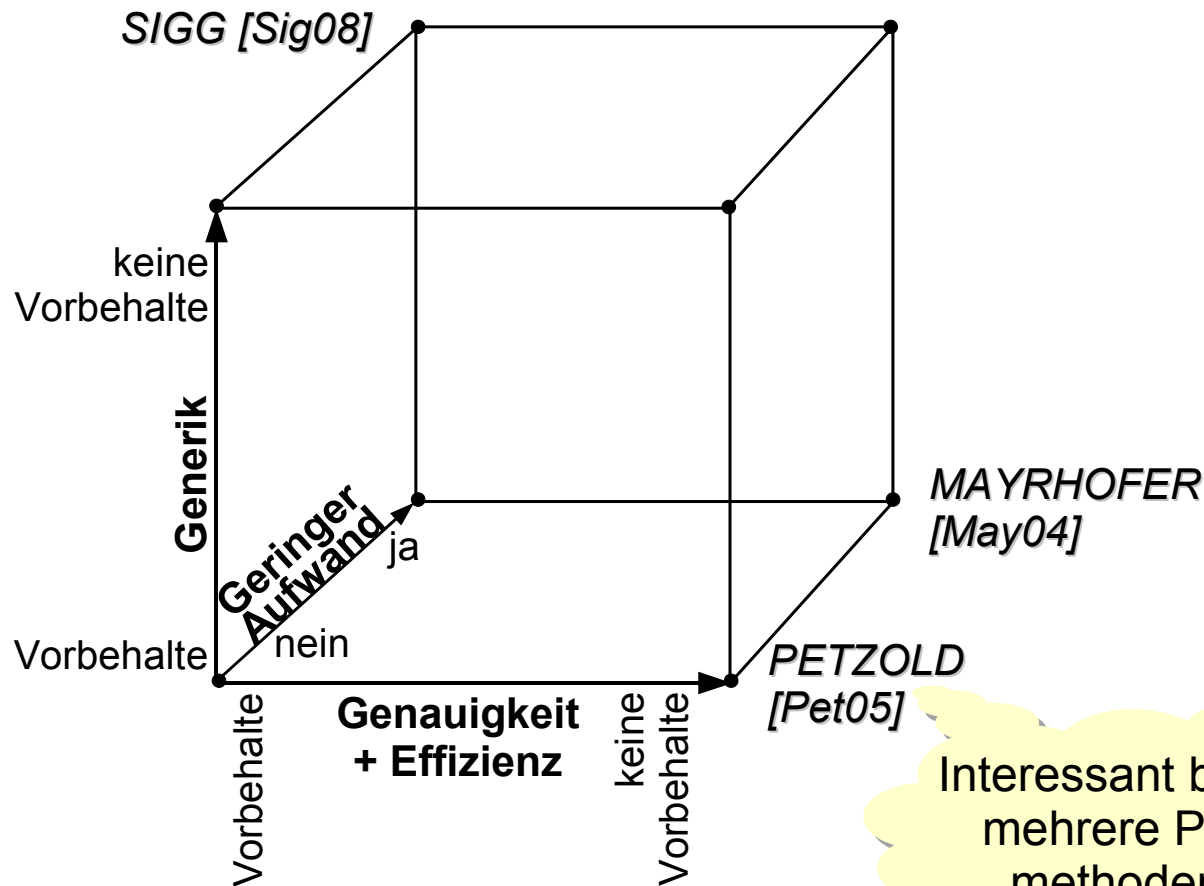
- ◆ Zusammengefasste Anforderungsgruppen (hellblau) und Abhängigkeiten von den vier Bereichen des Problemfeldes:



- ◆ Problemfeld
- ◆ **Bestehende Lösungsmöglichkeiten**
 - ◆ Verfahren zur Kontextdatenprognose
 - ◆ Hybridität und Prognosemethoden
- ◆ Verfahren der strukturierten Kontextdatenprognose
- ◆ Untersuchung der Ergebnisse und Demo
- ◆ Fazit und Ausblick

