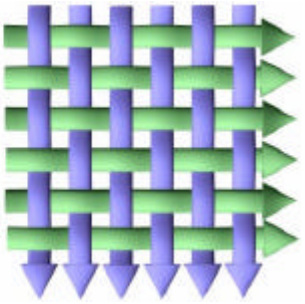
17. November 2003

Sonderforschungsbereich 559
Modellierung großer Netze in der Logistik



Technical Report 03004
ISSN 1612-1376

Standardisierte Beschreibung von Eingangsdaten für die Simulation auf Basis des Prozesskettenparadigmas

Teilprojekt M9:

Jochen Bernhard,
Thomas Fender,
Kay Hömberg,
Dirk Jodin,
Sigrid Wenzel

Fraunhofer IML
Joseph-von-Fraunhofer-Strasse 2-4

44227 Dortmund

Lehrstuhl Mathematische Statistik und industrielle Anwendungen
Vogelpothsweg 87

44221 Dortmund

Lehrstuhl für Förder- und Lagerwesen
Emil-Figge-Strasse 73

44227 Dortmund

Teilprojekt M1:

Falko Bause,
Marcus Völker

Lehrstuhl Informatik IV
August-Schmidt-Strasse 12

44221 Dortmund

Dortmund, den 17. November 2003

Inhalt

1 Einleitung

Für eine zielgerichtete Bestimmung und Aufbereitung der für die Modellierung von GNL benötigten Eingangsdaten ist eine systematische und standardisierte Klassifizierung, Beschreibung und Bewertung dieser Daten wünschenswert. In der Zusammenarbeit der Teilprojekte M1 „Simulation“ und M9 „Informationsgewinnung“ hat sich herauskristallisiert, dass eine standardisierte Form für eine zielführende Nutzung des Prozesskettenparadigmas und der darauf beruhenden Modellierungsmethodik ProC/B aus Anwendersicht sogar zwingend notwendig ist. Die dort für eine Modellbeschreibung notwendigen und typischen Eingangsdaten zur Beschreibung von Logistiknetzen und deren Darstellungsformen werden in diesem Bericht erläutert, wobei dabei das benötigten Qualitäts-, Quantitäts- und Granularitätsniveau (z.B. Aktualität, Genauigkeit, Betrachtungszeitraum etc.) besondere Berücksichtigung findet. Die Modellierung von GNL nach dem Prozesskettenparadigma ermöglicht dann eine Aggregierung aller benötigten Eingangsdaten auf dem notwendigen Niveau (siehe Kapitel 2). Dazu wird eine prozessunabhängige und standardisierte Beschreibung sowie eine Klassifizierung der systembeschreibenden Daten erarbeitet. In diesem Zusammenhang konnte auf eine Klassifizierung aus dem Bereich der Simulation von Logistik-, Materialfluss- und Produktionssystemen [VDI00] zurückgegriffen werden, die eine Dreiteilung der Daten in die Kategorie der Systemlastdaten, der Organisationsdaten und der Technischen Daten vorsieht. Dieser Ansatz kann für die Modellbildung von GNL übernommen werden, wobei die konkreten Ausprägungen der Eingangsdaten in der Regel auf eine andere Betrachtungsebene projiziert werden müssen (siehe Kapitel 3).

Die detaillierte Untersuchung der Eingangsdaten fokussiert sich in dieser Arbeit auf die Systemlast, da diese systembeschreibende Datenklasse den höchsten Komplexitätsgrad aufweist und somit einen hohen Bedarf an Datenanalyse und -aufbereitung notwendig macht. Anhand eines Anwendungsfalls aus Teilprojekt A5 wird die benötigte Systemlast zielorientiert ermittelt und die zugehörigen Daten mit geeigneten Datenanalyseverfahren extrahiert und die gewonnenen Ergebnisse validiert (siehe Kapitel 4). Mit Hilfe dieser experimentell erarbeiteten Ausgangsbasis kann eine standardisierte Beschreibung von Systemlasten zur Modellierung von GNL vorgenommen werden. Die generelle Klassifizierung der Systemlast wird dabei anhand der jeweils vorliegenden Eigenschaften und der zugehörigen Darstellungsform sowie den Abhängigkeiten zwischen diesen Eigenschaften durchgeführt (Kapitel 5.1). Mit Hilfe einer solchen standardisierten Beschreibung ist eine generelle Anleitung zur Systemlastmodellierung mit ProC/B möglich (Kapitel 5.2).

2 Eingangsdaten für ProC/B

2.1 Kurzeinführung in ProC/B

ProC/B ist eine auf dem Prozesskettenparadigma nach Kuhn [Kuh95] beruhende Beschreibungsmöglichkeit (vgl. [BBT+99, BBF+02]), um die Struktur und das Verhalten logistischer Netze zu modellieren. In der Darstellung von ProC/B-Modellen spiegelt sich diese Einteilung wider. In der Abbildung 1 ist eine Funktionseinheit (FE) eines ProC/B-Modells dargestellt. FEs sind mit Abteilungen realer Systeme vergleichbare Einheiten, die nach außen an definierten Schnittstellen Dienste anbieten und für ihre Diensterfüllung andere FEs verwenden können. Der mittlere Teil jeder FE beschreibt dabei das Verhalten der modellierten Real-Einheit und der untere Bereich die Struktur.

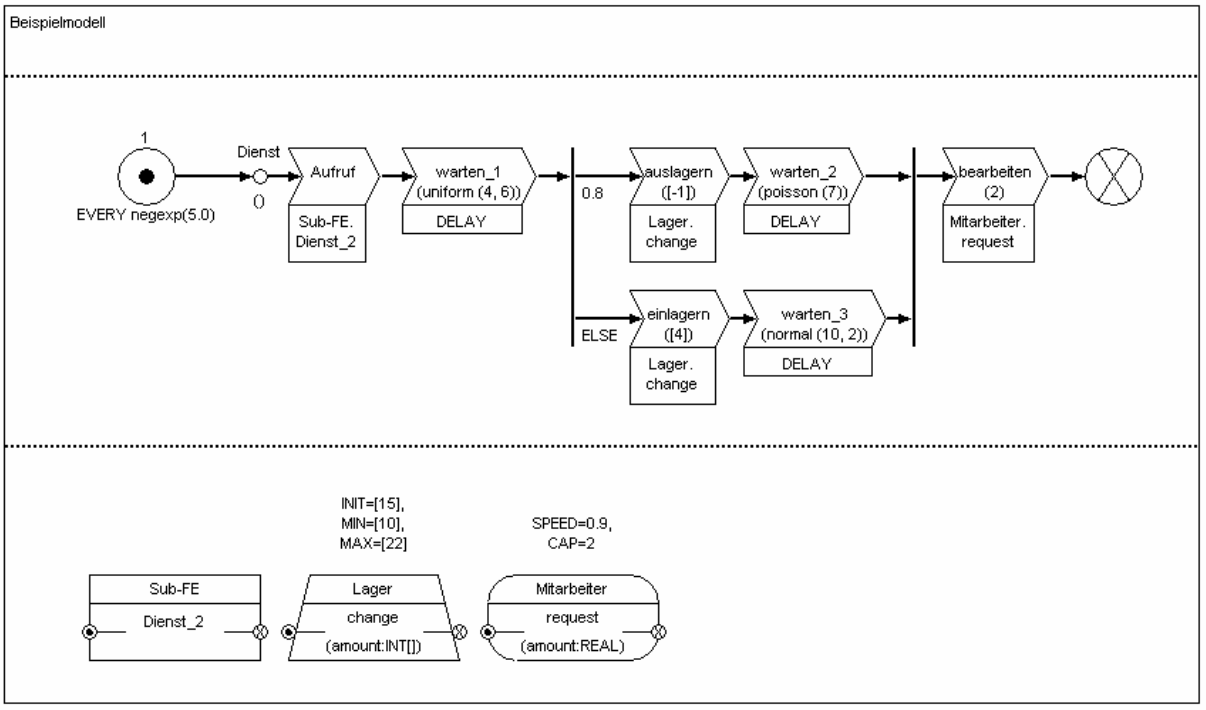


Abbildung 1: Eine Funktionseinheit.

Das Verhalten bestimmt sich durch die Dienstbeschreibungen anhand von Prozesskettenelementen (PKEs), die miteinander verkettet die Reihenfolge der modellierten Aktionen eines Prozesses beschreiben. Neben den PKEs, die Verzögerungen, Dienstaufrufe oder Hi-Slang-Code [BFK+93] enthalten können, sind ferner Schleifen und Konnektoren definiert. Als Konnektoren werden Oder-Konnektoren zur Ermöglichung alternativer Aktivitätsketten sowie Und-Konnektoren für nebenläufige Aktivitätsketten desselben Prozesses angeboten. Ferner sind Prozesskettenkonnektoren (PK-Konnektoren) definiert, die verschiedene Prozessketten einer FE miteinander synchronisieren.

Der Strukturbereich einer FE definiert, welche Sub-Einheiten innerhalb einer FE verfügbar sind. Diese Sub-Einheiten bieten ihrerseits Dienste an, die von der sie enthaltenden FE genutzt werden können. Diese Strukturgestaltung ist vergleichbar mit realen Strukturen, z. B. Firma - Abteilung - Subabteilung - Arbeitsplatz. Durch die Strukturbeschreibung in ProC/B entsteht eine Hierarchie von FEs, die sich baumartig gestalten lässt, jedoch auch mit sog. externen FEs Zugriffe auf andere FEs derselben Hierarchiestufe ermöglicht. Nach unten abgeschlossen wird eine solche Hierarchie durch standardisierte FEs, die den elementaren Zeit- bzw. Platzverbrauch modellieren (Server (siehe FE `Mitarbeiter` in Abbildung 1) bzw.

Counter (siehe FE `Lager` in Abbildung 1)) und nicht weiter verfeinert werden. Nach oben wird das Modell abgeschlossen durch eine Beschreibung der Umwelt. Dazu sind bei einem realen System eingehende und ausgehende Prozesse bzw. Dienstvorgänge zu beobachten, die den durch das Modell beschriebenen Bereich betreten bzw. verlassen. Das Betreten bzw. Verlassen wird in ProC/B durch Quellen bzw. Senken (vgl. Abbildung 1) modelliert. Senken terminieren Prozesse, während Quellen Prozesse erzeugen.

