



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Bachelorarbeit

Torben Woggan

Validierung von Gruppensimulationen

Torben Woggan

Validierung von Gruppensimulationen

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung

im Studiengang Angewandte Informatik
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer : Prof. Dr. Thomas Thiel-Clemen
Zweitgutachter : Prof. Dr. Stefan Sarstedt

Abgegeben am 26. August 2011

Torben Woggan

Thema der Bachelorarbeit

Validierung von Gruppensimulationen

Stichworte

Validierung, Gruppensimulation, Evakuierungssimulation, Simulation, Multiagentensystem, agentenbasierte Simulation, Analyse, Fußgängerdynamik

Kurzzusammenfassung

Diese Arbeit befasst sich mit verschiedenen Möglichkeiten zur Validierung von Gruppensimulationen. Hierzu wird über den bisherigen Forschungsstand aufgeklärt, indem zuerst allgemeine Techniken zur Validierung von Simulationen sowie zur Validierung von Multiagentensystemen, die auch bei der Validierung von Gruppensimulationen Anwendung finden können, vorgestellt werden. Dies wird für spezielle Techniken zur Validierung von Gruppensimulationen wiederholt. Die Verfahren werden auf ihre Vor- und Nachteile sowie Anforderungen untersucht. Auf Basis dieser Überlegungen wird ein Konzept zur verlässlichen Validierung von Gruppensimulationen in den unterschiedlichen Phasen der Entwicklung erstellt. Die Anforderungen des Konzepts und die Anwendung des Konzepts werden anhand eines Beispiels beschrieben.

Torben Woggan

Title of the paper

Validation of Crowd Simulations

Keywords

Validation, Crowd Simulation, Evacuation Simulation, Simulation, Multi-Agent System, Agent-Based Simulation, Analysis, Pedestrian Dynamics

Abstract

This paper deals with various possibilities for the validation of crowd simulations. Therefore, the existing research will be presented. First of all, general techniques for the validation of simulations and the validation of multi-agent systems which can also be applied to the validation of crowd simulations are shown and explained. After this the same will be done for special techniques for the validation of crowd simulations. The techniques will be analyzed for their advantages and disadvantages and their requirements. Based on these considerations, a concept for reliable validation of crowd simulations in the different phases of development will be created. The requirements of the concept and the application of the concept will be described with the help of an example.

Inhaltsverzeichnis

Tabellenverzeichnis	6
Abbildungsverzeichnis	7
1 Einleitung	9
1.1 Problem- und Themenbeschreibung	9
1.2 Frage- und Aufgabenstellung	10
1.3 Themenabgrenzung	10
1.4 Ziel der Arbeit und Bedeutsamkeit	11
1.5 Struktur der Arbeit	11
2 Überblick über den Forschungsstand	13
2.1 Validierung	13
2.1.1 Definitionen in der Literatur und Abgrenzung zur Verifizierung	13
2.1.2 Arten der Validierung und Validität	14
2.1.3 Bedeutsamkeit	16
2.2 Gruppensimulationen	17
2.2.1 Definition und Einsatzbereiche	17
2.2.2 Arten von Gruppensimulationen	17
2.2.3 Technische Umsetzungen	18
2.2.4 Bedeutsamkeit	20
2.3 Validierung von Gruppensimulationen	21
2.3.1 Generelle Vorgehensweisen	21
2.3.2 Verfahren zur Validierung von Simulationsmodellen	23

2.3.3	Probleme bei der Validierung von agentenbasierten Simulationen.....	26
2.3.4	Verfahren zur Validierung von agentenbasierten Simulationen	27
2.3.5	Probleme bei der Validierung von Gruppensimulationen	31
2.3.6	Verfahren zur Validierung von Gruppensimulationen	32
3	Erstellung eines Validierungskonzeptes.....	59
3.1	Entwicklungsphasen und die dazugehörige Validierung	59
3.2	Auswahl und Bewertung von geeigneten Verfahren.....	61
3.3	Erstellung eines konkreten Validierungskonzeptes	70
3.4	Validierung am Beispiel des WALK-Projektes	74
3.4.1	Vorstellung des WALK-Projektes.....	74
3.4.2	Entwicklungs- und Validierungsstand	77
3.4.3	Anforderungen des Konzeptes.....	79
3.4.4	Anwendbarkeit des Konzeptes.....	81
3.4.5	Anwendung des Konzeptes auf WALK	90
4	Schlussbetrachtung	97
4.1	Zusammenfassung der Ergebnisse.....	97
4.2	Fazit und Ausblick.....	98
5	Literaturverzeichnis	100

Tabellenverzeichnis

Tabelle 1: Vergleich der Vor- und Nachteile beim Vergleich der Simulation mit Übungen ..	63
Tabelle 2: Vergleich der Vor- und Nachteile der webbasierten Face Validation	64
Tabelle 3: Vergleich der Vor- und Nachteile der Nutzung von Immersion und Virtual Reality	65
Tabelle 4: Vergleich der Vor- und Nachteile der statistischen Validierung	66
Tabelle 5: Vergleich der Vor- und Nachteile der mathematischen Validierung	67
Tabelle 6: Vergleich der Vor- und Nachteile der Validierung durch die RiMEA.....	68

Abbildungsverzeichnis

Abbildung 1: Aufnahme und Simulation zu jeweils gleichen Zeitpunkten	34
Abbildung 2: Die realen Aufnahmen des Szenarios	35
Abbildung 3: Das Szenario der Aufnahme in der Simulation	36
Abbildung 4: Eine Testperson mit Head-Mounted Display und Sensoren.....	37
Abbildung 5: Einer der Bereiche in denen der Punkt p liegt.....	40
Abbildung 6: Vergleich von 40 Simulationsläufen mit der historischen Datenmenge einer anderen Simulation unter Nutzung von 3 Hauptkomponenten. 90 Prozent der Werte liegen unterhalb des Grenzwertes	41
Abbildung 7: Vergleich von 40 Simulationsläufen mit der historischen Datenmenge der gleichen Simulation unter Nutzung von 3 Hauptkomponenten. 3 Prozent der Werte liegen unterhalb des Grenzwertes	42
Abbildung 8: Darstellung von zwei Kurven von unterschiedlichen Simulationen, die die Anzahl des Replizierens (snap-point frequency) zum Grenzwert (distance threshold) angeben.....	44
Abbildung 9: Tatsächlich gemessene Anzahlen des Replizierens (snap-point frequency) zum Grenzwert (distance threshold) von zwei unterschiedlichen Simulationen	44
Abbildung 10: Das Fundamentaldiagramm, das den spezifischen Fluss in P/ms zur Dichte in P/m^2 angibt	49
Abbildung 11: Der architektonische Grundriss und die Startverteilung der Personen von Test 6.....	50
Abbildung 12: Der architektonische Grundriss der verschiedenen Stockwerke und der Rettungsweg von Test 8.....	51
Abbildung 13: Der architektonische Grundriss und die Startverteilung der Personen von Test 9.....	52
Abbildung 14: Der architektonische Grundriss, die Startverteilung der Personen und die zugewiesenen Rettungswege in Test 10.....	53

Abbildung 15: Der architektonische Grundriss und die Startverteilung der Personen von Test 11.....	54
Abbildung 16: Der architektonische Grundriss und die Startverteilung der Personen von Test 12.....	55
Abbildung 17: Der architektonische Grundriss und die Startverteilung der Personen von Test 13.....	56
Abbildung 18: Der architektonische Grundriss von Test 14	57
Abbildung 19: Das UML-Komponentendiagramm von WALK	76
Abbildung 20: UML-Sequenzdiagramm zum Zugriff vom GIS auf die Agenten	84
Abbildung 21: UML-Aktivitätsdiagramm zum Ablauf der Berechnungen bei der statistischen Validierung	87
Abbildung 22: UML-Aktivitätsdiagramm zum Ablauf der mathematischen Validierung.....	88
Abbildung 23: UML-Sequenzdiagramm zur Kommunikation zwischen Validierungs-Klasse, CommunicationSystem, GIS und SimulationManager	89

1 Einleitung

Dieses Kapitel gibt eine kurze Beschreibung von Gruppensimulationen und erklärt, warum das Validieren von diesen so wichtig ist und es nicht in der Forschung vernachlässigt werden darf. Es zeigt außerdem bereits einige der bei der Validierung auftretenden Probleme und dass es sich hierbei um keine trivialen Probleme handelt. Die in dieser Arbeit behandelten sowie nicht behandelten Punkte und das Ziel der Arbeit werden verdeutlicht.

1.1 Problem- und Themenbeschreibung

Mit Hilfe von Gruppensimulationen lässt sich das Verhalten von Individuen in großen Gruppen simulieren. Sie werden oft zum Simulieren von menschlichem Verhalten in bestimmten Situationen eingesetzt. Einer der wichtigsten Anwendungsbereiche ist hierbei die Entfluchtungsanalyse. Diese kann entweder makroskopisch betrachtet werden, indem Personenströme wie Flüssigkeitsströme simuliert werden, oder mikroskopisch in einer Gruppensimulation, in diesem Fall Evakuierungssimulation genannt, betrachtet werden. Dabei wird dann jede Person individuell simuliert. Diese Personen verfügen über individuelle Fähigkeiten, die ihr Verhalten charakterisieren (Meyer-König et al. 2009). Zu diesem Zwecke wird jedes Individuum in Gruppensimulationen als ein Agent modelliert. Evakuierungssimulationen simulieren die Evakuierung eines Ortes in einer Gefahrensituation und können auf diese Weise unter anderem zeigen, ob dieser in einem solchen Fall problemlos evakuiert werden kann oder wo Probleme auftreten.

Vor allem für Evakuierungssimulationen ist es sehr wichtig, dass die durch sie erlangten Ergebnisse möglichst nahe an der Realität liegen, da sie heutzutage oft eine entscheidende Rolle beim Entwurf und der Überprüfung der Sicherheit unter anderem von Gebäuden, Fahrgastschiffen und öffentlichen Plätzen spielen. Sollten die Ergebnisse einer solchen Simulation nicht korrekt sein, kann es zu schwerwiegenden Folgen wie Verletzten oder sogar Toten kommen.

Aber auch in anderen Gruppensimulationen sollen die Ergebnisse der Simulation möglichst nahe an der Realität liegen, sei es um in einer virtuellen Umgebung Menschenmengen realistisch darzustellen und dadurch den Spieler stärker in das Geschehen eintauchen zu lassen, so dass er realistisch auf die virtuellen Geschehnisse reagiert (Pelechano et al. 2008) oder um das Verhalten von Menschen in anderen Situation zu simulieren.

Um dies zu erreichen müssen die Gruppensimulationen validiert werden. Hierbei handelt es sich allerdings um eines der schwierigsten Probleme im Gebiet der Gruppensimulationen. Eines der Hauptprobleme ist aus Sicht vieler Forscher die nur geringe Menge an detaillierten Daten, mit denen die Ergebnisse der Simulationen verglichen werden könnten. Vor allem für die bereits erwähnten agentenbasierten Gruppensimulationen mit Entscheidungsfindung und unterschiedlichen Verhalten von Einzelnen ist dies ein großes Problem (Zhou et al. 2010).

Es gibt auch nicht das eine Verfahren zur Validierung, sondern viele verschiedene: „The difficulty for the validation task is that there is no [sic!] a universal approach“ (Xiang et al. 2005). Die Verfahren sind unter anderem davon abhängig, was simuliert wird und was wichtige Aspekte dieser Simulation sind, die somit nicht der Realität widersprechen dürfen.

1.2 Frage- und Aufgabenstellung

Viele Forscher haben sich mit den Möglichkeiten zur Validierung von Gruppensimulationen beschäftigt und Techniken entwickelt, um trotz der Probleme eine möglichst genaue Validierung zu ermöglichen.

Diese Arbeit klärt über den bisherigen Forschungsstand im Bereich der Validierung von Gruppensimulationen auf und präsentiert die einzelnen Ergebnisse. Des Weiteren werden allgemeine Validierungsmethoden aus dem Bereich der Validierung von Simulationen und dem Bereich der Validierung von Multiagentensimulationen auf ihre Tauglichkeit für die Validierung von Gruppensimulationen untersucht und vorgestellt. Auf diese Weise lässt sich ein Überblick darüber gewinnen, was es für Möglichkeiten für die Validierung von Gruppensimulationen gibt und wie man bei dieser vorgehen könnte.

1.3 Themenabgrenzung

Diese Arbeit beschäftigt sich nicht mit der Frage, wie Gruppensimulationen am besten implementiert werden sollten. Auch die Einsatzbereiche der Gruppensimulationen werden nur angeschnitten. Des Weiteren wird nicht auf Techniken zur Verifizierung von Gruppensimulationen eingegangen, es sei denn diese lassen sich nur schwer von den Validie-

rungstechniken abgrenzen. Das erstellte Validierungskonzept wird nur vorgestellt, eine Durchführung des Konzeptes findet nicht statt.

1.4 Ziel der Arbeit und Bedeutsamkeit

Ziel der Arbeit ist es, über die vorhandenen Möglichkeiten zur Validierung von Gruppensimulationen aufzuklären, diese zu bewerten und hierdurch in der Lage zu sein ein Validierungskonzept für Gruppensimulationen für die verschiedenen Phasen der Entwicklung zu erstellen. Dieses Konzept wird vorgestellt und es wird erläutert, welche Voraussetzungen die Gruppensimulation erfüllen muss, damit es angewandt werden kann. Hierdurch soll es Projekten, die an der Umsetzung einer eben solchen Gruppensimulation arbeiten, erleichtert werden, in den einzelnen Entwicklungsphasen zu validen Ergebnissen zu kommen, so dass letztendlich ein komplett validiertes fertiges System entsteht.

1.5 Struktur der Arbeit

Die Arbeit beginnt mit einem Überblick über den bisherigen Forschungsstand zum Thema Validierung von Gruppensimulation. Hierzu wird zuerst der Begriff Validierung definiert und andere Definitionen aus der Literatur genannt, da dieser Begriff in der Literatur durchaus unterschiedlich definiert wird. Danach wird auf die Bedeutsamkeit der Validierung eingegangen.

Als nächstes wird der Begriff Gruppensimulation genauer definiert und verschiedene Einsatzbereiche aufgezeigt. Auch hier wird auf die Bedeutsamkeit eingegangen. Des Weiteren wird ein kurzer Überblick über verschiedene Ansätze zur technischen Umsetzung von Gruppensimulationen gegeben.

Danach wird zunächst das generelle Vorgehen bei der Validierung beschrieben, das unabhängig von den verwendeten Verfahren zur Validierung ist. Als nächstes werden allgemeine Verfahren zur Validierung von Simulationsmodellen vorgestellt, welche sich auch für die Validierung von Gruppensimulation verwenden lassen. Dies wird mit den etwas spezielleren Verfahren zur Validierung von agentenbasierten Simulationen wiederholt. Zum Schluss dieses Teils wird dies auch für die ganz speziell auf Gruppensimulationen spezialisierten Verfahren durchgeführt.