Bestehende Verfahren zur Kontextdatenprognose

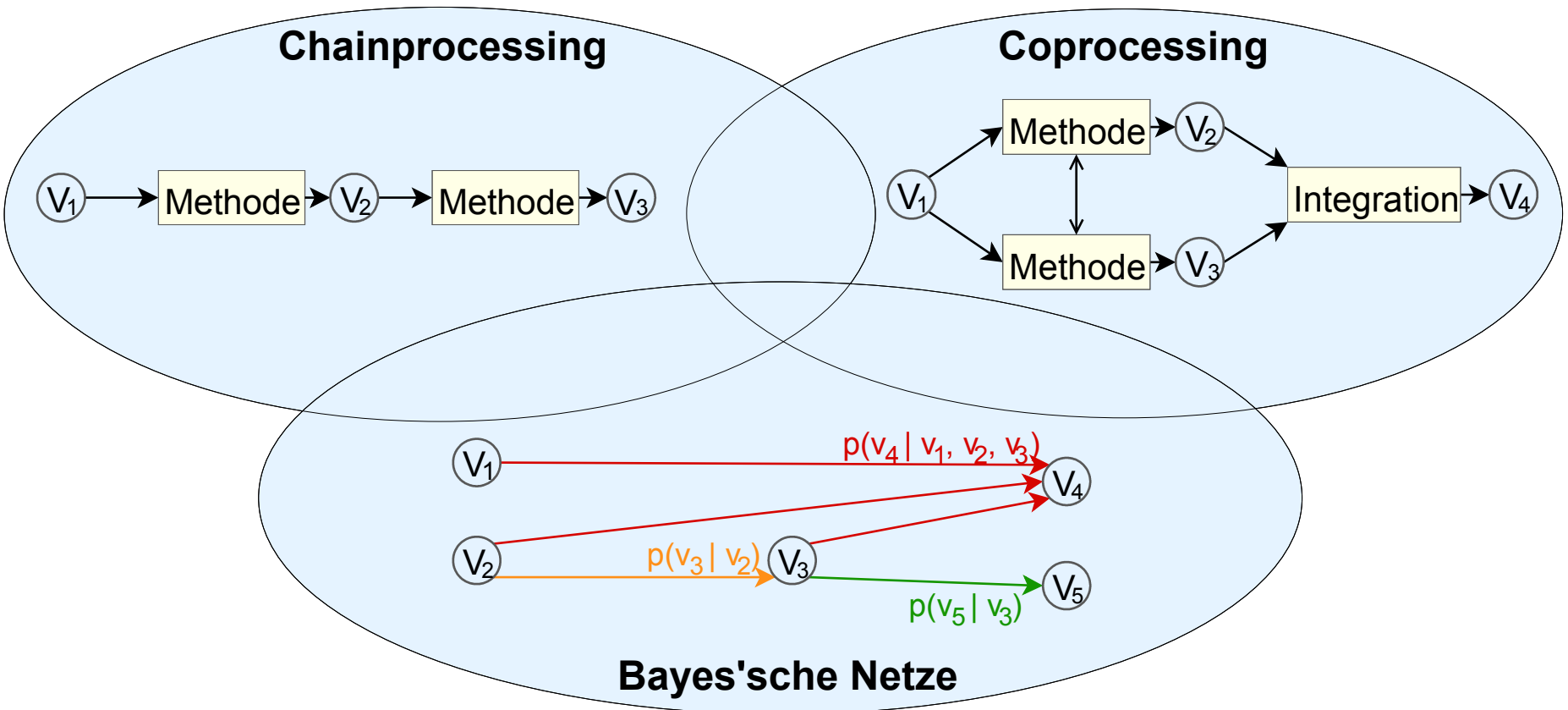
- ◆ Drei Verfahren im Hinblick auf obige Anforderungen analysiert:



Interessant bei Petzold:
mehrere Prognose-
methoden hybrid

- ◆ Offene Aufgabe: Genauigkeit + Effizienz trotz hoher Generik

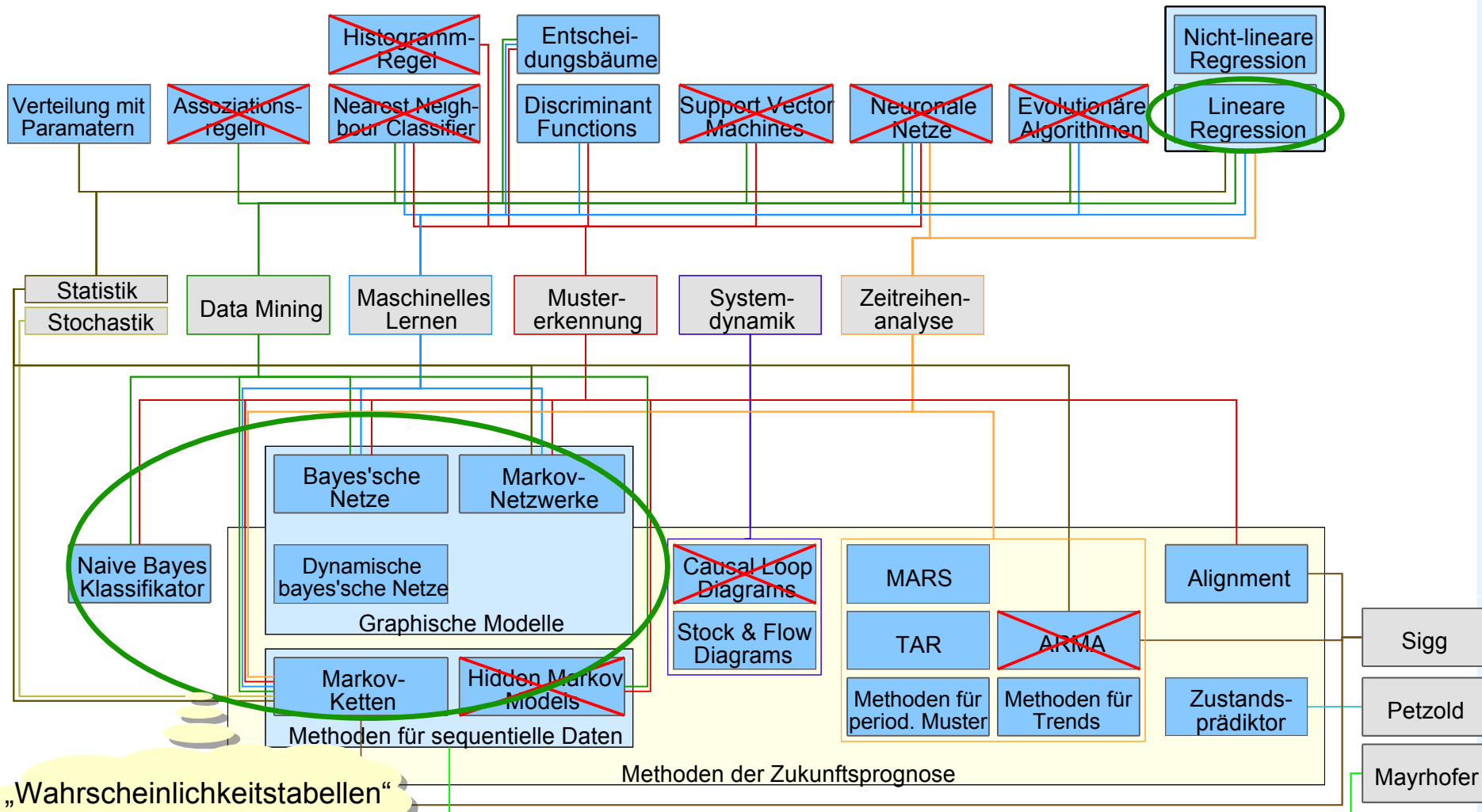
- ◆ Verbindung mehrerer Konzepte für Hybridität:



- ◆ Sicht auf bayes'sche Netze: Zuordnung von Methoden zu Variablen für Prognose der Variablenwerte

Prognosemethoden für hybriden Einsatz

- ◆ Auswahl sich ergänzender Methoden als Basissatz von Methoden



„Wahrscheinlichkeitstabellen“
als Methode abgeleitet

Kontextdatenprognose auf mobilen Geräten

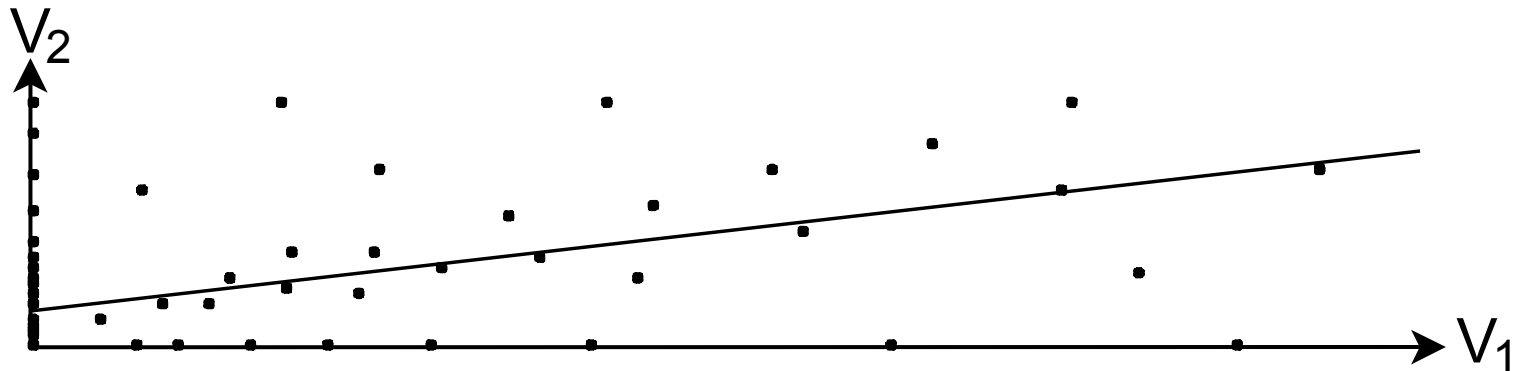
Folie 17: Bestehende Lösungsmöglichkeiten → Prognosemethoden für hybriden Einsatz

Prognosemethoden für hybriden Einsatz

- ◆ Wahrscheinlichkeitstabellen:
 - ◆ Anzahlen des Auftretens von Variablenwertkombinationen als Vorwissen

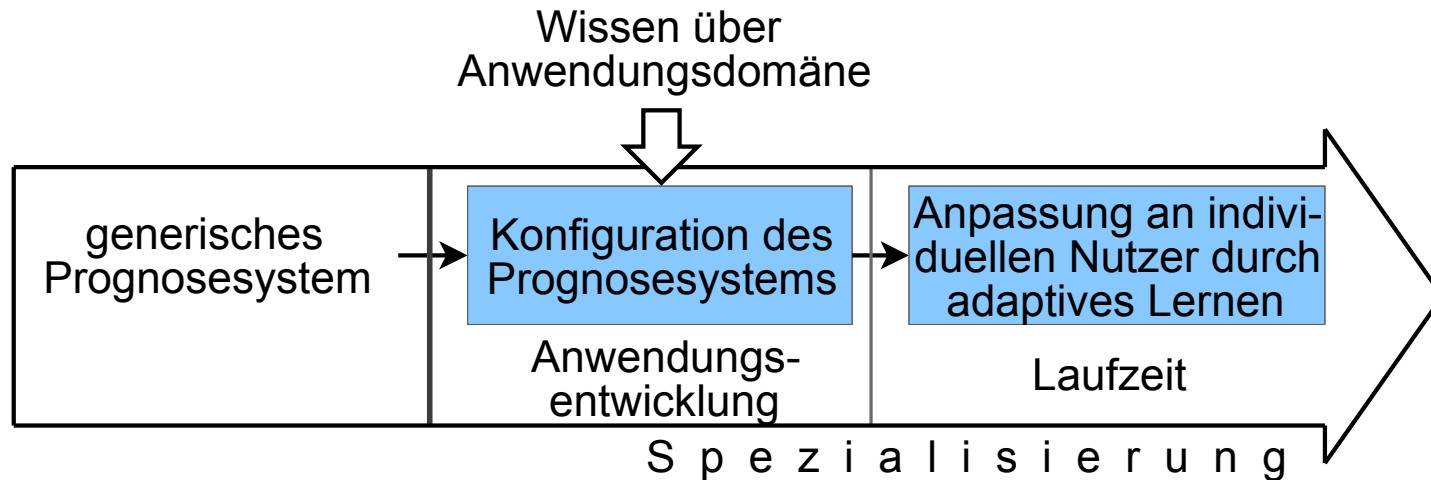
	$V_2 = \dots$	$V_2 = \dots$	$V_2 = \dots$
$V_1 = \dots$	9	24	80
$V_1 = \dots$	86	6	21
$V_1 = \dots$	92	16	4

- ◆ Lineare Regression:
 - ◆ Parameter a und b in $V_2 = a + b \cdot V_1$ als Vorwissen



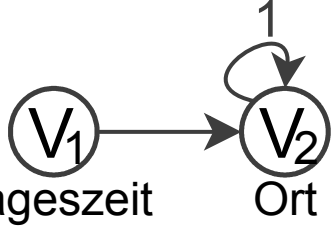
- ◆ Problemfeld
- ◆ Bestehende Lösungsmöglichkeiten
- ◆ **Verfahren der strukturierten Kontextdatenprognose**
 - ◆ Grundprinzipien
 - ◆ Prognosenetze
 - ◆ Prognose und Lernen
 - ◆ Architektur
 - ◆ Erweiterungen
- ◆ Untersuchung der Ergebnisse und Demo
- ◆ Fazit und Ausblick