2.2 Beschreibung von Eingangsdaten

2.2.1 Allgemeines

Ein Modell eines logistischen Systems wird erstellt, um bestimmte Kennzahlen ermitteln zu können. Dies muss sich auch in der Auswahl der nachzubildenden Teile realer Systeme widerspiegeln, da ein zu ungenaues Modell verfälschte Ergebnisse liefert und ein zu genaues Modell einen zu hohen Modellierungsaufwand und zu große Simulationszeiten zur Folge hat. Somit sind die zu wählenden Systemgrenzen und der Abstraktionsgrad direkt von der Aufgaben- und Zielstellung abhängig und führen zunächst zu einem konzeptionellen Modell. Das konzeptionelle Modell bestimmt damit maßgeblich die Art und den Umfang der benötigten Eingangsinformation.

2.2.2 Wahrscheinlichkeitsverteilungen

Die durch ProC/B angebotenen Modellelemente verfügen über Parameter, die das Verhalten der Prozesse beeinflussen. Vielfach sind solche Parameter nicht deterministisch sondern unterliegen Verteilungen. ProC/B bietet dazu die in Hi-Slang [BFK+93] definierten Wahrscheinlichkeitsverteilungen an, u. a:

- Cox-Verteilung

  cox (A, B: REAL) → REAL

  definiert eine COX-verteilte Pseudozufallszahl mit Rate A (reziproker Mittelwert) und Variationskoeffizient B (A > 0 und B ≥ 0.1).

- Erlang-Verteilung

  erlang (A, B: REAL) → REAL

  definiert eine Erlang-Verteilung mit Mittelwert 1/A und Standardabweichung $1/(A*\sqrt{B})$. Sowohl die Rate A als auch die Phasenanzahl B müssen größer als 0 sein.

- Linear-Verteilung

  linear (A, B: ARRAY OF REAL) → REAL

  A und B müssen eindimensionale REAL-ARRAYs sein mit gleichen Unter- und Obergrenzen lb und ub. Sie definieren eine kumulative Verteilung f. Das Ergebnis wird durch lineare Interpolation in der nicht-äquidistanten Verteilungstabelle ermittelt und ist definiert durch A und B.

  Die Funktion f ist definiert durch A[i] = f(B[i]) für lb = i = ub.

  Es müssen folgende Bedingungen eingehalten werden:

  A[lb] = 0, A[ub] = 1 und beide Arrays müssen monoton sein, z. B. für lb = i = ub gilt A[i] = A[i+1] und B[i] = B[i+1].

- Negativ-exponentiale Verteilung

  negexp (A: REAL) → REAL

definiert eine negativ-exponentiale Verteilung mit Mittelwert 1/A (bzw. Rate A). Dies entspricht der Wartezeit zwischen zwei Ankünften eines Poisson-verteilten Ankunftsprozesses mit Erwartungswert A für die Anzahl der Ankünfte je Zeiteinheit. A muss größer als 0 sein.

Zur Verwendung dieser Verteilung siehe auch die Quelle in Abbildung 1, die im Mittel alle 1/5 Zeiteinheiten einen Prozess erzeugt.

- Normal-Verteilung

  normal (A, B: REAL) → REAL

  definiert einer Normalverteilung mit Mittelwert A und Standardabweichung B (mit $B \geq 0$).

  Zur Verwendung dieser Verteilung siehe auch Abbildung 1 (PKE `warten_3`).

- Poisson-Verteilung

  poisson (A: REAL) → INTEGER

  definiert eine Poisson-Verteilung mit Parameter A. Das Ergebnis *n* wird definiert durch n+1 Basisziehungen u(i), so dass n die kleinste nicht-negative ganze Zahl mit u(0)*u(1)*...*u(n) < exp(-A) ist. Für negative A ist das Ergebnis 0.

  Zur Verwendung dieser Verteilung siehe auch Abbildung 1 (PKE `warten_2`).

- Gleich-Verteilung

  randint (A, B: INTEGER) → INTEGER

  definiert eine Gleich-Verteilung über die ganzen Zahlen A, A+1,..., B-1, B (mit A = B).

  uniform (A, B: REAL) → REAL

  definiert eine Gleich-Verteilung über die reellen Zahlen im Intervall [A, B] (mit A = B).

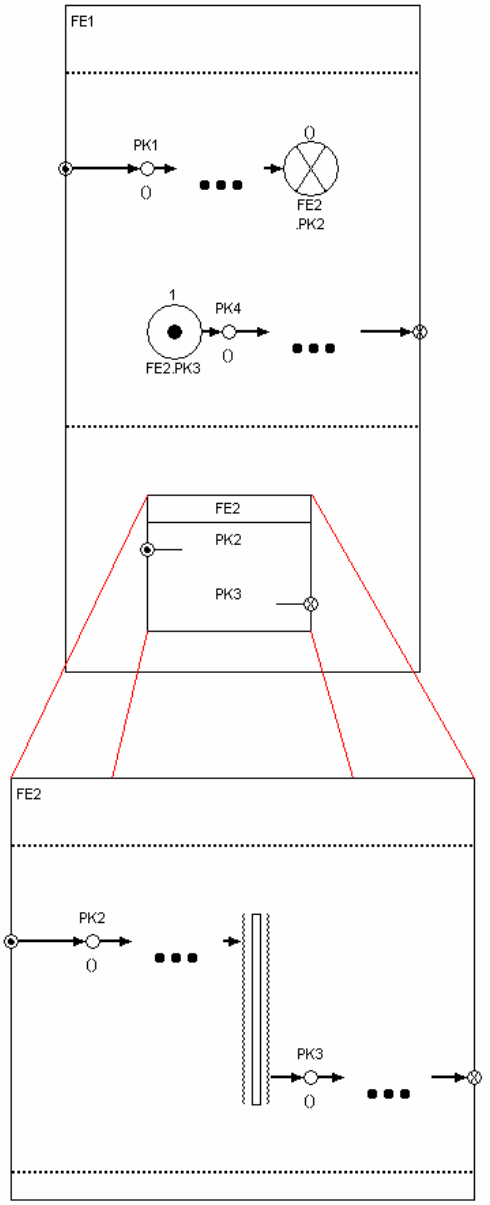
  Zur Verwendung dieser Verteilung siehe auch Abbildung 1 (PKE `warten_1`).

- Bernoulli-Variable

  draw (A: REAL) → BOOLEAN

  Das Ergebnis ist TRUE mit der Wahrscheinlichkeit A und FALSE mit der Wahrscheinlichkeit 1-A.

2.2.3 Bedingte Quellen und Senken sowie PK-Konnektoren



Neben dem in Abbildung 1 verwendeten Senkentyp (unbedingte Senke) verfügt ProC/B über sog. bedingte Senken. Diese ermöglichen es, die Erzeugung von Prozessen vom Terminieren anderer Prozesse abhängig zu gestalten. Dazu verweisen bedingte Senken auf die Dienste der anzustoßenden Quellen. In Abbildung 2 wird nach dem Durchlaufen der PK1 ein Prozess auf der Kette PK2 generiert. Die Zwischenankunftszeit zweier Prozesse der Kette PK2 ist dabei abhängig von der Laufzeit der Prozesse der PK1.

Eine Quellen- und Senkenfunktionalität ist ferner durch die PK-Konnektoren gegeben, da an PK-Konnektoren Prozessketten beginnen und enden können. Auch auf diese Weise ist es möglich, Prozesse zu erzeugen, wenn andere Prozesse bestimmte Zustände (deren Punkt durch den zugehörigen PK-Konnektor gekennzeichnet ist) erreichen. In Abbildung 2 ist die Erzeugung eines Prozesses der Kette PK3 vom Terminieren eines Prozesses der Kette PK2 abhängig.

Prozesse der Kette PK3 wiederum bedingen mit ihrer Terminierung das Erzeugen von Prozessen der Kette PK4 mit Hilfe einer sog. bedingten Quelle. Bedingte Quellen erzeugen dann Prozesse, wenn an der Senke des Dienstes, auf den die Quelle verweist, ein Prozess terminiert.

Durch Verwendung von bedingten Quellen und Senken sowie PK-Konnektoren lassen sich auch Ankunftsverhalten modellieren, die über die Möglichkeiten der oben beschriebenen Wahrscheinlichkeitsverteilungen hinausgehen.

Abbildung 2: Die Verwendung bedingter Quellen und Senken sowie PK-Konnektoren.

2.2.4 Parametrierung der Modellelemente

Die durch Quellen erzeugten Prozesse müssen den beobachtbaren eingehenden Dienstvorgängen des realen Systems bzgl. ihrer Menge und ihrer Eintrittszeitpunkte entsprechen. Um dies zu bewerkstelligen, verfügen Quellen über Parameter zur Beschreibung der zu erzeugenden Systemlast. Dazu gehören die Modellzeitpunkte der Erzeugung bzw. der Abstand zwischen 2 Modellzeitpunkten der Erzeugung (Zwischenankunftszeiten) sowie die Anzahl der Prozesse, die jeweils gemeinsam erzeugt werden sollen. Im Beispielmodell der Abbildung 1 wird durch eine unbedingte Quelle im Mittel alle 1/5 Zeiteinheiten ein Prozess erzeugt mit negativ-exponentiell verteilten Zwischenankunftszeiten.

Delay-PKEs erhalten eine Wartezeitbeschreibung, die evtl. die Dauer mehrerer im Modell nicht näher definierter Aktivitäten dargestellt und aus dem realen System vor der Modellierung ermittelt oder abgeschätzt werden muss. In Abbildung 1 wurden Delay-PKEs mit verschiedenen Verteilungen verwendet. Durch das PKE `warten_1` wird ein Warten von gleichverteilt zwischen 4 und 6 Zeiteinheiten ausgelöst, durch das PKE `warten_2` wird eine Poisson-verteilte Anzahl von Zeiteinheiten (mit Parameter 7) gewartet und durch das PKE `warten_3` eine normalverteilte Anzahl von Zeiteinheiten mit Mittelwert 10 und Standardabweichung 2.

Oder-Konnektoren benötigen Bedingungen zur Verzweigung. Neben durch Variablen und Parameter definierten Boole'schen Bedingungen kommen hierfür auch probabilistische Bedingungen in Frage. In Abbildung 1 wird die obere Aktivitätenfolge in 80% aller Fälle gewählt (Kantenwert `0.8`) und die untere in 20% (Kantenwert `ELSE`).

Server-Elemente müssen bzgl. der durch sie modellierten Anzahl an Bedieneinheiten parametriert werden. In Abbildung 1 definiert am Server `Mitarbeiter` die `CAP 2` somit 2 Mitarbeiter. Ferner ist eine Geschwindigkeit einzustellen, die bspw. die persönliche Verfügungszeit von Mitarbeitern regelt. So hat ein Server bzw. Mitarbeiter mit der Geschwindigkeit 0,9 eine persönliche Verfügungszeit von 10% (vgl. Abbildung 1). Durch den `request`-Aufruf durch das PKE `bearbeiten` wird dieser Server für 2 Zeiteinheiten in Anspruch genommen.