Der folgende Teil beginnt zunächst mit einem Überblick über die verschiedenen Entwicklungsphasen der Entwicklung einer Gruppensimulation und die dazugehörige Validierung. Aus den vorgestellten Verfahren werden die wichtigsten Verfahren ausgewählt und die Anforderungen, der Aufwand und die Vor- und Nachteile erläutert und die Verfahren dementsprechend bewertet. Anhand dieses Teils wird ein konkretes Validierungskonzept für die

einzelnen Phasen der Entwicklung von Gruppensimulationen erstellt und präsentiert, welches verschiedene Verfahren kombiniert, um eine bestmögliche Validierung zu erlauben.

Für dieses Konzept wird anhand eines Beispiels beschrieben, welche Anforderungen das Konzept an eine Gruppensimulation und die Rahmenbedingungen stellt und wie die Validierung durchgeführt werden würde. Hierfür wird zunächst das WALK-Projekt vorgestellt und über den bisherigen Entwicklungs- und Validierungsstand aufgeklärt. Dieses Projekt dient als Basis dieses Teils und es wird beschrieben, wie dieses zu erweitern ist, damit das Konzept angewendet werden kann.

Zum Schluss werden die Ergebnisse zusammengefasst und es wird ein Fazit aus der Arbeit geschlossen und ein Ausblick gegeben, der zeigt, in welchen Bereichen noch Forschungsbedarf besteht, um eine noch verlässlichere Validierung zu ermöglichen.

2 Überblick über den Forschungsstand

Dieses Kapitel dient dazu, einen Überblick über den aktuellen Forschungsstand im Gebiet der Validierung von Gruppensimulationen zu geben. Hierzu wird zuerst genauer auf Validierung und Gruppensimulationen eingegangen. Danach werden allgemeine und spezielle Verfahren zur Validierung von Gruppensimulationen vorgestellt. Als nächstes werden Verfahren zur Gewinnung von Daten für die Validierung erläutert. Zuletzt wird auf in der Forschung bisher unzureichend abgedeckte Bereiche eingegangen.

2.1 Validierung

In diesem Unterkapitel wird zuerst die Verwendung des Begriffs der Validierung in der Literatur geklärt und verdeutlicht, inwiefern sich die Begriffe Validierung und Verifizierung voneinander unterscheiden. Danach wird ein Überblick über verschiedene in der Literatur genutzte Differenzierungen zwischen Validierungsarten gegeben und es werden die Arten der Validierung beschrieben. Zuletzt wird aufgezeigt, warum Validierung von Gruppensimulationen solch eine große Bedeutsamkeit besitzt.

2.1.1 Definitionen in der Literatur und Abgrenzung zur Verifizierung

Das Validieren von Modellen dient dazu, zu zeigen, dass sich das Modell in seinem Anwendungsbereich in Bezug auf die festgelegten Ziele des Modells und der Simulation mit zufriedenstellender Genauigkeit verhält. Daraus folgt, dass sich Modell-Validierung damit beschäftigt, das richtige Modell zu bauen (Balci 1997).

Diese Definition entspricht inhaltlich der Definition, auf die sich auch viele andere Autoren in ihren Arbeiten beziehen. Diese stammt von Schlesinger et al. und definiert Modell-Validierung als „substantiation that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model“ (Schlesinger et al. 1979).

Es ist unnötig, darauf zu achten, dass das Modell sein Vorbild möglichst genau abbildet. Viel mehr praktischen Sinn macht es, das Modell so zu modellieren, dass es nur soweit dem Vorbild entspricht, wie dies von den Zielen der Simulationsstudie gefordert wird. Validierung beantwortet also, ob es unmöglich ist, das Modell und sein Vorbild im Rahmen des experimentellen Interesses zu unterscheiden (Zeigler et al. 2000).

Wenn man über Validierung spricht, ist es auch wichtig den Begriff Verifizierung zu definieren, da beide Techniken bei der Überprüfung der Korrektheit eines Modells und seiner Resultate oft zusammen auftreten und deshalb verwechselt werden können.

Modell-Verifizierung befasst sich damit, zu zeigen, dass das Computerprogramm des Modells und seine Implementierung korrekt sind (Schlesinger et al. 1979). Somit geht es darum, das Modell richtig zu bauen (Balci 1997).

Man kann daher sagen, dass es bei der Validierung darum geht, ein fachlich korrektes Modell zu bauen, während es bei der Verifizierung darum geht, dieses technisch korrekt umzusetzen.

2.1.2 Arten der Validierung und Validität

Eine Simulation kann nicht nur auf eine Art validiert werden, sondern auf viele verschiedene Arten. Aus diesem Grund muss zwischen verschiedenen Arten der Validierung beziehungsweise der daraus folgenden Validität unterschieden werden.

In der Literatur zu diesem Thema lassen sich viele verschiedene Arten der Validierung finden. Die Art der Validierung ist hierbei davon abhängig, auf welche Weise die Validierung ausgeführt wird oder in welcher Phase sie durchgeführt wird (Klügl 2008).

Sargent unterscheidet außerdem beim Modellierungsprozess einer Simulation zwischen der konzeptuellen Modellvalidierung (conceptual model validation), der Validierung im Betrieb (operational validation) und der Datenvalidität (data validity). Das konzeptuelle Modell ist die mathematische, logische oder wörtliche Repräsentation der Problementität, welche ein System, eine Idee, eine Situation oder ein Phänomen darstellt, das modelliert werden soll (Sargent 2008).

Die konzeptuelle Modellvalidierung definiert er so, dass ermittelt wird, ob die Theorien und Annahmen, die dem konzeptuellen Modell zugrunde liegen, korrekt sind und ob die Darstellung der Problementität „sinnvoll“ für den geplanten Zweck des Modells ist. Zur Überprüfung, ob diese Theorien und Annahmen korrekt sind, sollten mathematische Analysen und statistische Methoden auf Daten der Problementität angewandt werden. Des Weiteren sollte überprüft werden, ob die Theorien korrekt angewandt wurden. Auch jedes Teilmodell

und das Gesamtmodell muss evaluiert werden, um zu bestimmen, ob es für den geplanten Einsatzzweck des Modells sinnvoll und korrekt ist. Hierzu gehört, ob mit angemessenem Detail gearbeitet wurde und ob geeignete Strukturen, Logik und mathematische und kausale Zusammenhänge benutzt wurden.

Die Validierung im Betrieb wird als das Ermitteln, dass das Ausgabeverhalten des Modells ausreichende Genauigkeit für den geplanten Zweck des Modells besitzt, definiert. Bei dieser Art der Validierung wird das Modell benutzt. Alle gefundenen Mängel können aus jeder Phase der Modellentwicklung stammen, zum Beispiel aus der Entwicklung der Systemtheorien und aus nicht validen Daten. Wie die Validierung im Betrieb durchgeführt wird, hängt davon ab, ob die Problementität oder das Referenzsystem beobachtbar sind. Ein System ist beobachtbar, wenn es möglich ist, Daten über das Verhalten der Problementität im Betrieb zu sammeln. Ist das Referenzsystem beobachtbar, so kann das Ausgabeverhalten des Modells mit dem Ausgabeverhalten des Referenzsystems verglichen werden, ansonsten ist nur ein Vergleich mit anderen validen Modellen möglich. Der Vergleich mit dem Referenzsystem kann über grafische Darstellungen erfolgen wie Histogramme, Boxplots oder Streudiagramme. Bei beobachtbaren und unbeobachtbaren Referenzsystemen sind jeweils auch Vergleiche mit dem Referenzsystem beziehungsweise anderen validen Modellen unter Zuhilfenahme von statistischen Tests möglich. Außerdem ist es bei beiden Varianten möglich, das Ausgabeverhalten des Modells mit geeigneten Validierungstechniken zu untersuchen.

Die Datenvalidität wiederum soll garantieren, dass die Daten für das Bauen, Auswerten und Testen des Modells sowie das Ausführen der Modellexperimente zum Lösen des Problems geeignet und korrekt sind. Es gibt allerdings nicht viel, was getan werden kann, damit Datenvalidität gewährleistet werden kann. Um valide Daten zu erhalten, sollten gute Methoden zum Sammeln von Daten und zum Pflegen dieser Daten entwickelt werden. Diese Daten sollten außerdem noch auf Ausreißer untersucht werden. Für gefundene Ausreißer muss überprüft werden, ob diese korrekt sind oder es sich zum Beispiel um Messfehler handelt.

Klügl nennt beispielhaft noch weitere Validierungsarten: Empirische Validierung (empirical validation), statistische Validierung (statistical validation), strukturelle Validierung (structural validation) und Prozessvalidierung (process validation). Sie macht allerdings auch klar, dass sich noch andere Arten der Validierung in der Literatur finden lassen (Klügl 2008).

Klügl benutzt in ihrer Arbeit eine Charakterisierung der Validität in zwei Dimensionen, eine Dimension, die sich auf die Art der Testmethode bezieht und eine andere Dimension, die sich auf das betroffene Element des Modells bezieht. Zur ersten gehören unter anderem Face Validity und empirische Validität und zur zweiten Verhaltensvalidität (behavioral validity) und strukturelle Validität (structural validity).

Face Validation umfasst alle Methoden der Validierung, die sich auf die natürliche Intelligenz des Menschen stützen. Die Simulation wird von Menschen auf unterschiedliche Art und Weise betrachtet, so dass diese mit Hilfe ihres theoretischen und impliziten Wissens die Prozesse und Ergebnisse auf Sinnhaftigkeit und Glaubwürdigkeit überprüfen können. Sie wird auch Glaubwürdigkeitsprüfung (plausibility checking) genannt.

In der empirischen Validierung werden mit statistischen Messungen und Tests die Kennzahlen, die vom Modell produziert wurden, mit denen des Vorbilds verglichen.

Verhaltensvalidierung überprüft, ob das Eingabe- und Ausgabe-Verhalten des Modells mit dem des Vorbilds übereinstimmt, während die strukturelle Validierung die Validität der internen Struktur des Modells, wie kausale Beziehungen zwischen Variablen oder das logische Denken der Agenten, mit der des Vorbilds vergleicht.

Die Arten der Validität lassen sich auch auf eine gröbere Weise betrachten. So fassen Zeigler et al. diese in drei Stufen zusammen: Replikative Validität (replicative validity), voraussagende Validität (predictive validity) und strukturelle Validität (Zeigler et al. 2000).

Replikative Validität ist die niedrigste Stufe und bedeutet, dass das Modell gegebene Daten des Vorbilds mit akzeptabler Abweichung reproduzieren kann.

Die voraussagende Validität befindet sich auf der zweiten Stufe und ist eine stärkere Form der Validität. Sie fordert neben replikativer Validität auch, dass bisher ungesehenes Systemverhalten des Vorbilds vorausgesagt werden kann.

Die dritte Stufe, strukturelle Validität, setzt voraus, dass das Modell nicht nur fähig ist, die beobachteten Daten des Vorbilds zu replizieren, sondern dass das Modell auch die Zustandsübergänge des Vorbilds imitiert.

2.1.3 Bedeutsamkeit

Die Validierung ist von außerordentlicher Wichtigkeit, weil nur durch sie bewiesen werden kann, dass das Richtige simuliert wird und die Ergebnisse der Simulation (bei korrekter technischer Umsetzung) den Ergebnissen der echten Welt ähnlich genug sind.

Vor allem bei Simulationen, von deren Korrektheit Menschenleben abhängen können, wie Evakuierungssimulationen, muss gewährleistet sein, dass die Simulation in den betrachteten Punkten möglichst genau der Wirklichkeit entspricht und somit für ihre Aufgabe geeignet ist. Auf diese Eigenschaft muss Verlass sein, wenn die Evakuierungssimulation beim Entwurf von Gebäuden, Fahrgastschiffen und anderen öffentlichen Plätzen eingesetzt wird, um die Sicherheit dieser zu garantieren.

Auch in anderen Bereichen trägt eine Validierung der Simulation zu ihrer Nutzbarkeit bei. So zum Beispiel wenn ein Spieler durch realistisches Verhalten der Agenten der Simulation in das Geschehen eintauchen kann und sich dadurch wie in der Realität verhält (Pelechano et al. 2008). Generell lässt sich also sagen, dass Validierung zur Akzeptanz der Simulation und zu richtigen Ergebnissen führt.

2.2 Gruppensimulationen

In diesem Unterkapitel wird beschrieben, was Gruppensimulationen sind und wo sich ihre Einsatzbereiche befinden. Des Weiteren werden die verschiedenen Arten von Gruppensimulationen aufgezeigt und beschrieben. Ein weiterer Teil befasst sich mit den technischen Möglichkeiten zur Umsetzung von Gruppensimulationen. Zuletzt wird erläutert, warum Gruppensimulationen heutzutage so bedeutsam sind.

2.2.1 Definition und Einsatzbereiche

Gruppensimulationen (crowd simulations) sind Simulationen, die das kollektive Verhalten von Individuen simulieren. Es handelt sich bei diesem Forschungsfeld um ein noch relativ junges, in welchem die meisten Arbeiten in der Mitte und den späten Neunzigern durchgeführt wurden. Gruppensimulationen werden in vielen verschiedenen Anwendungsgebieten benutzt, so zum Beispiel in der Architektur, Computergrafik, Physik, Robotik, Sicherheitswissenschaft, Trainingssysteme und Soziologie (Thalmann und Musse 2007).

Ein spezielles Augenmerk wird in dieser Arbeit auf Evakuierungssimulationen gelegt. Diese werden auch als simulationsgestützte Entfluchtungsanalysen bezeichnet. Die „Richtlinie für Mikroskopische Entfluchtungsanalysen“ (Meyer-König et al. 2009) beschreibt den Zweck dieser als „Bestimmung der Entfluchtungsdauer von baulichen Anlagen und Freiflächen auf denen der freie ungehinderte Personenstrom beeinflusst wird und [...] Überprüfung der Konzeption und Leistungsfähigkeit von Flucht- und Rettungswegen, insbesondere der Lokalisierung von Bereichen mit signifikanten Stauungen“. Als „Entfluchtung“ wird das „In-Sicherheit-Bringen“ von Personen aus einem gefährdeten Bereich definiert.

2.2.2 Arten von Gruppensimulationen

Die Ansätze zur Simulation von Gruppen unterscheiden sich meistens abhängig vom Einsatzbereich, da sich die verschiedenen Anwendungsgebiete auf andere Aspekte des kollektiven Verhaltens fokussieren. Daraus resultieren verschiedene Modellierungstechniken. So kann das Verhalten der einzelnen Individuen in der makroskopischen Entfluchtungsanalyse vernachlässigt werden und die ganze Gruppe durch Fluss- und Netzwerkmodelle simu-

liert werden, oder die Entfluchtung kann mikroskopisch simuliert werden, wobei die Individuen einzeln nach mehr oder weniger komplexen Regeln handeln. In anderen Anwendungsbereichen können die Individuen durch physikalische Gesetze, Chaosgleichungen, Verhaltensmodelle oder soziologische Simulationen kontrolliert werden (Thalmann und Musse 2007).