- ◆ Generelle Methodik:

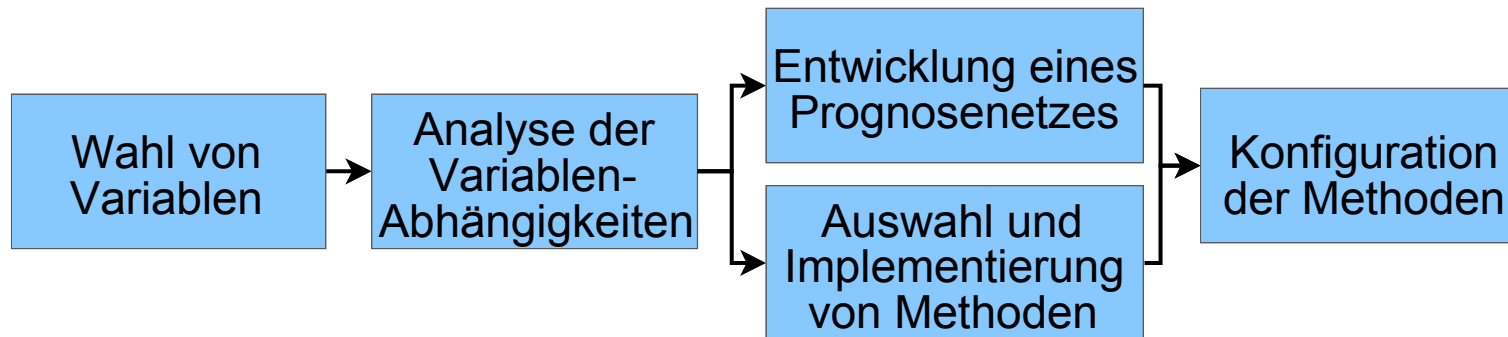


- ◆ Hybrider Einsatz mehrerer, austauschbarer Methoden
- ◆ Zusammenstellung von Methoden als Kern eines anwendungsspezifischen Prognosemodells
- ◆ Wahrscheinlichkeitstabellen und Regression als Hauptteil eines Basissatzes von Methoden zur Auswahl für Anwendungsentwickler
- ◆ Implementierbarkeit neuer Methoden durch Anwendungsentwickler, Prognosesystem als Framework

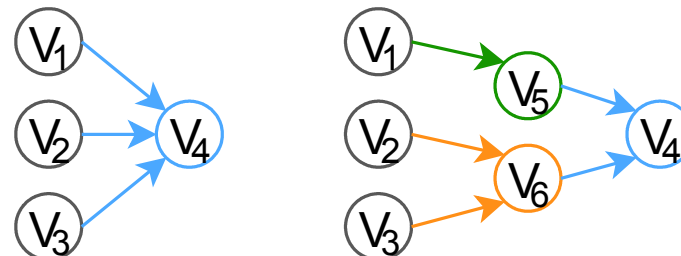
Prognosenetze

- ◆ Prognosenetz als Kern des Prognosemodells
- ◆ Ausgangspunkt für Zusammenstellung von Methoden, die den Variablen zur Prognose ihrer Werte zuzuordnen sind
- ◆ z.B. Prognose des Ortes aus Ort zuvor und Tageszeit: 
- ◆ Definiert als endlicher gerichteter Graph $N = (W, E)$ mit:
 - ◆ Knotenmenge W : Menge von Variablen $\{V_1, \dots, V_n\}$
 - ◆ Eine Kante $V_i \xrightarrow{\Delta} V_{i'}$ aus der Kantenmenge E mit $\Delta \geq 0$ drückt aus, dass bei Prognose des Wertes von $V_{i'}$ zum Zeitpunkt j der Wert von V_i zum Zeitpunkt $j-\Delta$ einbezogen wird
 - ◆ Keine Zyklen der Art $V_i \xrightarrow{\Delta_1} \dots \xrightarrow{\Delta_k} V_i$ mit $\sum_{l=1}^k \Delta_l = 0$
 - ◆ Abkürzende Schreibweisen: insbesondere $V_i \rightarrow V_{i'} := V_i \xrightarrow{0} V_{i'}$