Counter-Elemente müssen hinsichtlich ihrer Mindest- und Maximalkapazität sowie des Startbestandes parametrisiert werden. Bei fehlerhaften Grenzen können die in einer Simulation auftretenden Wartesituationen evtl. nicht mit der Realität übereinstimmen. Der Startbestand kann wesentlichen Einfluss auf die Konvergenzgeschwindigkeit einzelner Messwerte nehmen. In Abbildung 1 ist mit dem dortigen Counter ein `Lager` definiert. Sein Mindestbestand beträgt 10, sein Höchstbestand 22, sein Initialzustand ist 15. Der `change`-Aufruf des `auslagern`-PKEs beschreibt eine Entnahme von einem Element aus dem Lager, während der `change`-Aufruf des `einlagern`-PKEs eine Einlagerung von 4 Elementen modelliert.

3 Eingangsdaten für die Modellierung und Simulation von GNL

Zur Beschreibung von Materialfluss-, Produktion- und Logistiksystemen hat sich in Anlehnung an die VDI 3633, Blatt 1, eine Klassifikation der Eingangsdaten für die Simulation nach

- Systemlastdaten
- Organisationsdaten und
- technischen Daten

durchgesetzt, die auch in der logistischen Praxis Anwendung gefunden haben. Abbildung 3 zeigt die Zusammensetzung der simulationsrelevanten Daten laut VDI 3633, Blatt 1, auf.

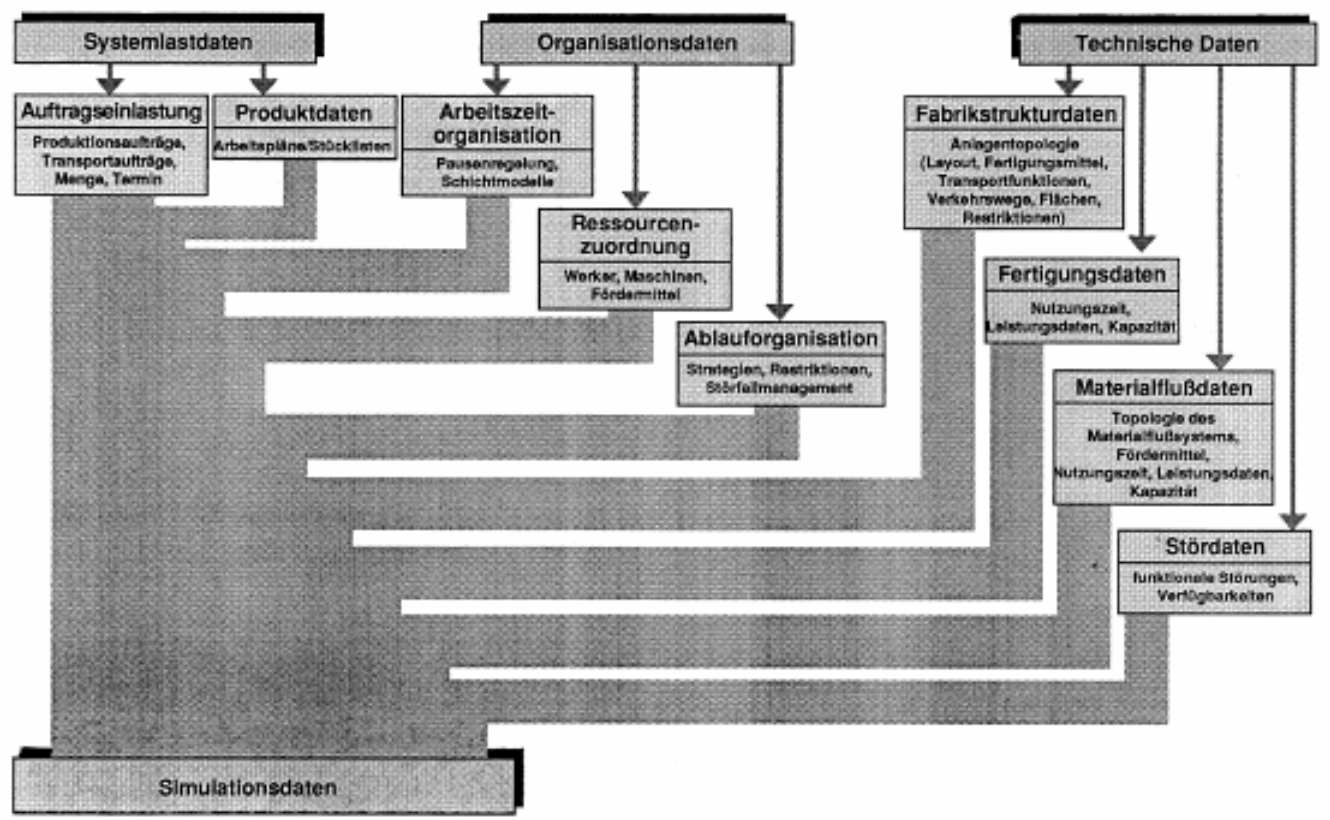


Abbildung 3: Zusammensetzung der simulationsrelevanten Daten [VDI00]

Eine Übertragung dieser Klassifikation auf die Eingangsdaten zur Modellierung von Logistiknetzen ist möglich. Wesentlich ist jedoch, dass sich die konkreten Inhalte entsprechend des neuen Bezugsrahmens *GNL* und des mit ihm verbundenen Abstraktionsgrades verändern:

- Systemlastdaten:
  Die Auftragseinlastung erfolgt über Produktions- und Transportaufträge z.B. auf der Basis von Kundenbestellungen; weniger relevant sind detaillierte Produktdaten wie Arbeitspläne und Stücklisten.
- Organisationsdaten:
  Die Arbeitszeitorganisation gilt in Analogie zu Abbildung 3, unter Arbeitszeiten sind beispielsweise die Betriebszeiten der am Netzwerk beteiligten Unternehmen zu fassen. Die Ressourcenzuordnung bezieht sich maßgeblich auf die unternehmens-übergreifende Disposition logistischer Ressourcen wie Transportmittel und Transportwege sowie am Netzwerk beteiligter Unternehmen. Die Ablauforganisation muss globaler unter dem Aspekt des Supply Chain Managements verstanden werden. Sie behandelt insbesondere

Strategien zur Organisation der Lieferketten oder Liefernetze sowie das Störfallmanagement.

- Technische Daten:
  Die technischen Daten beziehen sich weniger auf die klassischen innerbetrieblichen Fabrikstruktur-, Fertigungs-, Materialfluss- und Störddaten, sondern subsumieren die in den GNL anfallenden Daten zur Beschreibung der notwendigen Ressourcen für die unternehmensübergreifenden Abläufe (z. B. Umschlag- und Transportdaten) sowie die in GNL besonders relevanten Daten zur Informationsflussbeschreibung. So werden die Fabrikstrukturdaten um die Netzstrukturdaten wie die Verkehrswege, die Netztopologie, die Lage von Produktions- und Logistikstandorten usw., ergänzt. Die innerbetrieblichen technischen Leistungsdaten werden jetzt maßgeblich durch die technischen Daten der außerbetrieblichen Ressourcen, z.B. der Transportsysteme, wie Frachtraumkapazitäten von Flugzeugen, Geschwindigkeiten von Güterzügen oder die Containerkapazität von See- und Binnenschiffen ersetzt. Störddaten beziehen sich auf den Ausfall eines Zulieferers, einen gesperrten Transportweg, oder ein ausgefallenes Verkehrsmittel.

Die für die Simulation von Logistik- und Produktionssystemen notwendigen Eingangsdaten variieren erheblich in Umfang und Komplexität (vgl. [VDI00], [WMe93]). Darüber hinaus ist zwischen deterministischen und stochastischen Eingangsgrößen zu unterscheiden. Technische Daten mit Ausnahme von Störddaten beziehen sich häufig auf einzelne, diskrete und direkt aus der Anwendung übertragbare Größen, die je nach Verhaltenskomplexität der Systemkomponente determiniert oder stochastisch sein können (z.B. Lagerkapazität oder manueller Arbeitsplatz).

Organisationsdaten umfassen entweder ebenfalls relativ einfache, deterministische Eingangsgrößen (z.B. Schichtmodell) oder beziehen sich auf algorithmische Strategiebeschreibungen, die nicht direkt als Eingangsdatum verwendet werden können. Hingegen sind für die Übernahme realer Daten zur Beschreibung der Systemlastdaten und der Störddaten umfangreiche Datenbestände des Unternehmens zu analysieren und aufzubereiten. Für die Betrachtung von Logistiknetzen erschwert sich diese Situation noch dahingehend, dass nicht nur *ein* Unternehmen betrachtet wird, sondern innerhalb eines Unternehmensnetzes Daten verschiedener Unternehmen zur Beschreibung der Systemlasten und Störddaten betrachtet werden müssen. Damit ergibt sich insbesondere hinsichtlich der *Systemlast- und Störddaten* eine weitere, enorme Erhöhung des Datenumfangs und der Datenkomplexität.

Störddaten bzw. Störprofile der zu betrachtenden Ressourcen eines Systems sind zumeist nicht direkt verfügbare und nicht exakt bestimmbare Eingangsgrößen und werden somit häufig aufgrund von Annahmen zum künftigen Ausfall- und Reparaturgeschehens prognostiziert. Eine für die Beschreibung von Störddaten geeignete Kennzahl stellt die *Verfügbarkeit* dar. Die Verfügbarkeit betrachtet nicht nur die Ausfallwahrscheinlichkeit, sondern berücksichtigt reparierbare Systeme, die nach einer gewissen Zeit, wieder vom Defekt- in den Intaktzustand wechseln können. Die Verfügbarkeit ist das Verhältnis aus der Differenz zwischen Einsatzzeit und Ausfallzeit zur Einsatzzeit. Die Einsatzzeit stellt hier die Summe aus Defekt- und Intaktzeiten dar.