Prinzipiell lassen sich die Arten der Gruppensimulationen in zwei Hauptbereiche einteilen: In dem einen Bereich wird der Schwerpunkt der Simulation auf das realistische Verhalten der Individuen gelegt und in dem anderen wird der Schwerpunkt auf eine hochwertige Visualisierung gelegt.

Gruppensimulationen aus dem ersten Bereich besitzen oft nur eine einfache zweidimensionale Darstellung oder sogar keine Darstellung. Die Darstellung dient dem Verständnis der Resultate, ist aber nicht erforderlich. Die Verhaltensmöglichkeiten der Individuen sind meistens stark beschränkt und auf den Zweck zugeschnitten. Diese Art der Gruppensimulationen soll das Verhalten aus der Realität abbilden und wird in Evakuierungssimulationen, soziologischen Gruppenmodellen oder Modellen zur Gruppendynamik eingesetzt.

Im zweiten Bereich wird das Hauptaugenmerk auf die Visualisierung gerichtet. Das Ziel ist ein überzeugendes visuelles Ergebnis, der Realismus des Verhaltensmodells ist nur Nebensache. Solche Gruppensimulationen finden Anwendung in Filmen und Computerspielen. Eine realistische Visualisierung erfolgt nicht nur durch Verhaltensmodelle, sondern auch durch menschliche Eingriffe während des Produktionsprozesses. Die Gruppensimulation dient also als Hilfe für die Animatoren. Aus diesem Grund wird sich in diesem Bereich besonders auf das Rendering und die Animation konzentriert. Ein Beispiel für solch einen Einsatz sind heutzutage große Schlachtszenen in Filmen, da diese oft mit Gruppensimulationen erstellt werden. Hierfür gibt es unter anderem die Software „Massive“ (Massive Software 2011), welche ihren ersten Einsatz in der „Herr der Ringe“ Trilogie hatte.

Es gibt allerdings einen Trend, die beiden Bereiche zu verschmelzen, denn realistischere Verhaltensmodelle vereinfachen das Erstellen von visuell überzeugenden Ergebnissen. Vor allem im Bereich der Evakuierungssimulationen ist ein Trend zur besseren Visualisierung sichtbar. Des Weiteren ist sowohl realistische Nachbildung von Verhalten als auch eine hochwertige Darstellung in Trainingssystemen wichtig, um ein effektives Training zu ermöglichen (Thalmann und Musse 2007).

2.2.3 Technische Umsetzungen

Wie bereits erwähnt gibt es zwei verschiedene Ansätze zum Umsetzen einer Gruppensimulation, den mikroskopische Ansatz und den makroskopischen Ansatz. Makroskopische Modelle fokussieren sich auf das System als Ganzes, während mikroskopische Modelle sich mit

dem Verhalten und den Entscheidungen von Individuen sowie deren Interaktion mit den anderen Individuen der Gruppe befassen.

Der Bereich der makroskopischen Modelle umfasst Regressionsmodelle (regression models), Routenwahlmodelle (route choice models), Warteschlangenmodelle (queuing models) und kinetische Gas-Modelle (gas-kinetics models). Die Eigenschaften und Funktionsweisen dieser Modelle beschreiben Pelechano und Malkawi in ihrer Arbeit (Pelechano und Malkawi 2008).

Bei Regressionsmodellen werden mit Hilfe von statistisch fundierten Beziehungen zwischen Flussvariablen die Flussoperationen der Gruppe unter bestimmten Voraussetzungen vorhergesagt, wobei der Fluss von der Infrastruktur wie Treppen und Korridore abhängig ist.

Bei Routenwahlmodellen basiert die Wegfindung auf dem Konzept des Nutzens. Die Gruppenmitglieder wählen hierbei ihre Ziele so, dass der Nutzen des Weges besonders hoch ist. Der Nutzen kann sich unter anderem auf eine möglichst geringe Zeit oder einen möglichst einfachen Weg beziehen.

Warteschlangenmodelle benutzen Markov-Ketten-Modelle um zu beschreiben, wie sich die Gruppenmitglieder von einem Knoten zum nächsten Knoten eines Netzwerkes bewegen. Hierbei stellen Knoten Räume dar und die Verbindungen zwischen den Knoten können Türen und Durchgänge sein. Markov-Ketten-Modelle bestehen aus einer Menge an Zuständen mit Übergangswahrscheinlichkeiten. Bei jedem ausgeführten Schritt wird entweder ein Nachfolger aus der Übergangsverteilung ausgewählt oder es wird der wahrscheinlichste Nachfolger bestimmt.

In kinetischen Gas-Modellen wird die Bewegung der Gruppe mit der Bewegung von Flüssigkeiten und Gasen verglichen. Mit ihnen kann beschrieben werden, wie sich die Dichte und die Geschwindigkeit im Verlauf der Zeit verändert.

Im Gegensatz hierzu stehen die mikroskopischen Modelle. Mikroskopische Modelle sind agentenbasiert. Der agentenbasierte Ansatz ist besonders gut geeignet, da sich durch die Analogie zwischen den Agenten und den aktiven Elementen des Referenzsystems eine intuitive Struktur des Modells ergibt. Durch agentenbasierte Ansätze lassen sich Modelle in einem Detailgrad und in Anwendungsbereichen formulieren und behandeln, in denen dies vorher nicht möglich war (Klügl 2008).

Die mikroskopischen Modelle umfassen Modelle, die auf sozialen Kräften (social forces) beruhen, regelbasiert (rule-based) sind oder zelluläre Automaten (cellular automata) verwenden. Am verbreitetsten sind hierbei die Verwendung von sozialen Kräften und zellulären Automaten. Diese unterscheiden sich vor allem darin, dass beim Modell der sozialen Kräfte

der Raum als stufenlos angesehen wird und bei zellulären Automaten der Raum in ein Gitter aus gleichgroßen Zellen aufgeteilt ist. Auch für diese Modelle werden die Eigenschaften und Funktionsweisen beschrieben (Pelechano und Malkawi 2008).

In Modellen, die auf sozialen Kräften beruhen, ist die Bewegung eines Agenten in einem Raum von einem Mix aus sozio-psychologischen und physischen Kräften abhängig. Die Agenten wollen sich mit einer gewünschten Geschwindigkeit in eine bestimmte Richtung bewegen und passen ihre momentane Geschwindigkeit in einem bestimmten Intervall an. Außerdem versuchen die Agenten eine gewisse Distanz von anderen Agenten und Wänden zu halten. Hierdurch kommt es zu in der Wirklichkeit beobachteten Phänomenen, wie Bildung von Wölbungen an Durchgängen und dem Effekt, dass wenn die Agenten schneller laufen wollen, die Fortbewegung tatsächlich langsamer wird. Auch das Bilden von Schlangen und Gruppen ist vorhanden, da die Agenten dazu tendieren anderen Agenten zu folgen.

Regelbasierte Modelle finden vor allem in der Computergrafik Anwendung und werden dazu benutzt, um Schwärme und Herden von Tieren oder auch Gruppen von Menschen zu simulieren. Das bekannteste Modell ist das Boids-Modell von Reynolds, bei dem es sich um ein Partikelsystem mit den Boids genannten Entitäten handelt, die die Partikel darstellen. Jeder Agent ist unabhängig und navigiert abhängig von seiner lokalen Wahrnehmung der dynamischen Umgebung sowie den physikalischen Gesetzen, die seine Bewegung einschränken und einer Menge an Verhaltensweisen, die der Animator bestimmen kann. Dabei hält er sich an drei Regeln. Die erste Regel besagt, dass ein Agent so zu steuern hat, dass er bei seinen lokalen Nachbarn kein Gedränge auslöst. Die zweite Regel besagt, dass der Agent in die durchschnittliche Richtung der lokalen Nachbarn steuern soll. Die letzte Regel besagt schließlich, dass der Agent zur durchschnittlichen Position der lokalen Nachbarn steuern soll.

Das letzte erwähnte Modell, das Modell der zellulären Automaten, teilt den Raum in ein Gitter aus einheitlichen Zellen auf. Eine Zelle ist entweder frei oder besetzt und fasst nur einen Agenten. Jeder Agent befindet sich in einer bestimmten Zelle und bewegt sich von einer Zelle zur nächsten. Dies passiert in diskreten Zeitschritten. Der Wert einer Zelle ist von den Werten seiner Nachbar-Zellen abhängig. Die Werte der Zellen werden jeweils gleichzeitig aktualisiert und hängen von den Werten der Nachbar-Zellen im vorherigen Zeitschritt sowie von einer Menge lokaler Regeln ab.

2.2.4 Bedeutsamkeit

Durch Gruppensimulationen wird es ermöglicht, das Verhalten von Individuen in Situationen zu simulieren, die in der Realität nur schwer oder nicht umsetzbar sind. So lässt sich die Evakuierung von einem noch nicht gebauten Gebäude noch nicht in der Realität analysieren, sondern muss mit einer Evakuierungssimulation erfolgen oder per Handrechenregeln

berechnet werden, wobei letzteres die Individuen vernachlässigt. Hierdurch wird es ermöglicht, Gefahren in der Architektur zu erkennen und diese zu beheben.

In anderen Bereichen lässt sich durch Gruppensimulationen das Verhalten von Individuen in Gruppen beobachten und studieren oder sie können Animatoren bei der Animation von großen Massen von Individuen unterstützen und den Vorgang erleichtern.

2.3 Validierung von Gruppensimulationen

In diesem Unterkapitel wird aufgezeigt, wie man Gruppensimulationen validiert. Dazu werden zuerst generelle Vorgehensweisen zur Validierung vorgestellt. Danach werden allgemeinere Verfahren aus den Bereichen der Validierung von Simulationsmodellen und agentenbasierten Simulationen vorgestellt, die auch während des Entwicklungsprozesses einer Gruppensimulation Anwendung finden können. Außerdem werden die Probleme, die es nötig machen, spezielle Verfahren für agentenbasierte Simulationen und Gruppensimulationen anzuwenden, erläutert. Zum Schluss werden speziell für die Validierung von Gruppensimulationen entworfene Verfahren vorgestellt.

2.3.1 Generelle Vorgehensweisen

In der Literatur werden oft Validierung, Verifizierung und Akkreditierung zu VV&A zusammengefasst (Balci 1997). Ein ähnlicher Begriff ist VV&T, welcher für Validierung, Verifizierung und Testen steht (Balci 1994). Bei VV&A und VV&T handelt es sich nicht um eine Phase oder einen Schritt während des Entwicklungsprozesses, sondern um eine durchgehende Aktivität während des ganzen Entwicklungsprozesses. Hierdurch lassen sich die Fehler in der Phase, in der sie aufgetreten sind, finden und beheben und nicht erst wenn es zu spät ist, um noch etwas daran zu ändern.

Zu der Frage wer die Validierung durchführt gibt es verschiedene Ansätze. Am häufigsten verbreitet ist es, dass das Entwicklungsteam selbst auf Basis der Ergebnisse von Tests und Evaluierungen, die im Modellentwicklungsprozess durchgeführt wurden, entscheidet, ob das Modell valide ist. Hierbei handelt es sich um eine subjektive Einschätzung (Sargent 2008). Diese Vorgehensweise widerspricht jedoch der Feststellung von Balci, dass Validierung und Verifizierung Unabhängigkeit erfordert, um Voreingenommenheit der Entwickler zu verhindern.

Eine weitere Möglichkeit ist es, die Benutzer des Modells in die Entscheidung, ob das Modell valide ist, miteinzuziehen. Dies ist vor allem bei kleineren Entwicklungsteams sinnvoll und hilft außerdem bei der Glaubwürdigkeit des Modells (Sargent 2008). Allerdings ist auch diese Vorgehensweise nicht komplett von den Entwicklern unabhängig.

Um Unabhängigkeit der Validierung zu garantieren, gibt es den Ansatz (Sargent 2008) der „unabhängigen Verifizierung und Validierung“ (independent verification and validation). Hierbei wird eine dritte unabhängige Partei für die Entscheidung, ob das Modell valide ist, eingesetzt. Diese Partei ist sowohl von dem Entwicklungsteam als auch von den Auftraggebern oder Benutzern unabhängig. Die dritte Partei braucht allerdings ein genaues Verständnis des geplanten Ziels des Modells, um die Verifizierung und Validierung durchführen zu können. Wann die dritte Partei ihre Arbeit durchführt kann unterschiedlich sein. Diese kann parallel zur Entwicklung des Modells durchgeführt werden oder nachdem dieses entwickelt wurde.

Ist ersteres der Fall, so erhält das Entwicklungsteam von der dritten Partei während der Entwicklung des Modells Rückmeldungen zu der Verifizierung und Validierung. Das Entwicklungsteam sollte erst zur nächsten Entwicklungsphase fortschreiten, wenn das Modell die Verifizierungs- und Validierungsanforderungen der aktuellen Phase erfüllt. Dies ist die von Sargent empfohlene Vorgehensweise für den Einsatz einer dritten Partei.

Alternativ kann die Verifizierung und Validierung durch die dritte Partei nach der Fertigstellung des Modells durchgeführt werden. Die Aufgabe der dritten Partei kann nun unterschiedlich sein. Im einfachsten Fall kann sie die vom Entwicklungsteam durchgeführte Verifizierung und Validierung beurteilen. Sie kann allerdings auch eine nachträgliche komplette Verifizierung und Validierung durchführen. Dies kann allerdings sehr kostspielig und zeitaufwendig sein. Aus diesem Grund ist es sinnvoller, nur die bereits durchgeführte Verifizierung und Validierung überprüfen zu lassen.

Ein weiterer Ansatz zur Bestimmung, ob ein Modell valide ist, ist die Benutzung eines Scoring-Modells (scoring model). Dabei werden auf subjektive Weise verschiedene Aspekte des Validierungsprozesses mit Punkten bewertet und in Kategorien zusammengefasst. Ein Modell wird nun als valide angesehen, wenn es in allen Kategorien und in der Gesamtpunktzahl eine bestimmte vorher festgelegte Punktgrenze erreicht. Dieser Ansatz wird in der Praxis nur selten benutzt.

Ein Vertreter dieser Vorgehensweise ist Balci (Balci 1997). Seine Prinzipien zur Verifizierung, Validierung und Akkreditierung (VV&A) besagen, dass das Ergebnis der VV&A nicht nur als eine binäre Variable angesehen werden soll. Dadurch würde man beim Modell nur zwischen absolut korrekt oder komplett falsch unterscheiden.