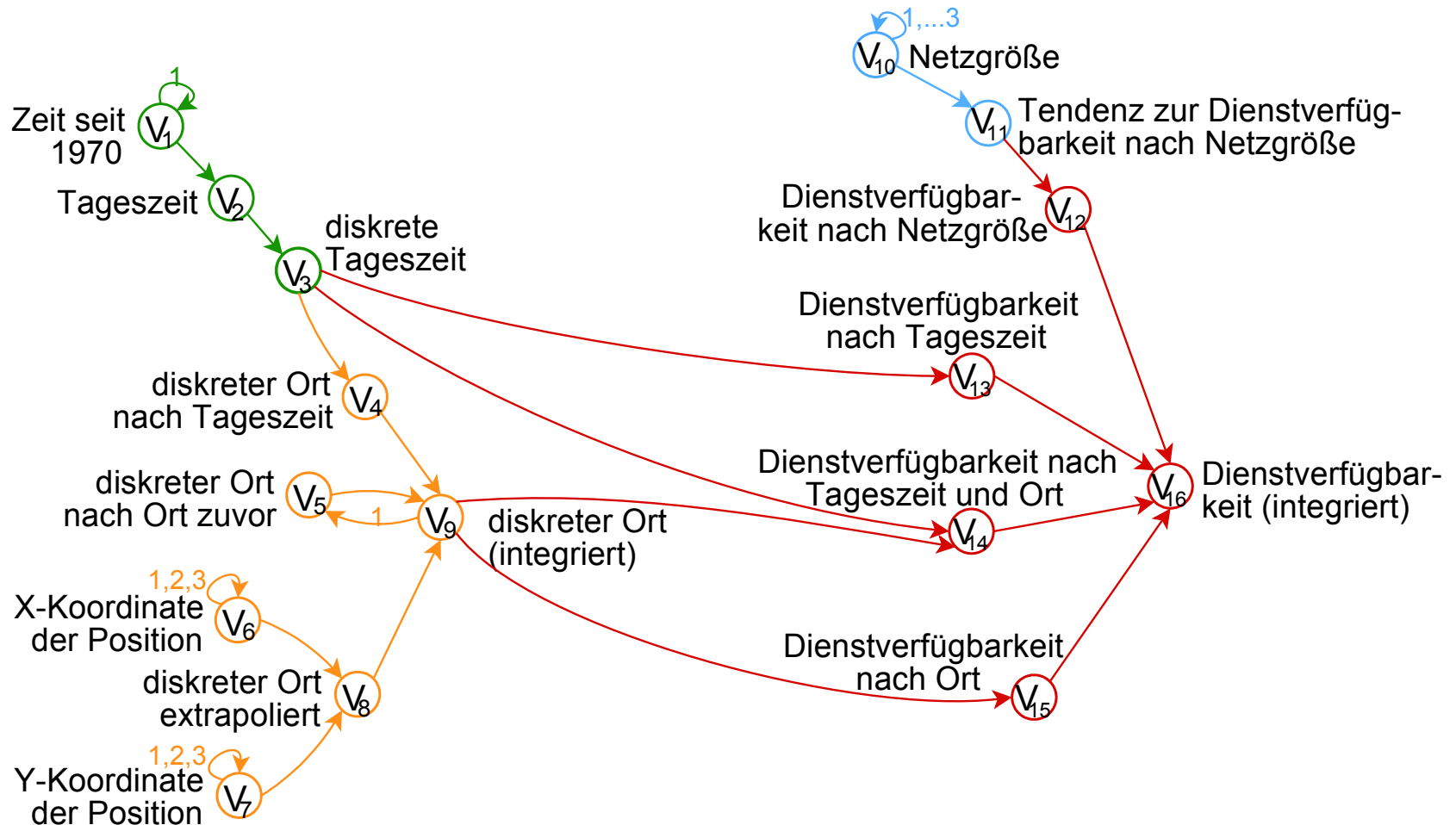
- ◆ Systematisches Vorgehen zur Entwicklung von Prognosemodellen:



- ◆ Netzmuster entwickelt für Coprocessing, Ereignisse, Flüsse und Relationen
- ◆ Netzmuster für Coprocessing (rechts) und entsprechendes Netzfragment ohne Coprocessing (links):

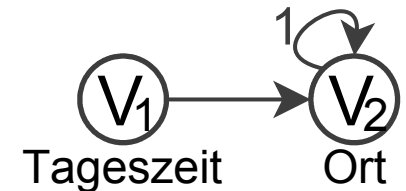


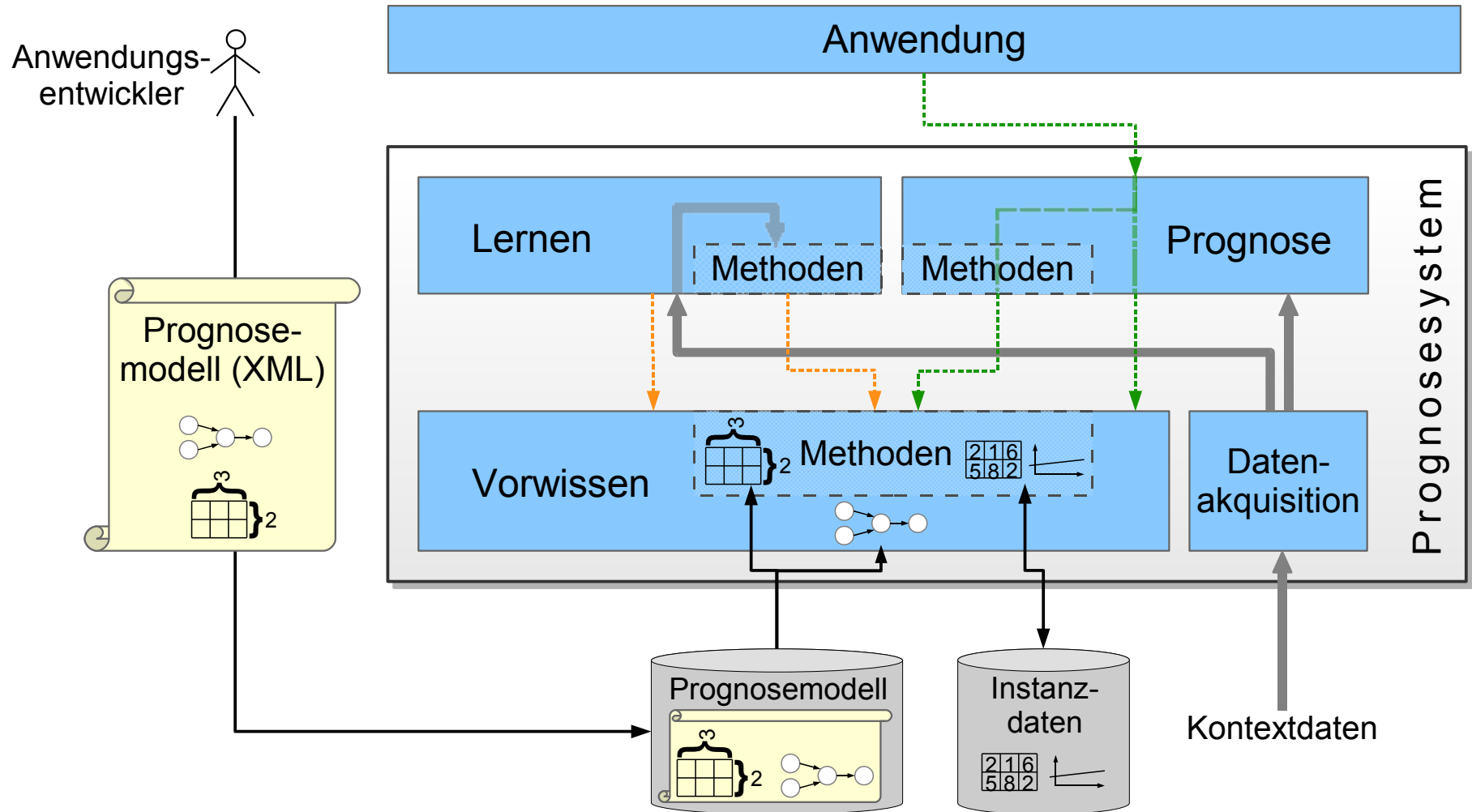
- ◆ Prognosenetz für Prognose der Dienstverfügbarkeit (vgl. Beispiel 2):