In der Praxis besteht jedoch das Problem, dass die Prognose der Verfügbarkeit nur mit erheblichen Einschränkungen möglich ist [VDI83]. Der statistisch gesicherte Nachweis der Verfügbarkeit im praktischen Betrieb ist erst nach entsprechend langer Beobachtungszeit möglich. Unter der Annahme, dass die Wahrscheinlichkeiten für Ausfälle zeitabhängig, aber für disjunkte Zeitintervalle statistisch unabhängig sind, kann der Poisson-Prozess zur mathematischen Darstellung der Ausfälle genutzt werden. Die Abbildung der Störabstände ist damit exponentialverteilt (vgl. hierzu beispielsweise auch [Arn98], S. 236ff. Gleiches gilt für die Störddauer. Vom Bovert entwickelt ein Modell zur Prognose der Verfügbarkeit, welches auch die zeitlich differenzierte Betrachtung erlaubt [Bov00]. Als beste Annäherung in Logistiksystemen wurden für das Reparaturverhalten die Lognormal-Verteilung und für das

Ausfallverhalten die Weibull-Verteilung ermittelt. Allerdings besteht die Schwierigkeit bei der anwendungsspezifischen Parametrisierung der Verteilungen darin, dass die anzugebenden Parameter häufig nicht genau bestimmt werden können. Alternativ kann eine Histogrammverteilung der gemessenen Stör- und Funktionszeiten verwendet werden. Zu beachten ist, dass die Einsteuerung realer Störddaten in das Simulationsmodell i.d.R. dem Zufallsprinzip zur Abbildung einer Störung entgegensteht. Die Ableitung der obigen Kenngrößen aus realen Störddaten mittels Parameterschätzung ist allerdings grundsätzlich möglich und sinnvoll, um die Störverteilung möglichst exakt an die realen Störddaten zu approximieren.

Zur Beschreibung der *Systemlast* kann neben der direkten Nutzung eines oder mehrerer in der Realität vorgegebenen Auftragsvolumina ebenfalls eine theoretische, ggf. statistisch verteilte Systemlast für die Darstellung der Aufträge verwendet werden. Grundsätzlich lassen sich folgende Systemlastbeschreibungen zur Nutzung unterscheiden (vgl. u.a. [WMe93]):

- Reale Betriebsdaten können über die mitprotokollierten Aufträge, die mindestens über Start- und Zielpunkt des Auftrags, über einen Generierungszeitpunkt t_i und eine Auftragskennung spezifiziert sein müssen, direkt als Systemlasten verwendet werden.
- Auftragsmatrizen (z. B. als Start-Ziel-Matrizen für Transportaufträge) erlauben die Betrachtung einer Systemlast auf Basis von Durchschnittskenngrößen für ein betrachtetes Zeitintervall.
- Verteilungen z. B. zur Festlegung der Anzahl zu generierender Aufträge oder der Zwischenankunftszeiten können theoretisch angenommen bzw. durch fachspezifische Kenntnisse über das Auftragsverhalten abgeleitet oder auf der Basis des Gesamtauftragsvolumens, der vorliegenden Auftragsmatrizen mit Dichteschätzungsmethoden berechnet werden, sowie durch die empirische Verteilung approximiert werden. Unter der Annahme, dass die Anzahl der zu generierenden Aufträge poissonverteilt ist, können die Zwischenankunftszeiten exponentialverteilt beschrieben werden.
- Ein weitere Ansatz stellt die algorithmische Beschreibung der den Systemlasten zugrunde liegenden Gesetzmäßigkeiten und ein daraus resultierend programmtechnische Generierung der Auftragsvolumina dar.

Die aufgezeigten Varianten sind mit unterschiedlichen Genauigkeitsgraden verbunden. Im konkreten Anwendungsfall ist in Abhängigkeit von den system- und aufwandsspezifischen Gegebenheiten abzuwägen, welche Vorgehensweise zu wählen ist. Es ist zu beachten, dass eine statistische Verteilung der Systemlast insbesondere die Betrachtung beliebiger Zeiträume und die beliebige Variationen der Auftragslasten zur Prognose ermöglicht. Für die hier vorgestellten Arbeiten ist die programmtechnische Generierung der Auftragsvolumina zunächst von geringer Relevanz.

Zur detaillierten Untersuchung von Eingangsdaten fokussieren sich die folgenden Untersuchungen zunächst auf die Systemlastdaten. Dies liegt zum einen in der Datenkomplexität der Systemlasten für GNL und der daraus resultierenden notwendigen Datenanalysen begründet; zum anderen lassen – im Gegensatz zu Störprofilen – Systemlasten i.d.R. keine direkte Ableitung der zu verwendenden statistischen Verteilung zu. Darüber hinaus ist auch die Problematik der Parameterschätzung bzgl. der gewählten Verteilung (in Analogie zu den Störungen) gegeben.

Zur Untersuchung der Simulationseingangsdaten für GNL haben die Teilprojekte M1 und M9 als Anwendungsprojekt im SFB 559 das Teilprojekt A5 ausgewählt. Die Modelle des Teilprojektes zeichnen sich durch die Betrachtung mehrerer Flughäfen und dort wiederum verschiedener Unternehmen aus. Die Systemlastdaten werden aus heterogenen, qualitativ und quantitativ sehr unterschiedlichen Datenbeständen mehrerer Unternehmen zusammengestellt und unterliegen einer stark vernetzten Informationsstruktur. Sie lassen damit eine herausragende Übertragung der Ergebnisse auf andere Teilprojekte zu. Darüber hinaus sind die Daten für die geplanten Untersuchungen ohne wesentliche Veränderung und Einschränkungen verfügbar.

4 Anwendungsfall A5 und deren Eingangsdaten

Die Untersuchung eines isolierten Luftfrachtknotens im Rahmen der Arbeiten des Anwendungsteilprojektes A5 haben in der ersten Phase des SFB gezeigt, dass der Frachtumschlag an einem Luftfrachtknoten im Wesentlichen von den erbrachten Vorleistungen vorhergehender Knoten abhängt. Im Rahmen der zweiten Phase des SFB-Teilprojektes A5 wird auf Basis der bisher durchgeführten Untersuchungen ein erweitertes Modell entwickelt, welches das für die Beziehungen und Wechselwirkungen zwischen Flughäfen relevante Teilsystem im Sinne eines Ausschnittes des Gesamtnetzes abbildet.

Eine typische Prozesskette für Luftfrachttransporte ist gekennzeichnet durch den Vorlauf, den Hauptlauf und den Nachlauf. Die in diesem Fall betrachtete Prozesskette besteht aus 2 Umschlagknoten und einer Quelle und einer Senke unter Berücksichtigung einer zeitpunktgeführten Steuerung. Die beiden Umschlagknoten innerhalb der Kette besitzen jeweils eine Hubfunktion, so dass damit eine direkte Beziehung zwischen zwei zeitpunktgeführten Hubknoten Hub B und Hub C vorliegt. Jeder Feeder-Flughafen (Quelle) hat ein spezifisches Frachtaufkommen mit unterschiedlichen Zielflughafen-Anteilen (Senke). Die an Hub B ankommende Fracht wird dort umgeschlagen (sammeln, vereinzeln, lagern, weiterleiten). Die Betrachtung innerhalb des Simulationsmodells fokusiert sich auf die Fracht, die über den Hub C den Zielflughäfen zugeführt wird. Hauptuntersuchungsziel ist die Auswirkung verschiedener Umschlagsstrategien auf Hub B und C auf das Gesamtsystem bzgl. beispielsweise Durchlaufzeit, Termintreue sowie Ressourcenverbrauch.

Ausgangspunkt für die hier beschriebenen Analysen und ermittelten Eingangsdaten ist eine feststehende Prozesskette mit festgelegten Systemgrenzen sowie zu untersuchenden Strukturen und Ressourcen. Die zu ermittelnden Eingangsdaten beziehen sich daher hauptsächlich auf die Spezifikation der Systemlast.

4.1 Ermittlung der Systemlastdaten

Anhand der Aufgaben und Zielstellung wurde der Bedarf an Information identifiziert. Daraus ergab sich, dass pro Hub Aussagen zu den Sendungen, zum Flugplan und zum Flugzeugtyp notwendig sind. Diese gliedern sich wie folgt:

Sendungen

- Start-, Zielflughafen, Routing:
  Festlegung der Route über die Umschlagflughäfen der Fracht
- Startzeitpunkt, spätester Zielzeitpunkt zur Modellierung von Termintreue
- Gewicht, Anzahl pro Sendung, evtl. Volumen

Flugplan

- Start- und Landzeitpunkt der Flugzeuge an den Umschlagflughäfen
- Routing der Flugzeuge
- Flugzeugtyp (Kapazität bzw. Zuladungsvolumen und -gewicht, Geschwindigkeit, verwendbare Frachtladungsträger), im Speziellen das reine Frachtzuladevolumen
- Anteil der Expressfrachtsendungen an der Gesamtfracht pro Flugzeug

Flugzeugtyp

- Zuladungsvolumen (Fracht)
- Zuladungsgewicht (Fracht)
- Geschwindigkeit

- verwendbare Frachtladungsträger

Auf der Basis der oben beschriebenen benötigten Information zur Definition der Systemlast wurden im Systemumfeld vorhandene, sekundär erhobene Informationen recherchiert. Folgende Informationen wurden für die weitere Analyse herangezogen:

Sendungen (Daten für Hub B, eines Transportanbieters über einen Monat)

- Start-, Zielflughafen
- Flugnummer und Datum der Flüge, mit denen die Fracht an Hub B ankommt bzw. abtransportiert wird
- Gewicht, Frachtstückanzahl pro Sendung

Flugplan (Daten des Hub B, alle Ankünfte und Abflüge über einen Monat)

- Flugnummer
- Datum und Uhrzeit (tatsächlich / geplant)
- Routing des Flugs
- Flugzeugtyp, mit Abfluggewicht

Aktuelle Flugpläne beliebiger Flughäfen (Plandaten pro Flughafen oder Fluggesellschaft)

- Start- und Landezeit
- Flugnummer

Flugzeugtyp (Hersteller, Flughafenumfeld)

- Zuladungsvolumen und –gewicht (gesamt)
- Geschwindigkeit
- verwendbare Frachtladungsträger

Flughafen-Code-Zuordnung

- IATA, ICAO, Klartextbeschreibung

Road-Feeding-Systematik

- LKW-Anlieferung und Abtransport per Flugnummer identifizierbar

Informationen bzgl. der Sendungsdaten für den ausgesuchten Hub B ist nur für einen Carrier vorhanden. Für diesen sind allerdings die meisten Informationen aus verschiedenen Datenbanken abrufbar: Angaben zu den Sendungen, die Start- und Zielflughäfen, sowie für Zwischenlandungen bis zu zwei Hubs. Es fehlen aber Informationen zu weiteren Zwischenlandungen, so dass das vollständige Routing nicht immer abgebildet werden kann. Ein weiterer wichtiger Punkt ist das Fehlen eines exakten Startzeitpunktes einer Sendung, so dass hierfür die Abflugzeit am Startflughafen genutzt werden muss.