Im Gegensatz dazu ist Sargent gegen den Einsatz eines Scoring-Modells. Dies begründet er mit fünf Punkten:

1. Ein Modell kann die Punktegrenze erreichen und trotzdem einen Fehler haben, der behoben werden muss.
2. Die Subjektivität des Ansatzes versteckt sich hinter den Werten und lässt die Ergebnisse objektiv erscheinen.
3. Die Punktegrenzen müssen meistens auf eine subjektive Weise festgelegt werden.
4. Durch die Bewertung kann ein zu hohes Vertrauen in das Modell entstehen.
5. Die Bewertungen können benutzt werden, um Modelle auf Basis dieser miteinander zu vergleichen und eines davon als besser anzusehen.

Sargent gibt auch eine Empfehlung, welche Schritte mindestens in einer Modellvalidierung durchgeführt werden sollten. Bevor das Modell entwickelt wird, sollte zwischen dem Entwicklungsteam, den Auftraggebern und, falls möglich, den Benutzern eine Abmachung geschlossen werden, die den allgemeinen Validierungsansatz und eine Minimalmenge an Validierungsprozessen spezifiziert. Danach sollte die geforderte Genauigkeit der für den Anwendungszweck des Modells interessanten Ausgabewerte festgelegt werden. Dies sollte vor der Entwicklung des Modells oder in einer frühen Entwicklungsphase geschehen. Außerdem sollten die Annahmen und Theorien, die dem Modell zugrunde liegen, wo immer möglich getestet werden. In jedem Schritt der Entwicklung des Modells (model iteration) sollte mindestens Face Validation auf dem konzeptuellen Modell durchgeführt werden. Das gleiche gilt für das Untersuchen des Verhaltens des Modells durch Nutzung des programmierten Modells (computerized model). Zumindest im letzten Schritt sollten dann, falls möglich, Vergleiche zwischen dem Verhalten und den Ausgabedaten des Modells und des Referenzsystems durchgeführt werden. Hierbei sollten mindestens einige verschiedene Bedingungen abgedeckt werden, am besten sogar viele verschiedene Bedingungen. Zuletzt schlägt Sargent noch vor, dass immer eine Validierungsdokumentation angefertigt wird. Sollte ein Modell des Weiteren für einen für einen längeren Zeitraum benutzt werden, so sollten periodisch erneute Überprüfungen der Validität durchgeführt werden.

2.3.2 Verfahren zur Validierung von Simulationsmodellen

In der Literatur lassen sich speziell für die Validierung von Gruppensimulationen erdachte Verfahren finden. Allerdings umfassen diese nicht immer die allgemeineren Verfahren für Simulationsmodelle. Aus diesem Grund werden in diesem Unterkapitel Verfahren zur Validierung von Simulationsmodellen erläutert, da diese sich meistens auch für Gruppensimulationen anwenden lassen.

Prinzipiell kann zwischen zwei Typen von Validierungsverfahren unterschieden werden: Subjektive Verfahren und objektive Verfahren. Subjektive Verfahren basieren auf subjektiven Einschätzungen, während objektive Verfahren mathematische Verfahren oder statistische Tests benutzen. Im Normalfall wird zur Validierung eine Kombination aus Verfahren eingesetzt (Sargent 2008).

Sargent gibt einen großen Überblick über mögliche Verfahren zur Validierung von Simulationsmodellen. Eine Möglichkeit zur Validierung stellt die Darstellung des Modells während der Ausführung dar. Hierzu nennt Sargent das Beispiel einer Simulation, bei der die Bewegung von Teilen durch eine Fabrik dargestellt wird. Diese Technik wird bei Gruppensimulationen meistens bereits eingesetzt, da sich so die Geschehnisse in der Simulation sehr viel leichter verfolgen lassen, als wenn nur reine Zahlenwerte ausgegeben würden. Die Darstellung der Geschehnisse kann auch einen wichtigen Punkt für die Face Validation darstellen.

Eine andere Möglichkeit, die grafische Darstellungen benutzt, ist das Anzeigen von im Zusammenhang mit dem Modell interessierenden Werten, zum Beispiel in Graphen, während die Simulation läuft. Dadurch wird das dynamische Verhalten der beobachteten Werte visualisiert und es kann nachvollzogen werden, ob sich diese korrekt verhalten. Bei Gruppensimulation könnte unter anderem die Dichte an Individuen an bestimmten Stellen, die Anzahl an Individuen in einem Bereich oder in einer Evakuierungssimulation die Anzahl der bereits evakuierten Personen angezeigt werden.

Bei der Face Validation werden Experten, die sich mit dem System auskennen, danach befragt, ob das Modell und dessen Verhalten sinnvoll erscheinen. So kann unter anderem überprüft werden, ob die Logik im konzeptuellen Modell korrekt ist und ob das Modell zu einer Eingabe die richtige Ausgabe erstellt. Für dieses Validierungsverfahren können von den Experten reine Daten verglichen werden, die als Zahlen oder Verlauf von Graphen dargestellt werden, oder die bereits erwähnte Darstellung des Modells während der Ausführung kann verwendet werden, was Bewertungen auf subjektiverer Basis ermöglicht. Dies kann hilfreich sein, wenn nicht genug konkrete Daten des Referenzsystems zur Verfügung stehen, wie dies bei Gruppensimulationen oft der Fall ist.

Eine ähnliche Möglichkeit zur Validierung ist der Turing Test. Bei diesem werden Experten, die sich mit dem Referenzsystem auskennen, gefragt, ob sie zwischen dem Referenzsystem und dem Modell unterscheiden können.

Eine weitere allgemeine Möglichkeit zur Validierung besteht im Vergleich des Modells mit anderen Modellen. Hierbei können Ergebnisse des Modells wie zum Beispiel Ausgaben validiert werden, indem sie mit anderen validen Modellen verglichen werden. Dies lässt sich auch für Gruppensimulationen anwenden, allerdings ist zu beachten, dass es vor allem bei Gruppensimulationen schwer ist zu bestimmen, ob das andere Modell tatsächlich valide ist, da die Validierung ansonsten falsche Ergebnisse liefert.

Liegen Aufzeichnungen von Daten aus dem Referenzsystem vor, so kann mit diesen eine Validierung des Modells durchgeführt werden. Hierzu wird ein Teil der Daten dazu verwen-

det, das Modell zu bauen und der Rest der Daten wird für die Überprüfung, ob sich das Modell so wie das Referenzsystem verhält, eingesetzt.

Degenerierungstests (degenerate tests) überprüfen, ob das Verhalten des Modells über einen längeren Zeitraum noch korrekt ist. Dazu müssen zuerst geeignete Werte für die Eingabe und die internen Parameter gewählt werden. Als Beispiel gibt Sargent hier eine Warteschlange eines Servers, bei der die Ankunftsrate von Anfragen höher ist als die Rate mit der diese abgearbeitet werden können. Die durchschnittliche Anzahl an Anfragen in der Warteschlange sollte also steigen.

Mit Ereignis-Validierung (event validation) wird das Vorkommen von Ereignissen im Modell mit dem des Referenzsystems verglichen. In Evakuierungssimulationen könnten Ereignisse wie die Evakuierung eines Agenten, das Auftreten von bestimmten Menschendichten oder auch Unfälle mit tödlichem Ausgang beobachtet werden.

Auch das Testen von Randbedingungen (extreme condition tests) gehört zur Validierung von Modellen. Hierzu wird das Modell auf Korrektheit auch bei Kombinationen von unwahrscheinlichen Randbedingungen überprüft.

Handelt es sich um ein stochastisches Modell, so sollte die interne Validität (internal validity) überprüft werden. Hierzu wird das Modell viele Male mit den gleichen Daten ausgeführt. Wird dabei eine hohe Variabilität der Ergebnisse festgestellt, das heißt es fehlt Konsistenz, so kann dies die Ergebnisse des Modells zweifelhaft erscheinen lassen. Ist dieses Verhalten für die Problementität typisch, so ist die Eignung der Verfahrensweise oder des Referenzsystems fragwürdig. Gruppensimulationen besitzen oft gewisse stochastische Eigenschaften, so dass auch hier überprüft werden muss, dass die Ergebnisse der Simulation, bei Evakuierungssimulationen zum Beispiel die Evakuierungszeiten, nicht zu weit voneinander entfernt sind.

Um zu prüfen inwiefern Änderungen an Eingabeparametern und internen Parametern das Verhalten des Modells und seine Ergebnisse beeinflussen, werden Sensitivitätsanalysen (sensitivity analysis) eingesetzt. Bei den Änderungen der Parameter sollten die gleichen Effekte auftreten wie in dem Referenzsystem. Auf diese Weise lässt sich auch herausfinden, ob Parameter überhaupt einen Einfluss auf das Verhalten und die Ergebnisse der Simulation besitzen oder ob sie möglicherweise sogar weggelassen werden können.

Bei der voraussagenden Validierung (predictive validation) wird das Modell zur Voraussage von Verhalten des Referenzsystems benutzt. Danach werden das Verhalten und die Ergebnisse des Referenzsystems mit der Voraussage verglichen.

Ein weiterer Ansatz zum Validieren von Simulationsmodellen ist das Verfolgen des Verhaltens von verschiedenen Typen von bestimmten Entitäten in Simulationsläufen, um zu bestimmen, ob die Logik des Modells korrekt ist und ausreichende Genauigkeit erreicht wird.

Des Weiteren gibt es mehrstufige Validierung (multistage validation). Diese kombiniert die Validierungsmethoden Rationalismus (rationalism), Empirismus (empiricism) und positive Volkswirtschaftslehre (positive economics). In ersterer wird davon ausgegangen, dass jeder bestimmen kann, ob die dem Model zugrunde liegenden Annahmen stimmen. Mit Hilfe von logischen Folgerungen lässt sich dann das korrekte Modell entwickeln. Im Empirismus dagegen müssen alle Annahmen und Ergebnisse empirisch validiert werden. Das Konzept der positiven Volkswirtschaftslehre wiederum besagt, dass ein Modell nur in der Lage sein muss die Zukunft vorauszusagen. Welche Annahmen und welche Struktur das Modell besitzt, das heißt kausale Zusammenhänge oder Mechanismen, spielt keine Rolle. In der mehrstufigen Validierung werden nun die Annahmen mit Hilfe von Theorie, Beobachtungen und dem Allgemeinwissen entwickelt, danach soweit es möglich ist empirisch getestet und die Zusammenhänge zwischen Ein- und Ausgabe des Modells mit dem Referenzsystem verglichen.

Aufgrund der Allgemeinheit der hier vorgestellten Validierungstechniken, lassen sie sich alle auch für Gruppensimulationen anwenden. Doch weitere konkretere Verfahren sind nötig, um die Validität von Gruppensimulationen zu gewährleisten.

2.3.3 Probleme bei der Validierung von agentenbasierten Simulationen

Die Validierung von agentenbasierten Simulationen ist nicht nur auf Grundlage der bisher vorgestellten Validierungstechniken für Simulationsmodelle möglich, sondern sie erfordert besondere Ansätze. Grund dafür sind einige Eigenschaften von agentenbasierten Simulationen, die problematisch bei der Validierung sein können. Diese Probleme beschreibt Klügl in ihrer Arbeit (Klügl 2008).

Ein großes Problem ist, dass für empirische und statistische Validierung charakteristische Nummern, die das System beschreiben, gefunden werden müssen. Für das Agentensystem im Groben ist dies meistens kein Problem, da hierfür zum Beispiel die Agenten gezählt werden können. Es ist jedoch problematisch, das individuelle Verhalten der Agenten, wie der gewählte Weg, auf eine allgemeine aber aussagekräftige Art und Weise zu beschreiben.

Des Weiteren gibt es noch keine Möglichkeiten Zeitabläufe zu validieren. Da agentenbasierte Simulationen allerdings oft für die Simulation von Dynamik, also der Bewegung von Agenten, eingesetzt werden, ist dies ein wichtiger Punkt. Ein Grund hierfür ist, dass brauchbare Daten zu Zeitabläufen kaum vorhanden sind. Es ist daher nur indirekt möglich, diese

Dynamik zu validieren, indem der resultierende Zustand in der Simulation mit dem des Referenzsystems verglichen wird.

Weitere Probleme ergeben sich dadurch, dass Agenten in agentenbasierten Simulationen Einfluss auf die Umgebung, andere Agenten oder zusammengefasste, organisatorische Strukturen haben können. Dadurch kann es zu chaotischen Effekten kommen, da Parameteränderungen nicht-lineare Auswirkungen haben. Solch ein Verhalten lässt sich nur schwer validieren, da minimale Ungenauigkeiten zu einem komplett anderen Ergebnis führen können.

Auch der Umfang der durchzuführenden Validierung ist ein Problem, da nicht nur Ein- und Ausgabeverhalten des Gesamtsystems mit dem Referenzsystem verglichen werden müssen, sondern dies auch für Untergruppen von Agenten und sogar einzelne Agenten durchgeführt werden muss. Hier liegt auch gleich das nächste Problem, denn es lassen sich zwar Daten für die empirische Validierung des Gesamtsystems finden, jedoch ist dies auf der Ebene der Agenten problematisch. So kann zum Beispiel der Personenfluss auf einfache Art und Weise gemessen werden, die Überwachung einer einzelnen Person jedoch erfordert größeren Aufwand und kann in manchen Ländern sogar datenschutzrechtliche Probleme machen. Um die Simulation auf der Agentenebene auf statistischer Grundlage zu validieren, sind diese Daten jedoch notwendig.

Ein letztes Problem ist, dass es möglich ist, dass Face Validation nicht durchführbar ist, weil das richtige Verhalten der Agenten nicht bekannt ist. Der Grund hierfür ist, dass, falls das Modell durch Überparametrisierung ein zu großes Maß an Freiheit besitzt, empirische Validierung nicht ausreicht, weil es durch Anpassung der Parameter immer möglich ist, das Modell auf die Daten abzustimmen. Somit entfällt die Möglichkeit, die Hypothese auf der das Modell beruht, zu widerlegen. Dies führt wiederum dazu, dass nicht entschieden werden kann, welche Hypothese korrekt ist. Aus diesem Grund wurde in den Sozialwissenschaften als Abhilfe das Konzept der generativen Hinlänglichkeit (generative sufficiency) entworfen, bei dem der Fokus in der Entwicklung von agentenbasierten Simulationen nicht mehr auf formaler Korrektheit liegt, sondern darauf, dass das Modell die beobachteten Phänomene generieren kann.