Prognose und Lernen

- ◆ Lernen:
 - ◆ Separat pro Methode
 - ◆ Jede Methode lernt aus Trainingsdaten über ihre Variable und die Elternvariablen, z.B. die von V_2 : $(v_{1,j}, v_{2,j-1}, v_{2,j})$
- ◆ Prognose:
 - ◆ Koordination der Methoden
 - ◆ Rekursiver Ablauf nach Eingangskanten, z.B. für Prognose von $v_{2,0}$:
 1. Prognose von $v_{1,0}$
 2. Prognose von $v_{2,-1}$
 3. Anwendung der Methode von V_2 auf $v_{1,0}$ und $v_{2,-1}$
 - ◆ Potentiell mehrere Prognoserunden simulieren mehrere in der Realität mögliche Fälle, daher Wahrscheinlichkeiten als Ausdruck von Unsicherheit
 - ◆ Wahrscheinlichkeiten für Variablenwerte oder für Zeitpunkte, zu denen ein Variablenwert eintritt, als gewünschtes Prognoseergebnis

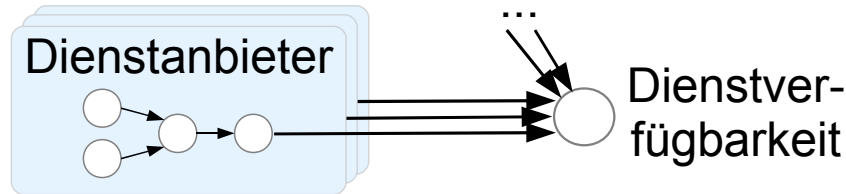




Erweiterungen des Basissystems

- ◆ Netzfragmentinstanzen:

- ◆ Abbildung der Zusammenhänge mehrerer Instanzen der Realität, z.B.



- ◆ Drei Varianten von Verteilung:

- ◆ Entfernte Nutzung von Prognosen
- ◆ Entfernte Datenakquisition
- ◆ Austausch von Vorwissen

- ◆ Unsicherheit:

- ◆ Verbesserte Angabe von Unsicherheit für Anwendung
- ◆ Bevorzugung von Methoden mit sichereren Ergebnissen beim Coprocessing

- ◆ Problemfeld
- ◆ Bestehende Lösungsmöglichkeiten
- ◆ Verfahren der strukturierten Kontextdatenprognose
- ◆ **Untersuchung der Ergebnisse und Demo**
 - ◆ Realisierbarkeit
 - ◆ Messungen zu Genauigkeit und Effizienz
 - ◆ Diskussion
 - ◆ Demo
- ◆ Fazit und Ausblick

- ◆ Prototypische Implementierung (ohne Erweiterungen des Basissystems)
- ◆ Erprobung anhand von Beispiel 1 und Beispiel 2 aus Perspektive der Anwendungsentwickler:
 - ◆ Erstellung von Prognosemodell:

```
<generalSettings ... />  
<nominalTypeSet> ... </nominalTypeSet>  
<measureAdapterSet> ... </measureAdapterSet>  
<methodPreknowledgeSet> ... </methodPreknowledgeSet>  
<variableSet> ... </variableSet>
```

- ◆ Abfragen von Prognosen, z.B.:

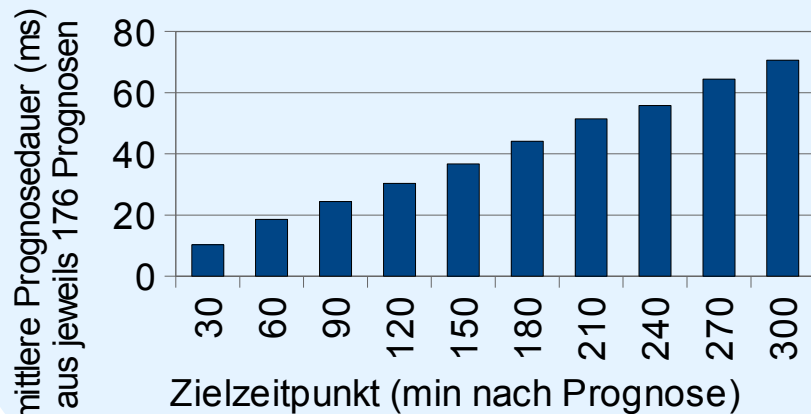
```
predictedVarValueDistribution = predictionSystem.predictState(  
    variableName, targetPointOfTime, iterationCount);
```