Der Flugplan an Hub B ist für den formulierten Informationsbedarf vollständig, die Flugpläne für Hub C sowie für die Start- und Zielflughäfen sind nur in eingeschränkter Version über sekundäre Informationsquellen, wie beispielsweise das Internet, verfügbar. Allerdings stehen diese nur als aktuelle Daten zur Verfügung und passen somit zeitlich nicht zu den Flugplandaten an Hub B.

Die Flugzeugtypen sind nur im Flugplan für Hub B aufgeführt, ebenso das zugeladene gesamte Frachtgewicht, wobei keine speziellen Daten für Expressfracht existieren.

4.2 Datenanalyse

Die Daten wurden von den verschiedenen Ansprechpartnern und Datenquellen vollständig in digitaler Form erhoben. Anschließend erfolgte eine einheitliche Ablage der Daten mittels quellenspezifischer Importfunktionen, eine parameterspezifische Bennennung und Datentyp-anpassung, Indizierung und Definition von Relationen in einem relationalen Datenbanksystem. Als vorbereitende Hilfsmittel für die Aufbereitung und Analyse der Daten wurden des Weiteren verschiedene SQL-Abfragen und Filter innerhalb der Datenbank definiert und zur Verfügung gestellt.

Mit Hilfe von Datenbankabfragen und multivariater deskriptiver Statistik kann eine Häufigkeitsverteilung der Sendungsströme erstellt werden mit Fokus darauf, dass alle Sendungen über den Hub B abgefertigt werden. Das Ankunftsverhalten der Sendungen wird mit deskriptiver Zeitreihenanalyse auf saisonale Effekte, wie Tages- bzw. Wochenrhythmus, und Trends untersucht. Dabei werden sowohl Ähnlichkeiten bei den Mengenniveaus als auch im strukturellen Verteilungsverhalten im zeitlichen Verlauf aufgedeckt.

Eine Klassifizierung der Flugzeugtypen nach der realistischen Frachtzuladung im Gegensatz zu einer Klassifizierung nach dem Abfluggewicht muss vorgenommen werden, um eine Verfälschung bei der möglichen Zuladung der Fracht zu vermeiden. Dies geschieht mit statistischen Verfahren zur Clusteranalyse aus den Ankunfts- und Abflugdaten in Hub B. Dabei sind die Größe des Datensatzes, sowie fehlerhaft eingegebene oder fehlende Daten eine besondere Herausforderung. Dazu sind Clustermethoden nötig mit schnellen Algorithmen und guten robusten Eigenschaften.

4.3 Validierung der Ergebnisse

Sendungen Hub B

Zur Bestimmung des Ankunftszeitverhaltens von Sendungen wurden zunächst alle Ankünfte von Sendungen innerhalb des betrachteten Zeitraums von 4 Wochen und deren Gewicht in ein Halbstundenraster eingeteilt und visualisiert (siehe Abbildung 4). Die Darstellung der Ankünfte mit halbstündiger Genauigkeit ist zur Auswertung ausreichend, da für eine kleinere Aufteilung nicht genügend Messwerte und damit Lücken in der Darstellung und eine größere Aufteilung ein zu grobes Abbild der Ankunftsverteilung ergeben.

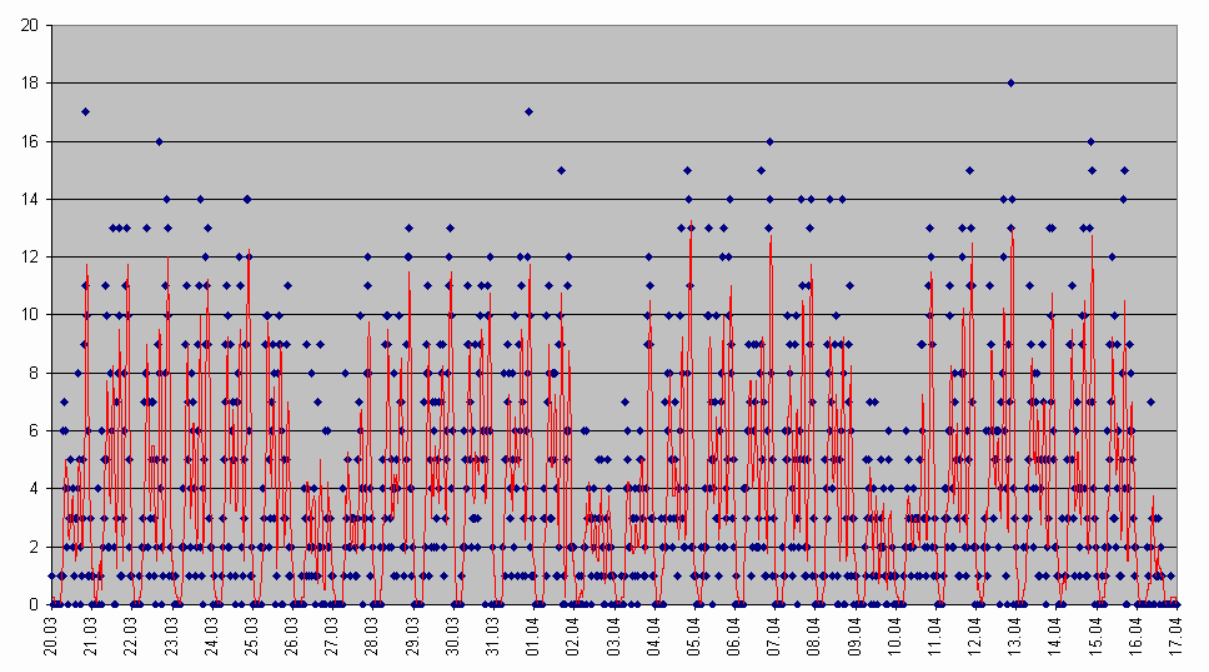


Abbildung 4: Ankommende Sendungen im Halbstundenraster an Hub B.

Anhand dieser deskriptiven Darstellung konnte sowohl ein typisches Tagesprofil als auch ein typisches Wochenprofil bezüglich der Anzahl ankommender Sendungen auf dem Hub B identifiziert werden. Die jeweiligen Tagesprofile unterscheiden sich demnach nicht strukturell, aber sehr wohl in der absoluten Menge ankommender Sendungen. Deshalb wurden mit Hilfe des Wochenprofils die Unterschiede des mittleren Sendungsaufkommens und der Streuung an den jeweiligen Tagen ermittelt (siehe Abbildung 5). Dies diente als Basis zur Normierung der Wochentage, um ein einheitliches Niveau für ein typisches relatives Tagesprofil zu erhalten (siehe Abbildung 6). Zur Ermittlung eines spezifischen Sendungsaufkommens zu einer bestimmen Tageszeit an einem bestimmten Wochentag muss der Tagesmittelwert mit dem relativen Tagesprofil verrechnet werden.

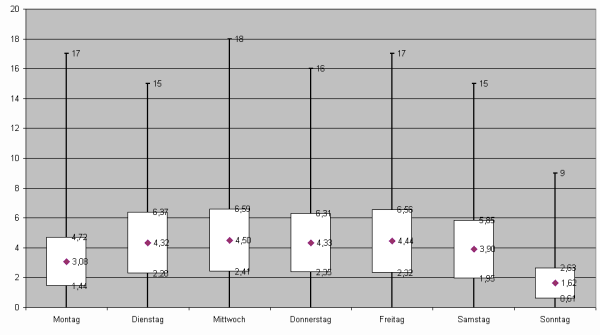


Abbildung 5: Tagesverteilung der Anzahl der Sendungen im Wochenprofil (Ankunft)

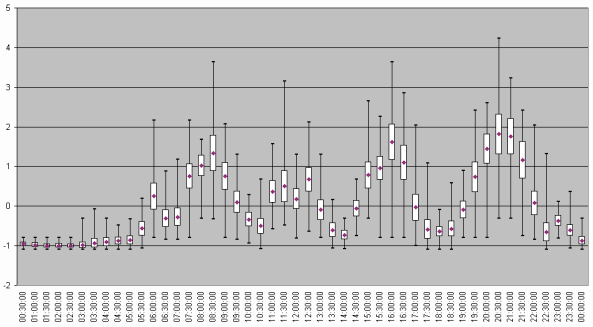


Abbildung 6: relatives Tagesprofil der Anzahl der Sendungen im Halbstundenraster (Ankunft)

Durch diese Aufbereitung der Daten wird eine deutliche Vereinfachung der Eingangsdaten für die Simulation gewährleistet. Hierbei reicht es aus, das relative Tagesprofil und die jeweiligen Niveaus der Wochentage zu verwenden. Gleichzeitig wurde verifiziert, dass die Ankünfte der Sendungen über vier Wochen ein ähnliches Verhalten aufweisen, so dass die Datenbasis dahingehend verbreitert werden kann, dass pro Messzeitpunkt (d. h. pro halbe Stunde) 28 Messwerte (4 Wochen x 7 Tage) zur Verfügung stehen. Damit können sowohl die Mittelwerte als auch die Standardabweichungen im relativen Tagesprofil als aussagekräftig bezeichnet werden.

Nach der Analyse und Validierung liegen folgende Eingangsdaten vor:

- Wochenverteilung der Ankunftssendungen an Hub B: Häufigkeitsverteilung der Ankunftssendungen pro Wochentag (mit Parametern Mittelwert und Standardabweichung, sowie Minimum und Maximum)
- Tagesverteilung der Ankunftssendungen an Hub B: relative, normierte Häufigkeitsverteilung der Anzahl an Ankunftssendungen pro halbe Stunde (mit Parametern Mittelwert und Standardabweichung, sowie Minimum und Maximum)
- Wochenverteilung der Abflugssendungen an Hub B: Häufigkeitsverteilung der Anzahl an Abflugsendungen pro Wochentag (mit Parametern Mittelwert und Standardabweichung, sowie Minimum und Maximum)
- Tagesverteilung der Abflugssendungen an Hub B: relative, normierte Häufigkeitsverteilung der Anzahl an Abflugsendungen pro halbe (mit Parametern Mittelwert und Standard-abweichung, sowie Minimum und Maximum)

Flugplan Hub B

Zur Abbildung des Flugplans auf Hub B, d. h. des Ankunftszeitverhaltens bezüglich landender und startender Flugzeuge, wurden wiederum alle Landungen und Starts innerhalb des betrachteten Zeitraums von vier Wochen in ein Halbstundenraster eingeteilt und visualisiert.