2.3.4 Verfahren zur Validierung von agentenbasierten Simulationen

Für die Validierung von agentenbasierten Simulationen empfiehlt Klügl daher ein aus vier verschiedenen Phasen bestehendes Vorgehen (Klügl 2008). Die in diesen Phasen angewandten Validierungsverfahren sind Face Validation, Sensitivitätsanalyse, Kalibrierung und statistische Validierung. Voraussetzung für dieses Vorgehen ist, dass die konzeptuelle Validierung und die Verifizierung bereits durchgeführt wurden und nun erstmals ein ausführbares Modell zur Verfügung steht, welches Animationen und Ausgabedaten erstellen kann. Da

agentenbasierten Simulationen stochastische Eigenschaften besitzen, müssen die Validierungsverfahren für verschiedene Simulationsläufe wiederholt werden.

Für die Validierung von agentenbasierten Simulationen gibt Klügl eine Grundstruktur vor. Begonnen wird mit der Face Validation. Für agentenbasierte Simulationen sollten nun mindestens drei Arten der Face Validation durchgeführt werden, bei denen der Fokus auf verschiedene Aspekte der Simulation gelegt wird. Bei diesen Arten handelt es sich um Animationsbewertung (animation assessment), Ausgabebewertung (output assessment) und immersive Bewertung (immersive assessment).

Bei der Animationsbewertung bewertet ein Experte die Animation des ganzen simulierten Systems oder von Teilen im Hinblick darauf, ob es sich wie das Referenzsystem verhält. Hierzu muss die Animation einen angemessenen Detailgrad besitzen und alle relevanten dynamischen Aspekte der Elemente auf eine Art und Weise darstellen, die für einen Menschen leicht wahrnehmbar ist. Durch die Animation lässt sich die Entwicklung des Systems verfolgen. Dazu wird oft auf eine Vogelperspektive gesetzt, weil dadurch zum Beispiel die Bewertung von Menschenflüssen und Staubildung im Vergleich zum Referenzsystem ermöglicht wird. Wird der Fokus auf einzelne Agenten gesetzt, so kann deren bestimmtes Verhalten bewertet werden.

Bei der Ausgabebewertung überprüft ein Experte die Plausibilität der absoluten Werte der Ausgabe, der Beziehungen zwischen den verschiedenen Werten sowie die Tendenz der verschiedenen Ausgabewerte von unterschiedlichen Simulationsläufen. Dies kann für das ganze System oder die einzelnen Agenten durchgeführt werden. Wenn es möglich ist, die vorgesehenen Beziehungen zu formalisieren, zum Beispiel in Form von Constraints, kann die Ausgabebewertung automatisiert werden.

Die letzte Art der Face Validation ist die immersive Bewertung. Hierbei betrachtet ein Experte das System durch die Augen eines bestimmten Agenten und kann somit wahrnehmen, was dieser wahrnimmt und wie dieser darauf reagiert. Aufgrund dieser Informationen kann er nun entscheiden, ob das Verhalten des Agenten realistisch war. Ist es dem Experten möglich den Agenten zu kontrollieren, so kann er des Weiteren die Reaktionen der anderen Agenten auf seine Aktionen bewerten.

Die Animationsbewertung und die immersive Bewertung sind einfacher durchzuführen als die Ausgabebewertung, da nur die Animation betrachtet werden muss. Aus diesem Grund sollten sie vor der Ausgabebewertung ausgeführt werden, weil so, falls festgestellt wird, dass das Modell nicht valide ist, die schwierigere und aufwendigere Ausgabebewertung nicht mehr durchgeführt werden muss.

Danach sollte eine Sensitivitätsanalyse durchgeführt werden. Dies ist für Modelle mit vielen verschiedenen Parametern wie agentenbasierten Modellen wichtig, um den Effekt der verschiedenen Parameter und deren Werte zu zeigen. Sollte sich durch die Sensitivitätsanalyse herausstellen, dass Parameter keinen Effekt haben, so können diese Parameter und mit ihnen verbunden Teile des Modells entfernt werden, so dass sich ein minimales Modell ergibt. Dies hilft unter anderem in Bezug auf Skalierungsprobleme durch eine große Anzahl an Parametern, Probleme beim Abbilden zwischen abstrakten Agentenparametern und konkreten kaum messbaren individuellen Parametern und anderen Parameterabhängigkeiten. Für die Sensitivitätsanalyse werden Parameter nach einem systematischen Ablaufplan verändert. Für jede Kombination von Parameterwerten muss mindestens ein Simulationslauf ausgeführt und bewertet werden, enthält das Modell stochastische Elemente sind mehrere Durchläufe nötig.

Das im vorherigen Schritt erstellte minimale Modell kann nun für den nächsten Schritt, die Kalibrierung, benutzt werden. Ziel der Kalibrierung ist es, die Parameter so zu setzen, dass das strukturell richtige Modell ein valides Ergebnis liefert. Hierbei handelt es sich um ein Optimierungsproblem, da aus einem gegebenen Eingabewert ein bestimmtes Ausgabeverhalten, das das des Referenzsystems ausreichend widerspiegelt, erzeugt werden soll und die dafür passenden Parameter gefunden werden sollen. Problematisch sind hierbei die Größe des Modells und fehlende Daten des Referenzsystems. Darüber hinaus ist es ein Problem, sinnvoll mit problematischen Parameter-Strukturen umzugehen. Wie die Kalibrierung ausgeführt wird, kann sehr unterschiedlich sein. So kann das ganze Modell als eine Blackbox betrachtet werden oder es können effiziente Optimierungsmethoden eingesetzt werden, um die „optimale“ Parameterbelegung zu finden. Sind Messungen für einzelne Agenten verfügbar, kann auch auf der Agentenebene kalibriert werden. Dann ist ein lokal optimales Agentenverhalten zum Beispiel durch Einsatz von verstärkendem Lernen (reinforcement learning) erreichbar, welches als Quelle für lokale Validierungsinformationen dienen kann, mit welchen wiederum lokale Kalibrierungen durchgeführt werden können. Hiernach sollte eine Plausibilitätsüberprüfung durchgeführt werden. Diese entspricht der Face Validation, muss allerdings nicht so intensiv ausgeführt werden.

Zuallerletzt wird eine statistische Validierung mit anderen Daten als den zuvor zur Kalibrierung benutzten Daten durchgeführt. Dies dient dazu, zu zeigen, dass das Modell nicht nur dazu optimiert ist, die gegebenen Daten zu reproduzieren, sondern auch valide Ergebnisse für neue Eingabewerte erzeugen kann.

Eine Fallstudie zur Validierung von agentenbasierten Simulationen haben Xiang et al. durchgeführt (Xiang et al. 2005). Diese umfasst die Bildung eines konzeptuellen Modells, die Implementierung sowie die Validierung eines agentenbasierten Modells zur Simulation von natürlichen organischen Materialien (NOM).

In dieser Fallstudie wird Face Validation bereits auf das konzeptuelle Modell angewandt. Hierzu überprüfen sechs Wissenschaftler aus verschiedenen Forschungsbereichen, in die das Thema fällt, das konzeptuelle Modell darauf, ob die aufgestellten Thesen sinnvoll erscheinen. Danach wird mit der Implementierung begonnen.

Für das implementierte Modell wird abermals Face Validation eingesetzt. Dieses Mal wird allerdings Visualisierung für die Validierung der Abläufe der Simulation eingesetzt. Das Verhalten und die Eigenschaften der Agenten und der gesamten Simulation werden hierbei mit den Beobachtungen aus dem Referenzsystem verglichen.

In der folgenden internen Validierung muss die statistische Integrität des Modells gezeigt werden, da dieses mit Zufallszahlen arbeitet. Verschiedene unabhängige Simulationsläufe mit den gleichen Eingabewerten sollten ähnliche Ergebnisse produzieren. Hierzu werden in dieser Fallstudie 450 Simulationsläufe mit verschiedenen Saaten für die Zufallszahlen ausgeführt und die Anzahl an Molekülen nach Abschluss der Simulation als Vergleichskriterien gewählt. Das Ergebnis sollte nun einer Normalverteilung entsprechen.

Für die Überprüfung der Validität im Betrieb werden zwei unterschiedliche Implementierungen des konzeptuellen Modells miteinander verglichen. Hierzu werden verschiedene zu vergleichende Variablen des Modells bestimmt und im Simulationsablauf aufgezeichnet. Wegen der stochastischen Eigenschaften des Modells werden nur die Mittelwerte aus 25 Simulationsläufen mit unterschiedlichen Saaten für die Zufallszahlen verglichen. Hierzu werden die Graphen des Verlaufs der Werte der Variablen miteinander verglichen. In der Fallstudie ergeben sich bei den unterschiedlichen Implementierungen verschiedene Graphen, die sich jedoch sehr ähnlich sehen. Die Autoren empfehlen die Durchführung von Tests der statistischen Signifikanz. Hierzu sollten Tests der Anpassungsgüte und Varianzanalysen durchgeführt werden, um sicherzustellen, dass die beobachteten Unterschiede auch statistisch unterschiedlich sind. Dies wird in der Fallstudie jedoch nicht durchgeführt.

Die Fallstudie kommt zu dem Schluss, dass der Vergleich von Modellen ein effektives Validierungsverfahren darstellt. Es fehlen jedoch trotz der wachsenden Popularität des agentenbasierten Modellierungsansatzes noch mathematische Grundlagen und eine formalisierte Validierungsmethodik. Die formalisierten Techniken kommen aus Bereichen, in denen die Systeme strukturiert sind, es eindeutige Regeln gibt und das Verhalten des Systems für die Modellierer transparent ist. Agentenbasierte Modelle werden jedoch oft für komplexe Systeme eingesetzt, welche unstrukturiert sind und bei denen ein unvollständiges Verständnis des Verhaltens des Systems bei den Modellierern vorhanden ist.

2.3.5 Probleme bei der Validierung von Gruppensimulationen

Bei der Validierung von Gruppensimulationen gibt es noch weitere Probleme als die Probleme, die durch die agentenbasierte Umsetzung begründet sind. Das bereits erwähnte Hauptproblem ist das Fehlen von ausreichend detaillierten Daten, mit denen der Vergleich der Ergebnisse des Modells mit denen des Referenzsystems möglich wäre. Vor allem im Bereich des Entscheidungsfindungsprozesses und dem Verhalten von Individuen fehlt es an diesen Daten. Sie werden jedoch für die Modellierung und Validierung von Gruppensimulationen wie Evakuierungssimulationen unter Berücksichtigung von psychologischen und sozialen Aspekten benötigt (Zhou et al. 2010).

Die Entscheidungsfindung kann großen Einfluss auf das Verhalten der Agenten haben, so etwa auf die Entscheidung welcher Fluchtweg bei einem Feuer gewählt wird, wie viele Personen einem Führer, der Anweisungen gibt, folgen würden und inwiefern die Personen den Informationen von Fremden Glauben schenken würden. Letztendlich beeinflusst dies die gesamten Bewegungen der Individuen in der Gruppe (Pelechano et al. 2008).

Das Fehlen dieser Daten liegt darin begründet, dass Daten zu psychologischen und sozialen Aspekten zwar in kontrollierten Experimenten gesammelt werden können, diese Daten aber zumeist nicht ausreichend sind, um die vielseitigen Verhaltensweisen in der Realität abzubilden. Des Weiteren ist es meistens der Fall, dass die Teilnehmer der Experimente emotional entspannt wirken. Dies kann dazu führen, dass sie sich anders als in einer realen Situation, wie zum Beispiel einer Evakuierung im Falle eines Feuers, verhalten (Zhou et al. 2010). So haben Studien nachgewiesen, dass steigender Stress die Wachsamkeit reduziert und dadurch vor allem die Orientierung in einem Gebäude erschwert. Auch sorgt Zeitdruck für eine Steigerung der Geschwindigkeit von Aktionen und eine subjektive Wahrnehmung von Informationen. In einer Notfallsituation wird die Fähigkeit Entscheidungen zu treffen durch viele verschiedene Faktoren beeinflusst wie die Unsicherheit, welche Änderungen die Umgebung durch die Situation erfahren kann und dem Zeitdruck. Ein weiterer Grund für Stress können zu viele oder zu wenige Informationen sein. So kann es sein, dass viele verschiedene Leute in einem Raum unterschiedliche Entscheidungen treffen und unterschiedliche Informationen von sich geben. Außerdem bestehen komplexe und sich dynamisch ändernde Situationen (Pelechano und Malkawi 2008). Das realistische Nachstellen von solchen Situationen in Experimenten kann entweder sehr schwer oder wie im Falle des Feuers sogar unmöglich sein (Pelechano et al. 2008).

Die psychologischen Aspekte, die die Routenwahl und das Verhalten der Personen beeinflussen, sind noch nicht wissenschaftlich fundiert untersucht worden und können deshalb nur durch statistische Verhaltensweisen implementiert werden (Meyer-König et al. 2009). Aus diesem Grund wird bei der Validierung von Gruppensimulationen oft die Entscheidungsfindung in gefährlichen Situationen nicht berücksichtigt und nur mit Hilfe von Kon-

struktionsrichtlinien die geregelte Evakuierung validiert (Pelechano et al. 2008). Es gibt jedoch auch Ansätze zur Validierung der Entscheidungsfindung in solchen Situationen.

Auch das Alter, die ethnische Herkunft und andere Gruppenzugehörigkeiten können das Verhalten der Individuen und Gruppe beeinflussen. Menschenmassen bestehen aus einer Mischung von unterschiedlichen Individuen (Malone et al. 2008). Studien haben zudem ergeben, dass Hausbesitzer bei der Evakuierung ihres Hauses die ihnen bekannten Wege den Fluchtwegen vorziehen (Pelechano und Malkawi 2008). Somit gibt es viele verschiedene Faktoren, die den Ablauf einer Gruppensimulation beeinflussen können und somit auch die Validierung erschweren.

Ein weiteres Problem liegt im Bereich der Face Validation von Simulationen. Bei der Face Validation handelt es sich um einen verbreiteten Ansatz zur Validierung von Modellen. Vor allem wenn nicht genügend konkrete Daten vorhanden sind, wird sie oft eingesetzt. Allerdings sind die Beurteilungen der Experten stark von deren individuellem Wissen und Erfahrungen mit ähnlichen oder sogar sehr verschiedenen Situation aus der Realität, zum Beispiel Evakuierungen, abhängig. Diese können einen starken Einfluss auf die Validierungsergebnisse haben (Zhou et al. 2010). Face Validation erzeugt somit stark subjektive und fragwürdige Ergebnisse (Banerjee und Kraemer 2010).

2.3.6 Verfahren zur Validierung von Gruppensimulationen

In diesem Unterkapitel werden verschiedene Verfahren zur Validierung von Gruppensimulationen vorgestellt. Diese umschließen den Vergleich der Simulationsergebnisse mit Experimenten, die Face Validation in verschiedenen Ausführungen, Verfahren unter Benutzung von Virtual Reality, mathematische und statistische Verfahren sowie Validierung auf Basis von fest vorgeschriebenen Tests einer Richtlinie. Die zuerst vorgestellten Verfahren sind subjektiver und erlauben ein offeneres Vorgehen, während die später vorgestellten Verfahren sehr enge Grenzen setzen und sehr konkrete Verfahren darstellen.