- ◆ Einsatz in der DEMAC-Middleware

Messungen zu Genauigkeit und Effizienz

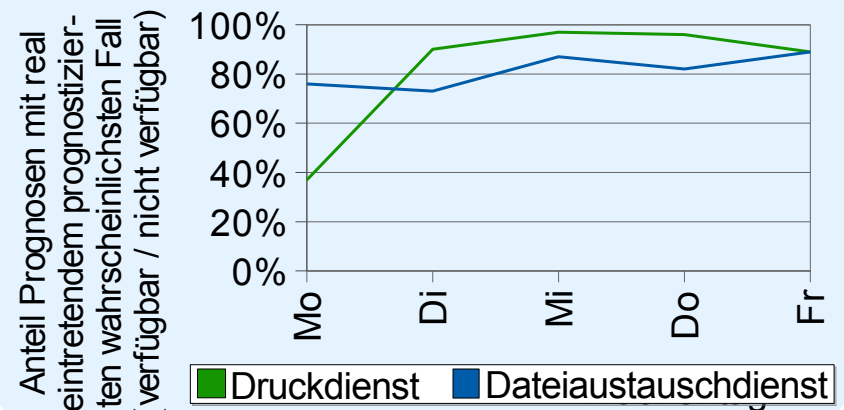
- ◆ Auf Beispiel 2 (Dienstverfügbarkeit) und obigem Netz basierendes Szenario:
 - ◆ Fiktive, realistische Kontextdaten: Nutzer mit mobilem Gerät, Mo-Fr
 - ◆ Druckdienst: auf der Arbeit verfügbar, Dateiaustauschdienst: selten, bei Anwesenheit von z.B. Smartphones verfügbar
- ◆ Erlerntes Vorwissen für Prognose benötigt maximal ca. 20 KB pro Dienst
- ◆ Rechenaufwand zum Lernen minimal (< 1% CPU-Last)
- ◆ Prognosedauer meist < 100 ms (Pentium M - CPU, 1,5 GHz)
- ◆ Prognosedauer:

Prognose der Dienstverfügbarkeit zu einem Zielzeitpunkt, 70 Prognoserunden

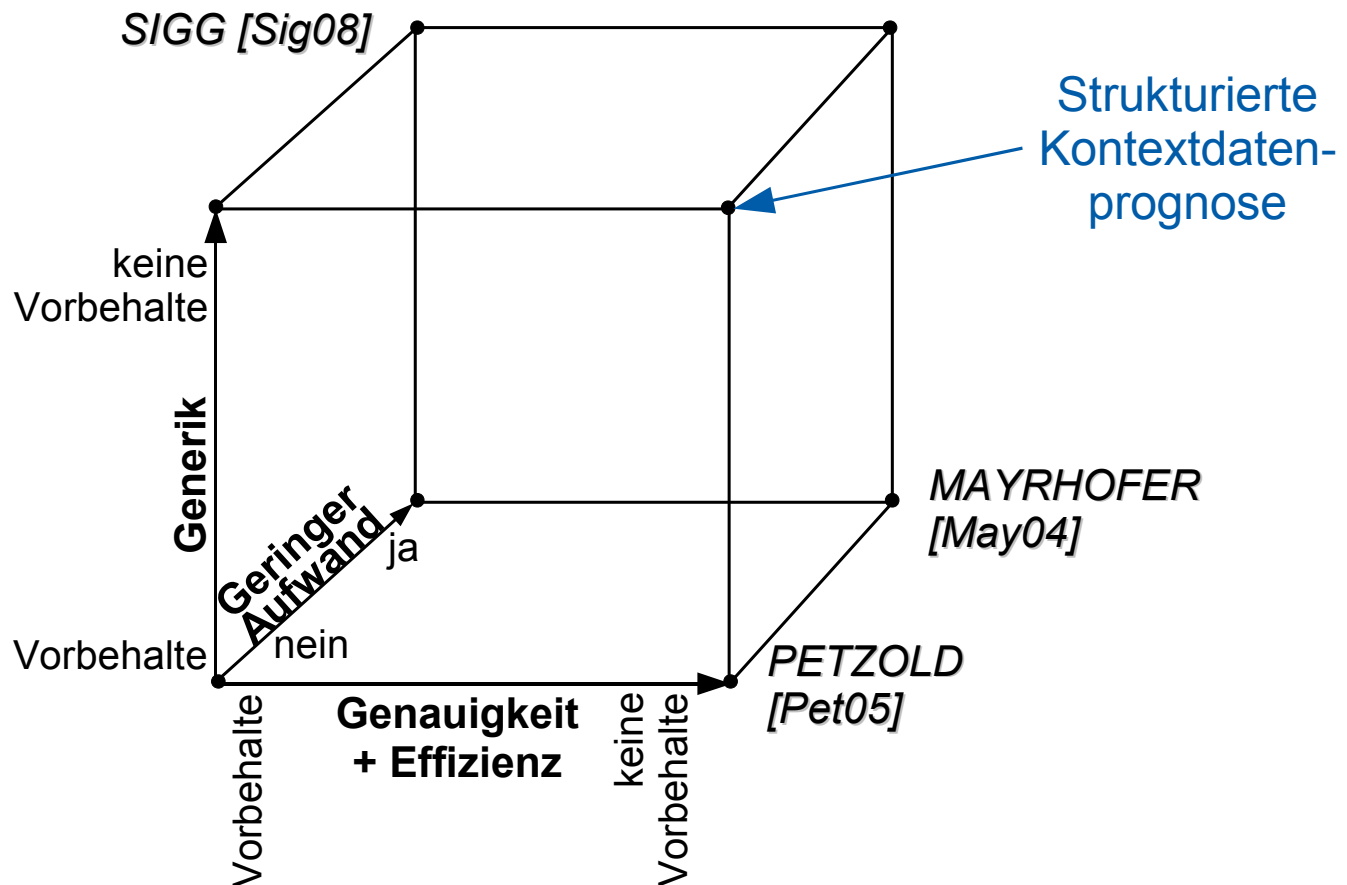


- ◆ Genauigkeit:

Prognose der Dienstverfügbarkeit 180 Minuten nach der Prognose



- ◆ Potential für hohe Genauigkeit + Effizienz trotz Generik, Skalierbarkeit
- ◆ Aufwand für Anwendungsentwickler in Kauf genommen