Die deskriptive statistische Analyse ergab, dass im Gegensatz zum Sendungsaufkommen kein oder ein zu vernachlässigendes Wochenprofil zu erkennen ist, welches in der Varianz des anschließend bestimmten Tagesprofils mit eingeschlossen ist. Die Normierung der Werte kann damit entfallen, trotzdem ist die Verbreiterung der Datenbasis auf 28 Messwerte pro halbe Stunde auch hier möglich.

Nach der Analyse und Validierung liegen folgende Eingangsdaten vor:

- Tagesverteilung der Flugzeugankünfte an Hub B: Häufigkeitsverteilung der Anzahl an Flugzeugankünften pro halbe Stunde (mit Parametern Mittelwert und Standardabweichung, sowie Minimum und Maximum)
- Tagesverteilung der Flugzeugabflüge an Hub B: Häufigkeitsverteilung der Anzahl an Flugzeugabflüge pro halbe Stunde (mit Parametern Mittelwert und Standardabweichung, sowie Minimum und Maximum)

Flugzeugtyp

Die Beschreibung der Transportkapazitäten beim An- und Abflug an den Hubknoten hängt entscheidend von der Frachtkapazität der jeweiligen Flugzeugtypen ab. In der Flugbewegungsdatenbank sind 149 verschiedene Flugzeugtypen registriert, inklusive der Spezialtypencodes, die häufig aus konstruktionstechnischen Gründen maßgeblich unterschiedliche Frachtkapazitäten beschreiben, erhöht sich die Anzahl auf 349 Typen.

Da andererseits das maximale Abfluggewicht und auch die tatsächlich transportierte Fracht bei vielen unterschiedlichen Typen sehr ähnlich sind, ist es sinnvoll die Flugzeugtypen bzgl. ihrer Frachtkapazität in Gruppen zu klassifizieren. Dabei muss notwendigerweise die real transportierte Fracht im Vordergrund stehen. Das maximale Abfluggewicht der Flugzeugtypen stellt nur einen unzureichenden Klassifizierungsparameter dar, denn sowohl das Flugzeug- als auch das Gepäckgewicht und die Anzahl der Passagiere verringern die tatsächliche Frachtkapazität.

Mit Hilfe der Daten zum Frachtgewicht sowie der Gepäckgewichts- und Postgewichtsdaten aber auch der Anzahl der Passagiere aus der Sendungsdatenbank von Hub B kann mit einer statistischen Clusteranalyse eine Klassifizierung bzgl. der tatsächlichen Transportkapazitäten der unterschiedlichen Flugzeugtypen durchgeführt werden. Dabei fungieren die letztgenannten Variablen als Kovariablen, die das eigentliche Klassifizierungsmerkmal Frachtgewicht zusätzlich beeinflussen. Da die Datenaufnahme zu den oben beschriebenen Variablen oft manuell durchgeführt wird, gibt es zahlreiche fehlende oder unplausible Datenwerte. Deshalb müssen robuste Klassifizierungsverfahren wie CLARA [KRo90] oder andere robustifizierte Verfahren – klassische Clusteranalyse mit robuster Vorverarbeitung durch geeignete Parameterschätzung [RDr99] – zum Einsatz kommen [Kuh03]. Das Ergebnis der unterschiedlichen Clusteranalysen ist die Klassifizierung der Flugzeugtypen in sechs Gruppen, wobei bei der Anwendung des Algorithmus CLARA eine Verteilung jedes einzelnen Flugzeugtyps auf die verschiedenen Cluster angegeben werden kann. Die Analysen mit robustifizierten Verfahren liefern dagegen eine eindeutige disjunkte Klassifizierung der Flugzeugtypen in sechs Cluster. Das letztere Ergebnis erlaubt somit für jeden Typ die Bildung einfacher Eingangsdaten bzgl. des Frachtgewichts, die Ergebnisse mit der Methode CLARA sind dagegen genauer und damit geeignet für Modellierungen mit sensitiveren Zielstellungen bzgl. des Frachtgewichts.

Nach der Analyse und Validierung liegen folgende Eingangsdaten vor:

- Eindeutige Klassifizierung der Flugzeugtypen in 6 Cluster: Kennzahlen zum Frachtgewicht, Passagierzahlen, Gepäckgewicht für jeden Cluster (Mittelwert und Streuung).
- Aufteilung jedes Flugzeugtyps auf 6 Cluster gemäß relativer Häufigkeiten und für jeden Cluster Bereitstellung von Kennzahlen zum Frachtgewicht, Passagierzahlen, Gepäckgewicht (Mittelwert und Streuung).

Die hier beschriebenen Eingangsdaten sind für die quantitative Beschreibung der Systemlast ausreichend. Erweiterte, untersuchungsspezifische Lenkungs- und Steuerungsalgorithmen innerhalb des Simulationsmodells benötigen hingegen weitere Informationen bzgl. der einzelnen Sendungen. Hierzu zählen insbesondere oben beschriebene Kovariablen und Informationen zum Routing einzelner Sendungen, deren Gewinnung und datentechnische Beschreibung aufgrund der sehr system- und untersuchungsspezifischen Anforderungen hier nicht weiter aufgeführt ist.

5 Klassifikation und Modellierung von Systemlasten

Die Komplexität von Systemlasten zur Modellierung von großen Logistiknetzen erfordert eine genaue Analyse des benötigten Qualitäts-, Quantitäts- und Granularitätsniveaus (z.B. Aktualität, Genauigkeit, Betrachtungszeitraum etc.) der zu verwendeten Eingangsdaten. Um die Modellierung der Systemlast für die Anwendung zu standardisieren, ist eine Klassifikation der Systemlast nach deren Eigenschaften, Abhängigkeiten und Darstellungsformen notwendig. Darauf aufbauend leiten sich Abbildungsvorschriften für die Modellierung mit dem Werkzeug ProC/B entsprechend ab.

5.1 Klassifikation der Systemlast

Die standardisierte Darstellung von Systemlasten zur Modellierung von GNL verlangt eine Klassifikation der Systemlast bezüglich Ihrer Eigenschaften und der zugehörigen Darstellungsform für typische Anwendungsfälle. Ein weiterer, wichtiger Punkt sind in diesem Zusammenhang die Abhängigkeiten zwischen verschiedenen Eigenschaften einer Systemlast. Insgesamt ergibt sich damit eine dreidimensionale Beschreibung der Systemlast, die durch eine Matrix visualisierbar ist (Abbildung 7). Die konkrete Klassifikation einer Systemlast ergibt sich in dieser Darstellungsform aus der Belegung der Schnittpunkte innerhalb dieser Matrix.

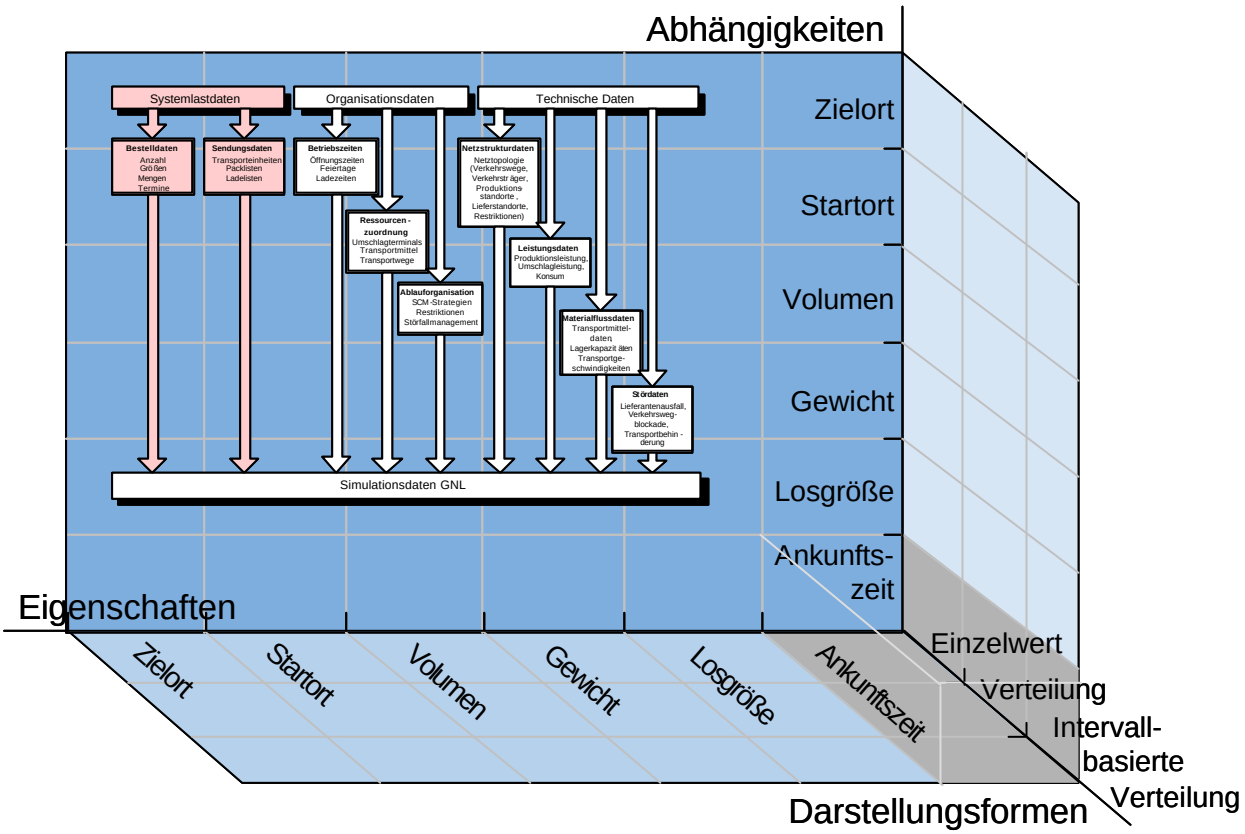


Abbildung 7: Eigenschaften und Darstellungsformen der Systemlast.

Die grundlegende, unabdingbare Eigenschaft einer Systemlast stellt die Ankunftszeit eines temporären Objektes dar. Erst die Angabe dieser Information ermöglicht die zeitliche Entzerrung der in ein System einfließenden Objekte und stellt damit die Beschreibung einer dynamischen Systembelastung über die Zeit in Analogie zur Realität sicher. Weitere Eigenschaften wie beispielsweise

- Identifizierung eines Objektes,
- Los-, Batch- oder Chargenzuordnung,
- Physikalische Eigenschaften (Gewicht, Volumen, etc.) eines Objektes,
- geografischer Startort und Zielort eines Objektes und
- weitere, systemabhängige Objekteigenschaften

sind dagegen optional und hängen direkt vom betrachteten Untersuchungsgegenstand und der Aufgabenstellung der Untersuchung ab. Sie sind die Eigenschaften, die die Restriktionen des Systems determinieren.