Vergleich mit Übungen

Begonnen wird mit einem Beispiel zum Vergleich der Simulation mit Übungen. Das Beispiel wurde von Klüpfel durchgeführt, welcher in seiner Arbeit die Implementierung und Validierung einer Evakuierungssimulation behandelt hat (Klüpfel 2003). Für die Validierung der Simulation wurden verschiedene Evakuierungsübungen durchgeführt, in der Simulation nachgebildet und ebenfalls ausgeführt und schließlich die Ergebnisse verglichen. Das Beispiel ist eine von drei Übungen, die zum Vergleich benutzt wurden.

Die Übung wurde in einem Multiplex-Kino durchgeführt. Bei den Teilnehmern handelte es sich um 100 Studenten, die einen Hut mit einer Nummer tragen mussten, damit sie auf den Videos, welche manuell ausgewertet wurden, identifizierbar waren. Die Teilnehmer wuss-

ten über die Evakuierungsübung Bescheid und es waren keine Gefahren vorhanden. Außerdem wurden die Teilnehmer aufgefordert vorsichtig zu sein, um Verletzungen zu vermeiden. Hierdurch ergibt sich allerdings eine Situation, in der sich nur bedingt das Verhalten im Falle einer echten Evakuierung nachstellen lässt, da die Teilnehmer wussten was auf sie zukommt und somit darauf vorbereitet waren und kein Stress vorhanden ist.

Die Sitzposition der einzelnen Teilnehmer wurde aufgenommen und somit der Startpunkt der Evakuierung für den Teilnehmer bestimmt. Während die Werbung im Kinosaal lief wurde nun ein Alarm ausgelöst und die Teilnehmer zum Verlassen des Gebäudes durch die markierten Ausgänge und Sammeln an einer bestimmten Platz außerhalb des Gebäudes aufgefordert. Daraufhin fingen die Teilnehmer mit der Evakuierung an. Eine Person wurde als evakuiert angesehen, wenn sie einen bestimmten Platz erreicht hatte. Die Teilnehmer wurden nun während der Evakuierung mit verschiedenen Kameras überwacht. Es gab zwei mögliche Wege, die die Teilnehmer nehmen konnten.

Das gleiche Szenario wurde in der vorher implementierten Evakuierungssimulation nachgestellt (Abb. 1). Nun wurden gemessene Zeiten der Übung für die gesamte Evakuierung sowie für die beiden Wege mit den Zeiten der Simulation verglichen. Betrachtet wurden jeweils die Gesamtzeit der Evakuierung, die durchschnittliche Zeit für die Evakuierung einer Person sowie der Medianwert. Ein weiterer Vergleich fand über die Darstellung der Anzahl der evakuierten Personen im Zusammenhang mit der vergangenen Zeit in einem Graphen statt. Des Weiteren wurden die Aufzeichnungen der Kamera mit den Bildern der Simulation verglichen, wodurch der Vergleich der Menschendichte zu einer bestimmten Zeit an einem bestimmten Ort und Staus beobachtet und verglichen werden konnten. Zuletzt wurden auch die individuellen Evakuierungszeiten der Teilnehmer sowie deren Wahl des Ausgangs des Kinosaals, von denen einer im vorderen und einer im hinteren Teil vorhanden waren, zwischen der Übung und der Simulation verglichen.

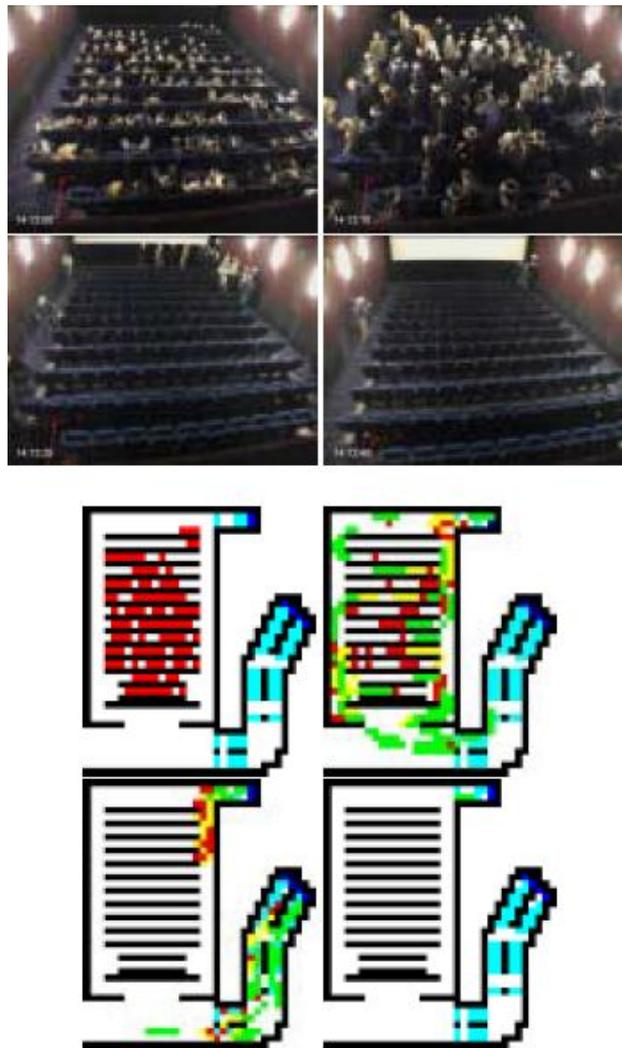


Abbildung 1: Aufnahme und Simulation zu jeweils gleichen Zeitpunkten (Klüpfel 2003)

Bei allen Vergleichen wurde, wenn Unterschiede zwischen Übung und Simulation auftraten, nach den Ursachen gesucht. So wurde unter anderem festgestellt, dass sich die Staus in der Simulation langsamer auflösten als in der Übung. Außerdem wählten die Personen in der Simulation die Ausgänge nur abhängig von der Geometrie während die Teilnehmer der Übung ihre Entscheidungen auch von den Staus vor den Ausgängen abhängig machten. Somit ließ sich feststellen, wo noch nicht ausreichend behandelte Bereiche in der Simulation waren. Außerdem wurden durch die Vergleiche die Auswirkungen der Entscheidungsfindung gezeigt. Eine Validierung dieser wurde aber nicht durchgeführt, da sie in der Simulation nicht berücksichtigt wurde.

Webbasierte Face Validation

Ein rein auf Face Validation basierendes Verfahren stammt von Fridman und Kaminka (Fridman und Kaminka 2010). Auch hier werden Videoaufnahmen zur Validierung benutzt, jedoch werden keine Experten für die Face Validation eingesetzt, sondern es findet ein webbasierter Ansatz Anwendung.

Bei diesem Verfahren werden Videoaufnahmen von verschiedenen Situationen, in denen Fußgänger unterwegs sind, mit Videos von verschiedenen Simulationen, die dieses Verhalten ebenfalls zeigen sollen, verglichen. Damit diese Vergleiche nicht von einem bestimmten Video abhängig sind, werden verschiedene Aufnahmen der echten sowie simulierten Fußgänger angefertigt. Die Videos werden jeweils auf eine Minute zurechtgeschnitten.

Nun wurde ein webbasiertes Experiment ausgeführt, an dem die Teilnehmer in ihrer Freizeit teilnehmen konnten. Die Teilnehmer hatten die Aufgabe, die ihnen gezeigten Aufnahmen der Simulation (Abb. 3) mit der dazugehörigen Aufnahme der echten Fußgänger zu vergleichen (Abb. 2). Dabei wurden nicht die einzelnen Fußgänger verglichen sondern das allgemeine Fußgängerverhalten.

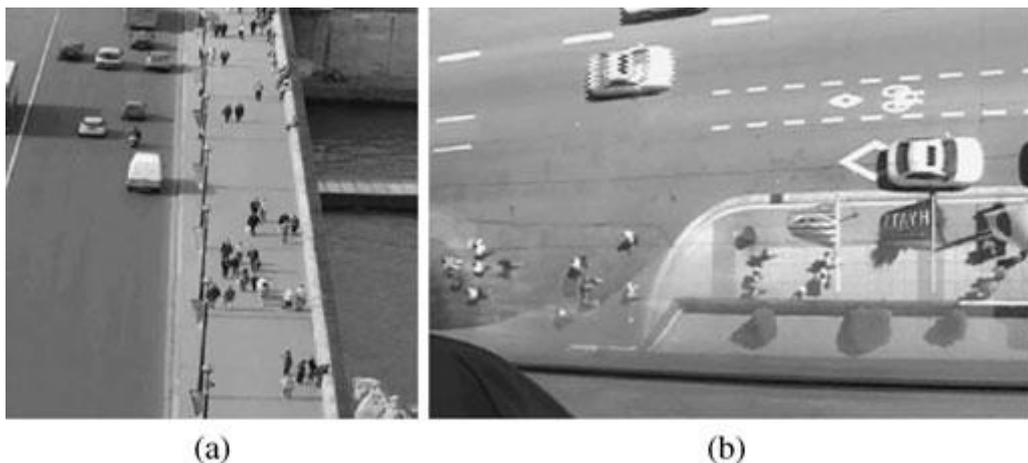


Abbildung 2: Die realen Aufnahmen des Szenarios (Fridman und Kaminka 2010)

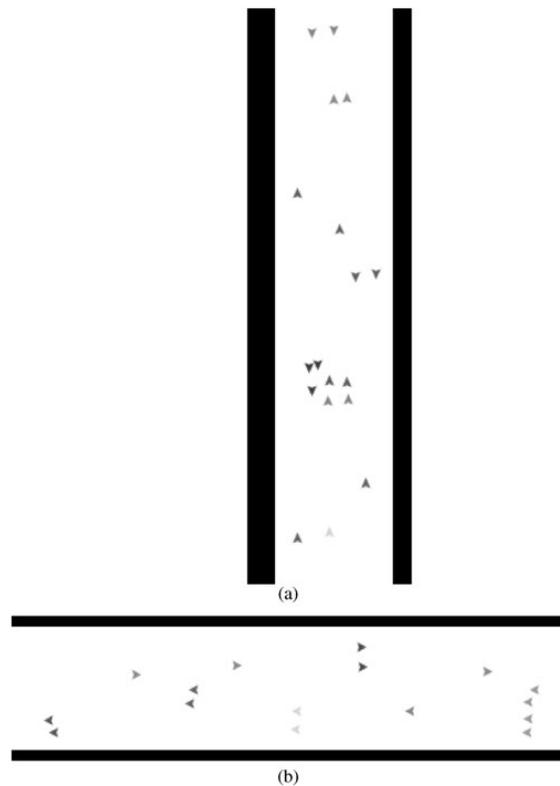


Abbildung 3: Das Szenario der Aufnahme in der Simulation (Fridman und Kaminka 2010)

Für die Teilnehmer begann das Experiment mit einer Beschreibung des Experiments und einer Trainingsphase. Nach dieser wurden den Teilnehmern zufällig ausgewählte Aufnahmen einer Situation gezeigt und zufällig ausgewählte Aufnahmen der Modelle, die das gleiche Verhalten darstellen sollten. Die Reihenfolge, in der die verschiedenen Aufnahmen gezeigt wurden, war ebenfalls zufällig, so dass sie keinen Einfluss auf den Ausgang des Experiments hatte. Die Teilnehmer mussten nun auf einer Skala von eins bis sechs bewerten, inwieweit die Modelle das Verhalten der echten Videoaufnahme widerspiegelten. Dazu konnten sie so oft wie sie wollten zwischen den Aufnahmen umschalten und ihre Bewertungen anpassen. Zuletzt wurden die Teilnehmer danach befragt, welches Modell nach ihrer Meinung das Verhalten am besten beziehungsweise am schlechtesten simuliert.

Virtual Reality und Face Validation

Ein anderes Verfahren, das auf den Einschätzungen von Teilnehmern eines Experimentes beruht, wurde von Pelechano et al. entwickelt (Pelechano et al. 2008). In diesem Experiment befinden sich die Teilnehmer in einer virtuellen Umgebung und sind von computergesteuerten Agenten umgeben. Sie tragen ein Head-Mounted Display und Sensoren, die ihre Blickrichtung bestimmen (Abb. 4).



Abbildung 4: Eine Testperson mit Head-Mounted Display und Sensoren (Pelechano et al. 2008)

Die Idee des Experiments beruht darauf, dass die Teilnehmer genauso wie in einer realen Situation reagieren, wenn das Geschehen um sie realistisch erscheint und sie sich dadurch in die nachgestellte Situation hineinversetzen können. Auf diese Weise lässt sich einerseits bestimmen inwieweit die vorgestellten Modelle diese Anforderung erfüllen, was einen Teil der Validierung darstellt. Andererseits können durch die daraus resultierenden realistischen Reaktionen der Teilnehmer auch Daten zu diesen in Experimenten gesammelt werden, die sonst so nicht möglich wären. So in etwa die Reaktionen bei einer Evakuierung im Falle eines Feuers. Im Gegensatz zu den bisher vorgestellten Verfahren lässt sich also erstmals auch die Entscheidungsfindung in gefährlichen Situationen validieren. Hierfür muss allerdings zuerst eine Simulation und Umgebung vorhanden sein, die eine starke Immersion, das „Hineinversetzen“, der Teilnehmer in die Situation erlauben. Ist dies der Fall, so können Simulationsläufe durchgeführt werden, um menschliches Verhalten zu erforschen und die gewonnenen Daten zur Validierung und Verbesserung des Modells eingesetzt werden.

Die realistischen Reaktionen auf das Geschehen in der Simulation beziehen sich sowohl auf die physiologische Ebene als auch auf die Verhaltensebene. Auch das emotionale und kognitive Verhalten gehören hierzu.

Wie stark der Teilnehmer sich in die Situation hineinversetzt lässt sich auf verschiedene Weisen messen. Die klassische Variante benutzt Fragebögen, die der Teilnehmer nach dem Experiment ausfüllen muss. Allerdings ist diese Art der Messung sehr stark von der subjektiven Sichtweise des Teilnehmers abhängig. Aus diesem Grund wurden andere Messmethoden entwickelt. So ist es unter anderem möglich, die Stärke der Immersion durch die sozialen Reaktionen oder Änderungen der Haltung eines Teilnehmers zu messen. Physiologische Messungen können zum Beispiel die Herzfrequenz bestimmen. Eine weitere Möglichkeit ist die Bestimmung der elektrodermalen Aktivität. Bei dieser wird das Absinken des Leitungswiderstands der Haut gemessen, das durch erhöhte Schweißsekretion bei emotional-affektiven Reaktionen ausgelöst wird. Andere Messungen umfassen die Bestimmung von Zeiten, die für eine Aufgabe benötigt werden, und von Fehlerraten zur Bestimmung der Leistung bei der Bearbeitung von Aufgaben sowie das Zählen von „Brüchen“, in denen der Teilnehmer sich nicht mehr in die Situation hineinversetzt.