DEMO

- ◆ Prognosesystem als wiederverwendbare Komponente für Anwendungen entwickelt
- ◆ Hohe Generik
- ◆ Trotzdem Potential für hohe Genauigkeit und Effizienz, dadurch auch Unterstützung mobiler Geräte
- ◆ Keine Out-Of-The-Box-Lösung, sondern Rahmen für anwendungsspezifische Ausgestaltung
- ◆ Viele Möglichkeiten für weitere Verbesserungen (z.B. Umsetzung von Erweiterungen des Basissystems)
- ◆ Offener großer Schritt: Verringerung des Aufwands für Anwendungsentwickler (z.B. durch Toolunterstützung)

Diskussion

- [Bis06] BISHOP, Christopher M.: Pattern Recognition and Machine Learning. Springer, 2006 (Information science and statistics). – ISBN 0–387–31073–8
- [Bra07] BRAMER, Max: Principles of Data Mining. Springer, 2007 (Undergraduate Topics in Computer Science). – ISBN 1–84628–765–0
- [DA99] DEY, A. K. ; ABOWD, G. D.: Towards a Better Understanding of Context and Context-Awareness / Georgia Institute of Technology. Version: 1999. <http://hdl.handle.net/1853/3389>. Atlanta, GA, USA : Georgia Institute of Technology, 1999 (GIT-GVU-99-22). – Forschungsbericht. – veröffentlicht auch in Proceedings of the CHI'00 workshop on 'situated interaction in ubiquitous computing' (CHI'00)
- [Fer07] Kapitel Pervasive Computing - connected > aware > smart. In: FERSCHA, Alois: Die Informatisierung des Alltags. Springer-Verlag, 2007. – ISBN 978–3–540–71454–5, S. 3–10
- [Hüb03] HÜBNER, Gerhard: Stochastik - Eine anwendungsorientierte Einführung für Informatiker, Ingenieure und Mathematiker. 4. Braunschweig : Friedr. Vieweg & Sohn, 2003. – ISBN 3–528–35443–7
- [Kri99] KRIEBEL, Stefan K. T.: A Combined Parametric and Nonparametric Approach To Time Series Analysis. Akad. Verl.-Ges., 1999 (Dissertationen zur Künstlichen Intelligenz (DISKI) 222). – ISBN 3–89601–222–3
- [KS97] KANTZ, Holger ; SCHREIBER, Thomas: Nonlinear time series analysis. Cambridge University Press, 1997 (Cambridge Nonlinear Science Series 7). – ISBN 0–521–55144–7

- [Kun08] KUNZE, C. P.: Kontextbasierte Kooperation: Unterstützung verteilter Prozesse im Mobile Computing, Universität Hamburg, Fachbereich Informatik, Verteilte Systeme und Informationssysteme, Diss., 2008
- [Mat03] Kapitel Vom Verschwinden des Computers - Die Vision des Ubiquitous Computing. In: MATTERN, Friedemann: Vom Verschwinden des Computers. Springer-Verlag, 2003, S. 1–41
- [May04] MAYRHOFER, R.: An Architecture for Context Prediction, Johannes Kepler Universität Linz, Diss., 2004.
<http://www.mayrhofer.eu.org/downloads/publications/PhD-ContextPrediction-2004.pdf>
- [Pet05] PETZOLD, Jan: Zustandsprädiktoren zur Kontextvorhersage in ubiquitären Systemen. Universitätsstr. 22, 86159 Augsburg, Universität Augsburg, Diss., 2005.
http://opus.bibliothek.uni-augsburg.de/volltexte/2005/155/pdf/dissertation_petzold.pdf
- [Rot02] ROTH, Jörg: Mobile Computing - Grundlagen, Technik, Konzepte. 1. Heidelberg : dpunkt.verlag, 2002. – ISBN 3–89864–165–1
- [Sat96] SATYANARAYANAN, M.: Fundamental challenges in mobile computing. In: PODC '96: Proceedings of the fifteenth annual ACM symposium on Principles of distributed computing. New York, NY, USA : ACM, 1996. – ISBN 0–89791–800–2, S. 1–7
- [Sat01] SATYANARAYANAN, M.: Pervasive computing: vision and challenges. In: IEEE Personal Communications 8 (2001), 8, Nr. 4, 10–17. <http://dx.doi.org/10.1109/98.943998>. – ISSN 1070–9916

- [Sch07] Kapitel Eingebettete Interaktion - Symbiose von Mensch und Information. In: SCHMIDT, Albrecht: Die Informatisierung des Alltags. Springer-Verlag, 2007. – ISBN 978–3–540–71454–5, S. 77–101
- [Sig08] SIGG, S.: Development of a novel context prediction algorithm and analysis of context prediction schemes, Universität Kassel, Diss., 2008.
<http://www.upress.uni-kassel.de/online/frei/978-3-89958-392-2.volltext.frei.pdf>
- [Tur06] TURJALEI, Mirwais: Integration von Context-Awareness in eine Middleware für mobile Systeme, Universität Hamburg, Fachbereich Informatik, Verteilte Systeme und Informationssysteme, Diplomarbeit, 2006.
<http://vsis-www.informatik.uni-hamburg.de/getDoc.php/thesis/407/TurjaleiDiplomarbeit.pdf>
- [VGS05] VOVK, Vladimir ; GAMMERMANN, Alexander ; SHAFER, Glenn: Algorithmic Learning in a Random World. Springer, 2005. – ISBN 0–387–00152–2
- [Wit02] WITTIG, F.: Maschinelles Lernen Bayes'scher Netze für benutzeradaptive Systeme, Universität Saarbrücken, Diss., 2002.
http://deposit.d-nb.de/cgi-bin/dokserv?idn=972323384&dok_var=d1&dok_ext=pdf&filename=972323384.pdf