Die typischen Darstellungsformen der einzelnen Eigenschaften sind je nach Genauigkeitsvorgabe der Aufgabenstellung (Granularität) als Einzelwerte, als Verteilung oder als intervallbasierte Verteilungen zu klassifizieren. Dabei können innerhalb einer Systemlastmodellierung für die verschiedenen Eigenschaften durchaus unterschiedliche Darstellungsformen gewählt werden. Die Verwendung von Einzelwerten ist eine sehr detaillierte Darstellungsform mit der stärksten Verknüpfung an die reale Datenbasis. Diese Darstellungsform hat allerdings den Nachteil, dass Generalisierungen im Sinne von Hochrechnungen auf andere Fallbeispiele bzw. zeitliche Prognosen nicht sinnvoll durchführbar sind. Eine verteilungsbasierte Darstellung der jeweiligen Eigenschaften der Systemlast ist dagegen durch die Glättung der Häufigkeitsverteilung sehr viel besser geeignet, generalisierende Ergebnisse einer Modellierung zu erlangen. Neben der Angabe der Verteilungsart sind die typischen Kennzahlen einer Verteilung ein Lagemaß (arithmetisches Mittel, Median) und ein Streuungsmaß (Standardabweichung). Je nach benötigter Genauigkeit der Eingangsdaten kann die Angabe der Lage und der Streuung ausreichend sein, genauer, aber auch aufwendiger, ist die Modellierung des Verteilungsfunktionals. Unter Berücksichtigung von Abhängigkeitsstrukturen zwischen den Eigenschaften einer Systemlast können intervallbasierte Verteilungen notwendig sein. Ein typisches Fallbeispiel ist, eine Häufigkeitsverteilung der Ausprägungen einer Eigenschaft, die bedingt wird durch eine weitere Systemlasteigenschaft. Dann ergibt sich für jede Ausprägungsklasse bzw. jedes Ausprägungsintervall dieser zweiten Eigenschaft eine eigene Verteilungsvorschrift, so dass als Eingangsdatum eine Verteilung der jeweiligen Klasse zur Verfügung stehen muss. Die Kennzahlen (Lage- und Streuungsmaße) der jeweiligen Verteilungen können dabei sehr unterschiedlich sein. Dagegen ist der Verteilungstyp in der Regel für alle Intervalle gleich, da davon ausgegangen wird, dass die grundlegende Verteilungsstruktur einer Eigenschaft über alle Klassen bzw. Intervalle erhalten bleibt.

5.2 Systemlastmodellierung in ProC/B

5.2.1 Standardverteilungen

Eine Modellierung der in ProC/B standardisiert definierten Verteilungen kann unmittelbar durch die Parameter an den Quellen erfolgen. Die in Kapitel 2.2.2 beschriebenen Wahrscheinlichkeitsverteilungen können dazu für die Zwischenankunftszeiten der Prozesse verwendet werden. Außerdem können die Verteilungen auch in arithmetischen Ausdrücken verwendet werden. Eine beispielhafte Verwendung findet sich in Abbildung 1.

5.2.2 Intervallbasierte Verteilungen

Zur Modellierung intervallbasierter Verteilungen von Zwischenankunftszeiten ist ein Steuerungsprozess erforderlich, der Prozesse auf der vorgangsbeschreibenden Prozesskette generieren lässt. Der Steuerungsprozess bildet die intern in Quellen vorhandene prozessgenerierende Schleife nach, erlaubt jedoch zusätzliche Modellierungen der Zwischenankunftszeiten. Dazu wird der Steuerungsprozess zum Zeitpunkt 0 gestartet und durchläuft dann einen ständigen Wechsel von Prozessgenerierung und Warten.

Ein Beispiel für eine auf Zeitintervallen beruhende Verteilungsvariation findet sich in Abbildung 8. Dort werden für verschiedene Tageszeitbereiche (eine Modellzeiteinheit entspricht hier einer Stunde) unterschiedliche Zwischenankunftsverteilungen gewählt. Von 20 bis 6 Uhr werden deterministische Zwischenankunftszeiten von einer halben Stunde verwendet. Danach sind sie bis 9 Uhr negativ-exponential verteilt mit einem Mittelwert von 15 Minuten (= 4 Ankünfte je Stunde). Von 9 bis 16 Uhr sind sie wiederum negativ-exponential verteilt, wobei die Verteilung um 0.05 Stunden verschoben wird, was hier einen Mittelwert von 9 Minuten und Mindestabstand von 3 Minuten ergibt, und von 16 bis 20 Uhr sind sie negativ-exponential verteilt mit einem Mittelwert von 20 Minuten. Die Auswahl der Verteilungen erfolgt im Modell dadurch, dass zunächst die aktuelle Uhrzeit berechnet wird (PKE `Uhrzeit_feststellen`). Dazu werden von der aktuellen Modellzeit (`time`) die Stunden bereits vollständig vergangener Tage abgezogen (`((time - 0.5) // 24` ergibt die Anzahl bereits vollständig vergangener Tage). Die ermittelte Uhrzeit wird an einem Oder-Konnektor ausgewertet und somit zu den verschiedenen Verteilungen der Zwischenankunftszeiten verzweigt. Innerhalb dieser Alternativen findet lediglich ein Wartevorgang statt, dessen Länge den Zwischenankunftszeiten der Alternative entspricht. Nach dem Warten erfolgt dann die Generierung eines nebenläufigen Prozesses auf einer zweiten Prozesskette, die den zu modellierenden Vorgang beschreibt (in Abbildung 8 durch das Element `zu_startende_PK` angedeutet).

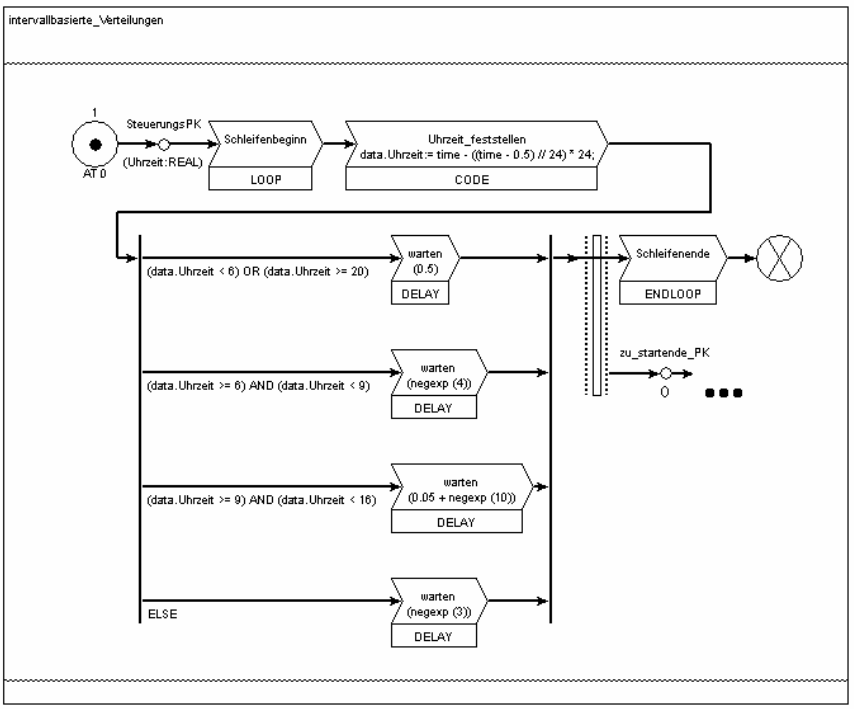


Abbildung 8: Modellierung intervallbasierter Verteilungen von Zwischenankunftszeiten.

Soll anders als in Abbildung 8 der erste Vorgangs-Prozess zum Zeitpunkt 0 generiert werden, so ist der PK-Konnektor vor den ersten Oder-Konnektor zu platzieren.

Auf ähnliche Weise können auch Intervalle nachgebildet werden, die auf anderen Parametern als der Zeit oder auf mehreren Parameter gemeinsam basieren.

Ferner können auch Intervalle nachgebildet werden, die Werte für Prozess-Parameter festlegen. In Abbildung 9 ist die intervallbasierte Festlegung eines Parameters `Gewicht` dargestellt, dessen Verteilung hier ebenfalls von der Zeit abhängig ist. Dazu wurde das Beispiel aus Abbildung 8 erweitert und in die Zwischenankunftszeit-Alternativen weitere Verzweigungen eingefügt. Damit die Parameterwerte in der Prozesskette `zu_startende_PK` verwendbar werden, sind ferner die Gewichtswerte am Prozessketten-Konnektor über eine globale Variable in diese Prozesskette zu kopieren.

Für den Fall, dass die Intervalle für einen Parameter nicht von den Intervallen der Zwischenankunftszeit abhängen, ist die Wertebestimmung dieses Parameters hinter die Festlegung der Zwischenankunftszeiten zu platzieren (vgl. Bestimmung des Parameters `Volumen` in Abbildung 9).

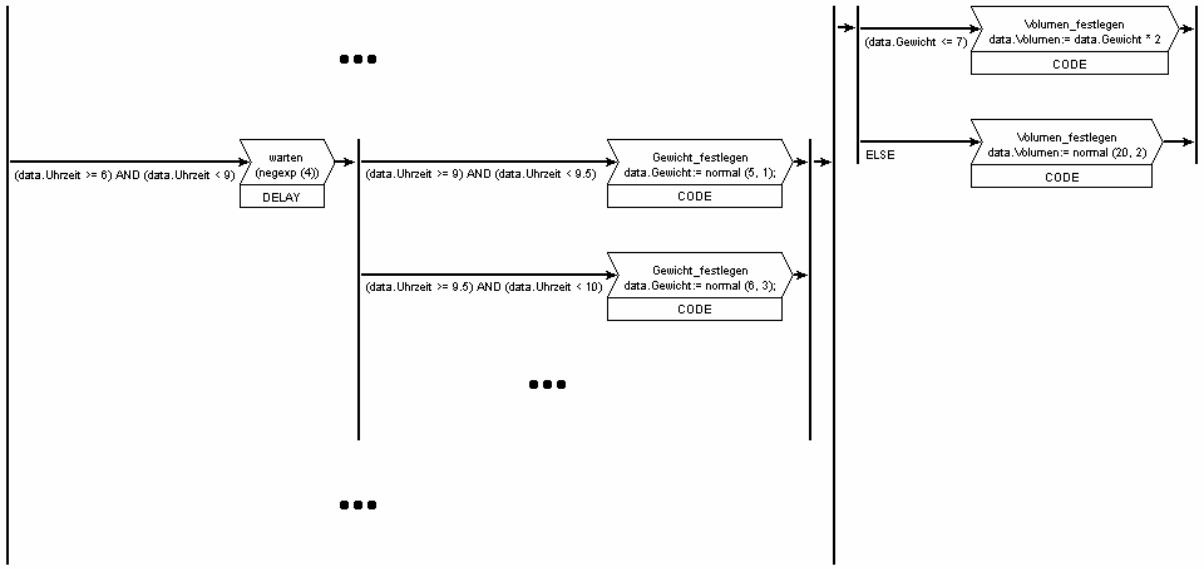


Abbildung 9: Modellierung intervallbasierter Verteilungen von Prozess-Parametern.