Eine stärkere Immersion des Teilnehmers kann unter anderem dadurch erreicht werden, dass der Teilnehmer Objekte in der virtuellen Umgebung manipulieren kann und auch mit den virtuellen Menschen kommunizieren kann. Wird auf eine unnatürliche Weise mit der virtuellen Umgebung interagiert, zum Beispiel durch eine Joystick-Steuerung, so kann dies ebenfalls einen Einfluss haben. Realistischere Interaktionsweisen wie das „Laufen auf der Stelle“ erzeugen bessere Ergebnisse. „Brüche“ können unter anderem durch Einflüsse aus der echten Umgebung ausgelöst werden, wie der Zusammenstoß mit einer Wand, Stolpern und Whiteouts. Whiteouts entstehen zum Beispiel dadurch, dass sich der Teilnehmer durch einen Agenten hindurch bewegt. Dadurch entsteht der gleiche Effekt, als wäre die virtuelle Umgebung plötzlich verschwunden. Auch die Technik zur Darstellung hat einen Einfluss. So sorgen niedrige Frameraten, ein geringeres Sichtfeld als in der Realität und eine niedrige Auflösung für schlechtere Resultate. Auch fehlende oder zu geringe physische und akustische Rückmeldungen bei Kollisionen verschlechtern die Resultate.

Weitere Effekte auf die Immersion haben die Art der Fortbewegung der Agenten im Modell und die Reaktion auf andere Agenten. So scheinen in Modellen, die auf sozialen Kräften beruhen, die Agenten zu vibrieren, während sie versuchen sich zu bewegen, was einen negativen Einfluss auf den Realismus der Darstellung der Simulation hat. Dagegen haben Modelle, die auf zellulären Automaten beruhen, das Problem, dass die Fortbewegung nicht kontinuierlich möglich ist, sondern nur zwischen miteinander verbundenen Zellen erfolgt. Hierdurch wird unter anderem auch die Möglichkeit der Bewegungsrichtung des Agenten eingeschränkt. Ein weiteres Problem dieser Modelle ist außerdem, dass Überschneidungen der Agenten nicht verhindert werden. Es kann zwar nur ein Agent in einer Zelle stehen, wenn sich jedoch zwei Agenten jeweils gleichzeitig in die Zelle des anderen Agenten begeben wollen, wählen sie dessen Nachbarzelle, da seine Zelle belegt ist. Dadurch schneiden sich die Bewegungsbahnen der Agenten. Weitere Probleme ergeben sich, wenn die Modelle Agenten nicht die Möglichkeit geben, andere Agenten zu schieben und wenn die Agenten

sich nicht gegenseitig über die Umgebung austauschen können und anderen keine Anweisungen geben können. Alle diese Probleme sorgen für eine geringere Realitätsnähe des Modells und sollten deshalb vermieden werden.

Statistische Validierung

Bei den bisher vorgestellten Verfahren handelt es sich um Verfahren, die auf Vergleichen und subjektiven Bewertungen basieren. Es gibt allerdings auch objektivere Verfahren zur Validierung von Gruppensimulationen. Die Arbeit von Malone et al. beschreibt ein Verfahren, das auf statistischer Validierung basiert und somit nicht von den Personen, die die Validierung durchführen, abhängig ist (Malone et al. 2008).

Bei dem Verfahren handelt es sich um ein multivariates Verfahren, das heißt, dass mehrere statistische Variablen gleichzeitig betrachtet werden. Die in diesem Fall untersuchten Variablen können bei einer Gruppensimulation mit Personen die durchschnittliche Dichte an Personen an den Ausgängen, Anzahl an Personen, die eine bestimmte Tür durchschritten haben, durchschnittliche Orientierung aller Individuen, durchschnittliche X-Position aller Individuen, durchschnittliche Y-Position aller Individuen, durchschnittliche Geschwindigkeit in X-Richtung, durchschnittliche Geschwindigkeit in Y-Richtung, durchschnittliche Dichte um jedes Individuum, Anzahl der Individuen, die in Richtung der Tür blickt und Anzahl der Individuen, die den Raum verlassen haben, sein. Anhand dieser Werte soll die Zusammensetzung einer Gruppe bestimmbar sein. Dies liegt unter anderem darin begründet, dass Eigenschaften wie die Geschwindigkeit einer Person, die Zufälligkeit ihrer Bewegungen und wie dicht die Person anderen Personen oder Hindernissen kommen will, als vom Alter abhängig angesehen werden können.

Alle diese statistischen Variablen sollen 29-mal pro Sekunde von der Simulation berechnet und gespeichert werden können. Dies entspricht der Bildfrequenz der am weitesten verbreiteten Videostandards und hilft dadurch beim Vergleich mit Videoaufnahmen. Da allerdings noch nicht genügend Videodaten zum Testen des Verfahrens vorhanden waren, wurde es bisher nur in Vergleichen von Simulationen erprobt.

Um das Verfahren trotzdem durchführen zu können, wurde eine zufällig bestimmte Menge von Bildern aus der Simulation ausgewählt und die dazugehörigen statistischen Werte erlangt. Hierzu wurden die Werte gemessen und berechnet und schließlich für alle betrachteten Bilder gemittelt, so dass für einen Simulationslauf ein einziger mehrdimensionaler Datenpunkt entstand. Dies wurde insgesamt 60-mal mit Simulationen durchgeführt, die die gleichen Parameter besaßen und sich nur in der Saat unterschieden, so dass schließlich 60 Datenpunkte vorhanden waren. Diese Daten waren die historische Datenmenge. Danach wurden die Werte von Parametern der Simulation soweit verändert, dass Simulationen entstanden, die sich im Aussehen unterschieden. Diese Simulationen konnten nun mit der historischen Datenmenge verglichen werden, wobei das Verfahren für Simulationen mit

den gleichen Parametern keine Unterschiede feststellen sollte und für Simulationen mit unterschiedlichen Parametern diese feststellen sollte.

Zum Prüfen inwieweit die statistischen Variablen eines Simulationlaufes mit den Vergleichsdaten übereinstimmen, wird bei diesem Verfahren das Prinzip der simplizierten Tiefe benutzt. Hierzu werden die Vergleichsdaten der historischen Datenmenge in einer Datenwolke angeordnet. Nun werden jeweils $d + 1$ der Datenpunkte der Vergleichsdaten miteinander verbunden, wobei d die Dimension der Werte bestimmt. Die Tiefe des Datenpunktes, der die statistischen Variablen des zu vergleichenden Simulationlaufes einer Simulation darstellt, wird dadurch bestimmt, in wie vielen der durch das Verbinden entstandenen Bereiche er liegt (Burr et al. 2004a).

Ein Beispiel (Abb. 5): Bei 2-dimensionalen Werten müssten jeweils immer 3 Datenpunkte der Datenwolke verbunden werden. Bestünde die Datenwolke aus 6 Punkten, so gäbe es 20 durch das Verbinden entstehende Bereiche. Ein Datenpunkt, der nun in 6 dieser Bereiche liegt, hat somit eine simplizierte Tiefe von $6/20$ also $0,3$ (Burr et al. 2004b).

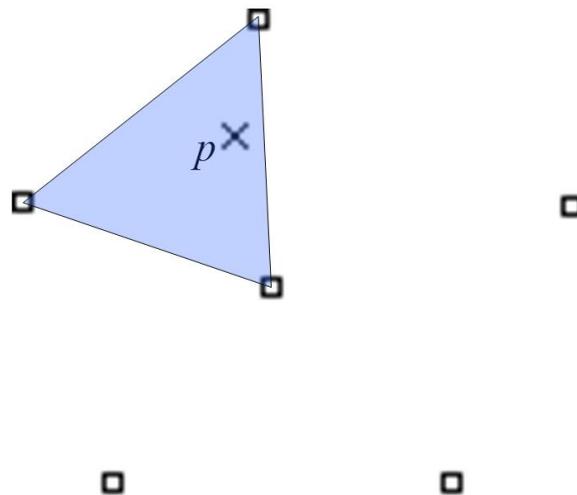


Abbildung 5: Einer der Bereiche in denen der Punkt p liegt (Burr et al. 2004b)

Die Tiefe des Datenpunktes gibt an, wie zentral er in der historischen Datenmenge liegt. Je höher der Wert der simplizierten Tiefe des Datenpunktes ist, desto zentraler liegt er. Der tiefste Datenpunkt ist der zentralste Punkt der Verteilung, der Median (Härdle und Simar 2004). Durch Hauptkomponentenanalyse lässt sich die Komplexität dieser Berechnung verringern, indem die Datensätze dadurch vereinfacht werden, dass eine Vielzahl von statistischen Variablen durch eine möglichst geringe Anzahl von Linearkombinationen, den Hauptkomponenten, ersetzt wird. Diese wurden in diesem Experiment durch die Eingabe der historischen Datenmenge in ein statistisches Softwarepaket bestimmt, so dass die Anzahl der Dimensionen auf 3 verringert werden konnte.

Die verschiedenen durch Parameteränderungen entstandenen Simulationen wurden nun jeweils 40-mal ausgeführt. Auf die gleiche Weise wie für die historischen Datenpunkte wurde erneut pro Simulation und pro Simulationslauf ein Datenpunkt berechnet. Zur Durchführung des Vergleiches wurde für jeden Datenpunkt einer Simulation die simplifizierte Tiefe des Punktes in der historischen Datenmenge bestimmt. Zusätzlich wurde ein bestimmter Wert festgelegt, ab dem der Datenpunkt als Ausreißer beziehungsweise eine andere beobachtete Simulation, angesehen wurde. So konnte bestimmt werden, wie viele der 40 Datenpunkte einer Simulation als Ausreißer oder andere Simulation zu interpretieren waren. Je weniger Ausreißer es gab, desto wahrscheinlicher war es, dass die gleiche Simulation betrachtet wurde.

Bei der Durchführung dieses Verfahrens durch Malone et al. ergab sich bei Vergleichen von unterschiedlichen Simulationen ein Anteil von 73% bis 100% Ausreißer (Abb. 6). Bei Vergleichen von gleichen Simulationen mit unterschiedlichen Saaten ergab sich ein Anteil von 0% bis 3% Ausreißer (Abb. 7). Somit wurden Unterschiede gefunden, wo sie vorhanden waren und nicht gefunden, wo sie nicht vorhanden waren. Die offene Frage bei dieser Art der Validierung bleibt allerdings, ab wann man Unterschiede wahrnehmen möchte und welche Unterschiede man erlaubt.

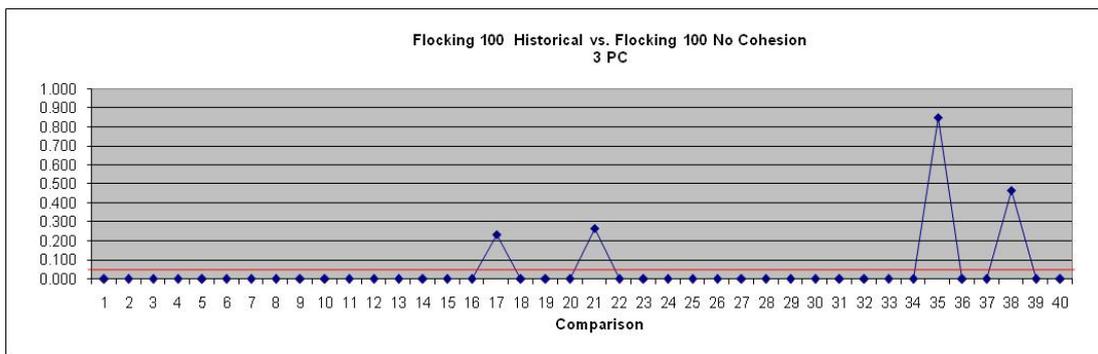


Abbildung 6: Vergleich von 40 Simulationsläufen mit der historischen Datenmenge einer anderen Simulation unter Nutzung von 3 Hauptkomponenten. 90 Prozent der Werte liegen unterhalb des Grenzwertes (Malone et al. 2008)

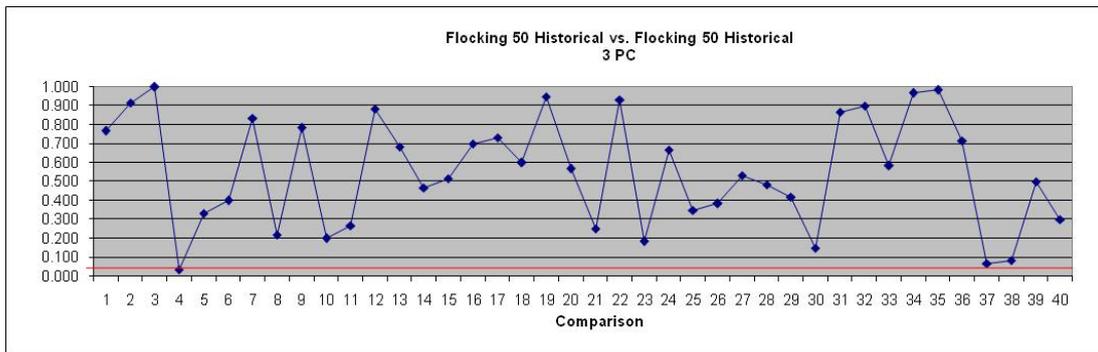


Abbildung 7: Vergleich von 40 Simulationsläufen mit der historischen Datenmenge der gleichen Simulation unter Nutzung von 3 Hauptkomponenten. 3 Prozent der Werte liegen unterhalb des Grenzwertes (Malone et al. 2008)

Mathematische Validierung

Ein mathematisches Verfahren zur Validierung von Gruppensimulationen wurde von Banerjee und Kraemer entwickelt (Banerjee und Kraemer 2010). Dieses Verfahren validiert allerdings nicht auf der Mikroebene, beziehungsweise der Agentenebene, sondern führt eine Validierung auf der Makroebene durch. Unterschiede zwischen Realität und Simulation auf der Mikroebene werden ignoriert, solange die Ergebnisse auf der Makroebene noch zueinander passen. Das Verfahren basiert auf der Überlegung, dass die Umgebung aus p konvexen Polygonen besteht, die die von den Agenten begehbaren Oberflächen darstellen. Es gibt einen Punkt, an dem die Agenten die virtuelle Umgebung verlassen, aber keinen Punkt, an dem Agenten die Umgebung betreten, das heißt die Anzahl der Agenten wird über den Zeitraum der Simulation abnehmen. Die Agenten befinden sich zu Anfang an einer vordefinierten Stelle und verändern ihre Position nur durch das simulierte Verhalten, die Umgebung zu verlassen. Zu gleichen Zeitpunkten des Szenarios in der Simulation und der Aufnahme aus der Realität wird jeweils die Distanz zwischen den Agentenverteilungen über die p Polygone berechnet. Bei komplett gleichem Verhalten auf makroskopischer Ebene wäre die Distanz bei jedem Vergleich immer 0. Die Vergleiche zwischen Simulation und Aufnahme erfolgen nicht bei gleichen Simulations- und Aufnahmezeiten, sondern bei gleichen Zeitpunkten in dem Szenario, das sie darstellen. Dies liegt daran, dass die Geschwindigkeiten in der Simulation und Aufnahme laufen, unterschiedlich sein können. Wenn also zum Beispiel eine Minute im Szenario der Aufnahme vergangen ist, sollen die Verteilungen im Szenario der Aufnahme mit den Verteilungen im Szenario der Simulation, in dem ebenfalls eine Minute vergangen ist, verglichen werden. Es sollen also die Zeiten der Simulation mit den Zeiten der Aufnahme verglichen werden, welche die gleichen Zeitpunkte des beobachteten Ablaufes darstellen sollen. Das heißt, dass für die Simulation und die Aufnahme der Abstand zwischen t und $t + 1$ unterschiedlich sein kann.

$$D(sim_t, scn_{t'}) = \sum_{i=1}^{i=p} \left(\frac{C_{sim}^t(R_i) - C_{scn}^{t'}(R_i)}{X(R_i)} \right)^2$$

D ist die Distanz der Verteilungen, sim_t ist die Zeit in der Simulation, $scn_{t'}$ ist die Zeit in der Aufnahme beziehungsweise des realen Szenarios und R_i ist das i -te Polygon. $C_x^t(r)$ gibt die Anzahl an Agenten auf dem Polynom beziehungsweise in der Region r zur Zeit t . $X(R_i)$ ist die maximale Kapazität der Region R_i .