5.2.3 Einzelwert-Darstellungen

Die Modellierung von Einzelwert-Darstellungen basiert auf der gleichen Grundidee wie die Modellierung der intervallbasierten Verteilungen (siehe Kapitel 5.2.2). Wiederum ist ein Steuerungsprozess erforderlich. Dessen Wartezeiten werden jedoch über eine in einem Parameter abgelegte Liste gesteuert, auf die mit Hilfe eines Index zugegriffen wird.

Ein Beispiel hierzu findet sich in Abbildung 10. Dort durchläuft der `index` zyklisch die Werte 1 bis 20, indem er in jedem Schleifendurchlauf erhöht, bei Erreichen des Wertes 21 jedoch auf 1 zurückgesetzt wird. Durch diesen Index erfolgt beim Delay-PKE `warten` der Zugriff auf die die Zwischenankunftszeiten beschreibende Liste (`Wartezeit`).

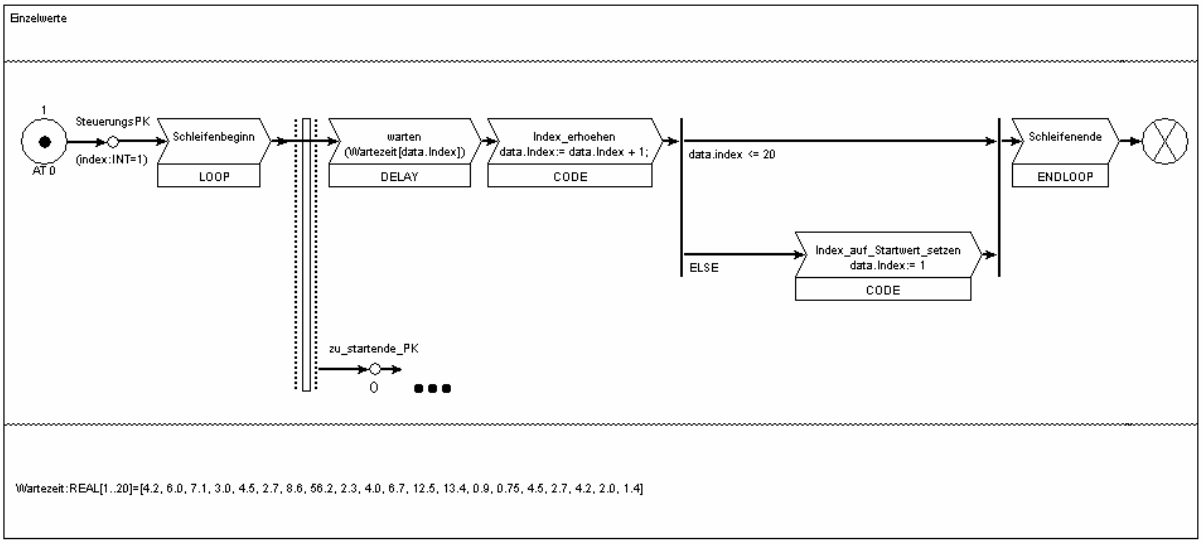


Abbildung 10: Modellierung von Einzelwert-Darstellungen.

Soll anders als in Abbildung 10 der erste Vorgangs-Prozess nicht zum Zeitpunkt 0 sondern erst nach Abwarten der ersten Zwischenankunftszeit generiert werden, so ist der PK-Konnektor hinter das Delay-PKE zu platzieren.

Ein Einbinden von Einzelwert-Darstellungen in Intervalle ist ferner auch möglich. Dazu ist die komplette Einzelwert-Darstellungs-Schleife in eine Intervall-Alternative einzufügen. Innerhalb der Alternative ist dann eine Initialisierung des Index-Wertes vor den Einzelwert-Schleifenstart zu stellen und das Einzelwert-Schleifenende vom intervallabgrenzenden Zustand abhängig zu machen (bei Zeitintervallen muss die Einzelwert-Schleife also enden, sobald die Uhrzeit außerhalb des zugehörigen Intervalls liegt).

6 Ausblick

Die Arbeiten zur standardisierten Beschreibung von Eingangsdaten für die Simulation auf Basis des Prozesskettenparadigmas haben gezeigt, dass einerseits eine Klassifikation für die Systemlast von modellgestützten Analyseverfahren definiert werden konnte und die Modellierung der Eigenschaften und Abhängigkeiten der Systemlast gut durch das dem Prozesskettenparadigma folgenden Werkzeug ProC/B unterstützt wird. Zur prozessorientierten Modellierung der Systemlast war insbesondere die Definition von nebenläufigen Steuerungsprozessen im Sinne eines Informationsflusses sowie die Verwendung des neu in das Werkzeug integrierte Modellierungselements LOOP hilfreich und zielführend. Hierdurch wurde die Modellierung auch sehr komplexer Systemlasten erst ermöglicht.

Die Übertragung der Systemlastklassifikation, bestehend aus Eigenschaften und Abhängigkeiten, auf verschiedenste Anwendungsgebiete der modellgestützten Analyse von GNL ist grundsätzlich gegeben, muss aber jeweils dem Untersuchungsgegenstand und der Aufgabenstellung angepasst werden. Zur weiteren Unterstützung des Anwender zur Bestimmung der systemrelevanten Eigenschaften wird in weiteren Arbeiten eine Einteilung der Systemlast-Eigenschaften in Klassen erfolgen, um darüber dem Anwender weitere Unterstützung zur vollständigen Formulierung seiner benötigten Systemlast zu geben.

Wie in Kapitel 3 dargestellt werden für eine modellgestützte Analyse neben den Systemlastdaten auch Organisationsdaten und Technische Daten zur Modellerstellung benötigt. In Analogie zu den im Bereich der Systemlast durchgeführten Arbeiten werden auch für diese Eingangsdaten entsprechende Klassifikationsschemas und Abbildungsvorschriften erstellt.

Die Bestimmung des typischen Informationsbedarfs sowie der daraus abzuleitenden Eingangsdaten für identifizierte Standardprozesse auf unterschiedlichstem Abstraktionsniveau in großen Netzen der Logistik ist ein weiteres Forschungsziel, das als solches auch im Antrag für die dritte Phase des SFB 559 beschrieben ist. Diese Arbeiten sollen zum Aufbau von Informationenkategorien im Anwendungskontext großer Netze der Logistik und in diesem Zusammenhang ebenfalls zur Entwicklung von Kriterien zur Bestimmung der Informationsgüte und einer systematisierten Bewertung der Informationsquellen führen.

7 Literatur

[Arn98] Arnold, D: Materialflusslehre. Vieweg Verlag, Wiesbaden, 1998.

[BBF+02] Bause, F.; Beilner, H.; Fischer, M.; Kemper, P.; Völker, M.: The Proc/B Toolset for the Modelling and Analysis of Process Chains. In: Field, T.; Harrison, P.-G.; Bradley, J.; Harder, U. (Hrsg.): Computer Performance Evaluation Modelling Techniques and Tools. 12th International Conference on Modelling Techniques and Tools (Performance Tools 2002), London, UK, 14.-17. April, LNCS 2324, Springer, Berlin Heidelberg New York, 2002, S. 51-70.

[BBT+99] Beilner, H.; Bause, F.; Tatlitürk, H.; van Almsick, A.; Völker, M.: Zum B-Modellformalismus - Version B1 - zur Vorbereitung automatisierter Analysen von Modellen logistischer Systeme hinsichtlich technischer, ökonomischer und ökologischer Ziele, SFB-Bericht 99002, Universität Dortmund, 1999.

[BFK+93] Büttner, M. (ed.); Fricke, F.; Klaaßen, O.; Nolte, S.; Stahl, H.; Weißenberg, N.: Hi-Slang Reference Manual for the Hierarchical Evaluation Tool HIT. Universität Dortmund, Informatik IV, 1993.

[Bov00] Bovert, E.-M. vom: Modellerstellung zur Verfügbarkeitsprognose komplexer Förder- und Lagersysteme. Dissertation, Universität Dortmund, Verlag Praxiswissen, Dortmund, 2000.

[KRo90] Kaufman, L.; Rousseeuw, P.J.: Finding Groups in Data: An Introduction to Cluster Analysis, Wiley, New York, 1990.

[Kuh95] Kuhn, A.: Prozessketten in der Logistik, Entwicklungstrends und Umsetzungsstrategien. Verlag Praxiswissen, Dortmund, 1995.

[Kuh03] Kuhls, S.: Robuste Clusteranalyseverfahren zur Klassifikation von Flugzeugtypen gemäß transportiertem Frachtgewicht. Diplomarbeit, Universität Dortmund, Fachbereich Statistik, Dortmund, 2003.

[RDr99] Rousseeuw, P.J.; Driessen, K. van: A Fast Algorithm for the Minimum Covariance Determinant Estimator, Technometrics 41, 1999, S. 212-223.

[VDI83] VDI 3581: Verein Deutscher Ingenieure, Düsseldorf, Zuverlässigkeit und Verfügbarkeit von Transport- und Lageranlagen, Beuth Verlag, Berlin, 1983.

[VDI00] VDI 3633 Blatt 1: Verein Deutscher Ingenieure, Düsseldorf, Simulation von Logistik-, Materialfluss- und Produktionssystemen, Entwurf. Beuth Verlag, Berlin, 2000.

[WMe93] Wenzel, S.; Meyer, R.: Kopplung der Simulation mit Methoden des Datenmanagement. In: Kuhn, A.; Reinhardt, A.; Wiendahl, H.-P. (Hrsg.): Handbuch Simulationsanwendungen in Produktion und Logistik. Fortschritte der Simulationstechnik. Vieweg Verlag, Braunschweig, 1993, S. 347-368.