Die Werte für $C_{scn}^t(R_i)$ sind für diskrete Zeitpunkte t_0, t_1, \dots, t_k gegeben, welche einen konstanten Zeitabstand von τ_{scn} besitzen und durch Schnappschüsse einer Videoaufnahme generiert werden können. Zum Start der Simulation wird die Verteilung in dieser so wie in dem realen Szenario gesetzt. Nun wird die Simulation für τ_{sim} Einheiten der Simulationszeit laufen gelassen und die Distanz der Verteilungen zwischen der Simulation und dem realen Szenario berechnet. Die berechnete Distanz wird mit einem Grenzwert δ , der die maximal zulässige Abweichung definiert und zwischen 0 und 1 liegt, verglichen. Liegt die Distanz unter dem Grenzwert, so läuft die Simulation einfach weiter, ansonsten wird die Verteilung des realen Szenarios zu diesem Zeitpunkt in der Simulation repliziert, dies wird als Snappoint bezeichnet, und die Simulation läuft danach weitere τ_{sim} Einheiten der Simulationszeit. Die Vergleiche werden nun wiederholt bis t_k erreicht wurde. Jetzt wird gezählt wie oft die Verteilung des realen Szenarios in der Simulation repliziert werden musste. Die Idee besteht darin, dass eine Simulation, bei der dies weniger oft durchgeführt werden musste, genauer für das betrachtete Szenario und den gewählten Grenzwert ist. Um zwei Simulationen für verschiedene Grenzwerte bei einem Szenario zu vergleichen, kann die Fläche unter der Kurve, die die Anzahl des Replizierens zum Grenzwert darstellt, betrachtet werden (Abb. 8, Abb. 9). Je geringer diese Fläche ist, desto genauer ist die Simulation.

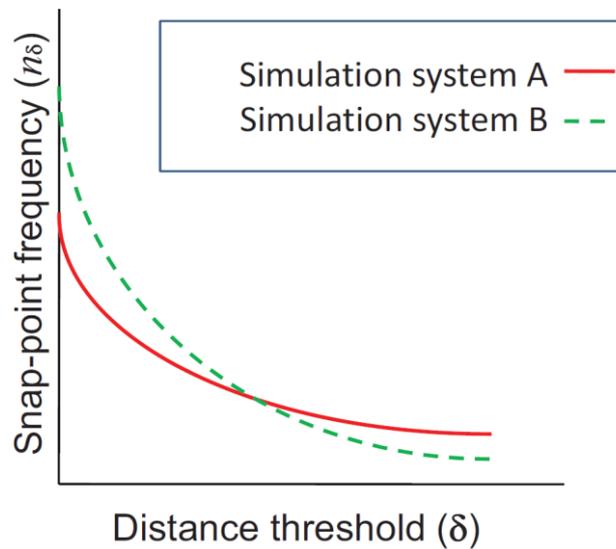


Abbildung 8: Darstellung von zwei Kurven von unterschiedlichen Simulationen, die die Anzahl des Replizierens (snap-point frequency) zum Grenzwert (distance threshold) angeben (Banerjee und Kraemer 2010)

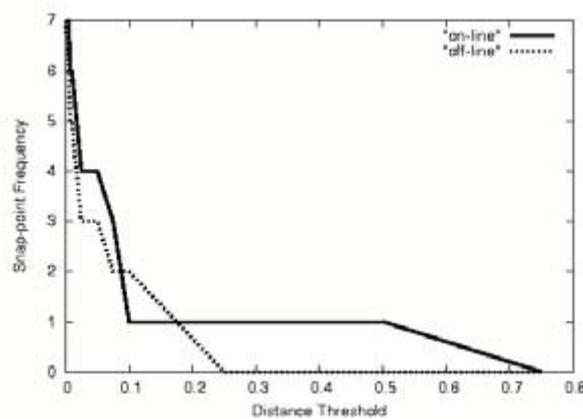


Abbildung 9: Tatsächlich gemessene Anzahlen des Replizierens (snap-point frequency) zum Grenzwert (distance threshold) von zwei unterschiedlichen Simulationen (Banerjee und Kraemer 2010)

Das Hauptproblem bei diesem Verfahren ist die Bestimmung von τ_{sim} , da diese dafür sorgen muss, dass ein Zeitpunkt t in der Simulation dem gleichen Zeitpunkt t in der Aufnahme beziehungsweise des realen Szenarios entspricht und somit für eine Synchronisierung zuständig ist. Es gibt zwei Verfahren zur Bestimmung von τ_{sim} : Im ersten Verfahren wird τ_{sim} vor dem Start der Validierung bestimmt, im zweiten Verfahren findet dies während der Validierung statt.

Im ersten Verfahren wird die Anzahl der Personen beziehungsweise Agenten, die die Umgebung zu verschiedenen Zeitpunkten verlassen haben, betrachtet und verglichen und daraufhin τ_{sim} so gewählt, dass die Verteilungen in der Simulation und der Aufnahme möglichst genau zusammenpassen. Zuerst wird eine Sequenz von der Anzahl der Personen, die die Umgebung zu einem bestimmten Zeitpunkt verlassen haben, erstellt. Dabei ist $C_x^t = \sum_i C_x^t(R_i)$, gibt also die Gesamtanzahl an Personen beziehungsweise Agenten in der gesamten Umgebung an.

$L_{scn}^j = C_{scn}^{t_0} - C_{scn}^{t_j}$ gibt die Anzahl der Personen im realen Szenario, die die Umgebung zum Zeitpunkt j verlassen haben. $\{L_{scn}^j\}_{j=1}^{j=k}$ ist nun die Sequenz, die die Personen im realen Szenario angibt, die zu den Zeitpunkten 1 bis k vom Start des realen Szenarios die Umgebung verlassen haben. Die Simulation wird mit der gleichen Startverteilung wie im realen Szenario laufen gelassen und die Gesamtanzahl der Agenten C_{sim}^t wird in einem kleinen Intervall Δ aufgezeichnet, wodurch eine Sequenz $C_{sim}^0, C_{sim}^\Delta, C_{sim}^{2\Delta}, \dots, C_{sim}^{m\Delta}$ entsteht, für die gilt $C_{sim}^{m\Delta} \leq C_{scn}^{t_k}$ und $C_{sim}^{(m-1)\Delta} \geq C_{scn}^{t_k}$, das heißt, dass der Zeitpunkt $m\Delta$ der erste gemessene Zeitpunkt in der Simulation ist, bei dem die Gesamtanzahl der Agenten kleiner oder gleich der Gesamtanzahl der Personen zum Schluss des realen Szenarios ist. Aus den gemessenen Werten lässt sich die Sequenz $\{L_{sim}^j\}_{j=1}^{j=m}$ bilden, in der L_{sim}^j die Anzahl der Agenten angibt, die die Umgebung der Simulation zur Zeit $j\Delta$ seit dem Start der Simulation verlassen haben. Ist das gewählte Δ klein genug, so ist sichergestellt, dass $m \gg k$ gilt. τ_{sim} kann nun als $\tau_{sim} \approx h^* \Delta$ geschätzt werden, wobei h^* eine Ganzzahl darstellt, welche durch die Lösung des diskreten Optimierungsproblems $h^* = \arg \min_h \sum_{s=1}^k (L_{sim}^{sh} - L_{scn}^s)^2$ unter der Einschränkung, dass $kh \leq m$ gilt, berechnet wird. Es wird also das h gesucht, für das die Abweichungen zwischen den Anzahlen an Agenten und Personen, die die Umgebung verlassen haben, am geringsten ist.

Die andere Möglichkeit τ_{sim} zu bestimmen, ist dies während der Validierung durchzuführen. Hierzu werden die aus dem Segment von t_{j-1} bis t_j gewonnenen Informationen dafür eingesetzt, um τ_{sim} für das Segment von t_j bis t_{j+1} anzupassen. τ_{sim}^j ist das τ_{sim} , das zum Zeitpunkt t_j berechnet wurde und für das darauf folgende Segment benutzt wurde.

$$\tau_{sim}^j = \tau_{sim}^{j-1} * \left(\frac{C_{scn}^{t_{j-1}} - C_{scn}^{t_j}}{C_{sim}^x - C_{sim}^y} \right)$$

Für x gilt $x = \sum_{t=0}^{j-2} \tau_{sim}^t$, für y gilt $y = x + \tau_{sim}^{j-1}$, für j gilt $j = 1, \dots, k - 1$ mit $\tau_{sim}^{-1} = 0$ und $\tau_{sim}^0 = t_1 - t_0$.

$\frac{C_{scn}^{t_{j-1}} - C_{scn}^{t_j}}{C_{sim}^x - C_{sim}^y}$ gibt also das Verhältnis der Anzahl der Personen, die die Umgebung des realen Szenarios verlassen haben zur Anzahl der Agenten, die die Umgebung der Simulation verlassen haben an. Somit lässt sich bestimmen, ob die Simulation zu schnell, das heißt der Wert der Gleichung ist kleiner als 1, oder zu langsam, das heißt der Wert der Gleichung ist größer als 1, läuft. Läuft die Simulation also zu schnell, so wird τ_{sim} für das nächste Segment verringert, läuft die Simulation zu langsam, so wird τ_{sim} erhöht. Dieses Verfahren ist weniger aufwendig als das andere Verfahren, allerdings sorgt es dafür, dass sich Über- und Unterkompensation der Schätzung abwechseln. Ist also k nicht groß genug, so kann es sein, dass die Schätzung keine Zeit dazu hat, sich zu stabilisieren. Aus diesem Grund ist das andere Verfahren als genauer anzusehen.

Validierung durch eine Richtlinie

Die letzte hier vorgestellte Möglichkeit eine Gruppensimulation zu validieren, beruht auf der Zuhilfenahme von Richtlinien. Als Beispiel wird hier die „Richtlinie für Mikroskopische Entfluchtungsanalysen“ (RiMEA) vorgestellt (Meyer-König et al. 2009). Die RiMEA ist eine Muster-Richtlinie, die eine Handreichung für die Genehmigungsbehörden darstellt. Sie soll standardisierte Kriterien für die Szenarien und Parameter bei der Fluchtwegberechnung definieren. Wie der Name bereits aussagt, werden in der RiMEA mikroskopische Gruppensimulationen betrachtet. Somit werden die Bewegungen jeder einzelnen Person dargestellt und jede Person verfügt über individuelle Fähigkeiten, die ihr Verhalten charakterisieren. In der Richtlinie wird jedoch deutlich darauf hingewiesen, dass bei der Fluchtwegberechnung nicht alle Einflüsse der Realität berücksichtigt werden können. Psychologische Aspekte, die unter anderem die Routenwahl und das Verhalten der Personen beeinflussen, werden nicht berücksichtigt, da diese noch nicht wissenschaftlich fundiert untersucht wurden. Die in der RiMEA behandelten Simulationen stellen somit idealisierte Fälle dar, bei denen sich die Personen gemäß der vorgegebenen Parameter und Routenvorgaben bewegen.

Die Richtlinie definiert vier Ziele, die die Entfluchtungsanalyse erfüllen soll:

1. Die Gesamtentfluchtungsdauer beziehungsweise die Entfluchtungsdauer von Teilbereichen von baulichen Anlagen soll statistisch erfasst werden und unter Berücksichtigung von sicherheitstechnischen Aspekten bewertet werden.
2. Wenn erforderlich soll sie den Nachweis führen, dass die geplanten oder bestehenden Flucht- und Rettungswege abweichend von den Dimensionierungsvorgaben des Bauordnungsrechts für die angenommenen Personenzahlen ausreichen.
3. Es soll gezeigt werden, dass die Fluchtvorkehrungen ausreichend flexibel sind für den Fall, dass bestimmte Flucht- und Rettungswege oder gesicherte Bereiche aufgrund eines Zwischenfalls nicht verfügbar sind.

4. Soweit möglich sollen signifikante Stauungen, die während der Entfluchtung aufgrund der normalen Bewegung von Personen entlang der Flucht- und Rettungswege auftreten, erkannt werden.

Ein wichtiger Bestandteil der RiMEA ist die Definition der in der Simulation verwendeten Population. Diese hat aus einzelnen Personen zu bestehen, welche individuell repräsentiert und aufgezeichnet werden. Die Leistungsfähigkeit von Personen oder Personengruppen wird durch Parameter festgelegt. Diese Parameter variieren zwischen den Individuen einer Population. Außerdem sind nach der Richtlinie bestimmte Vereinfachungen zulässig. Dazu gehört, dass sich Personen immer entlang der Flucht- und Rettungswege bewegen. Außerdem muss Gruppenverhalten nicht explizit berücksichtigt werden, sondern muss nur implizit über die Zuordnung von Fluchtwegen zu Personengruppen berücksichtigt werden.

Des Weiteren gibt die Richtlinie Standardvorgaben für die Altersverteilung der Population, die Personendichte abhängig vom Gebäudetyp (zum Beispiel Kaufhaus oder Bürogebäude), die ungehinderten Gehgeschwindigkeiten auf der Ebene und auf Treppen abhängig vom Alter und Geschlecht der Person und die Reaktionsdauer der Personen in Abhängigkeit von Eigenschaften wie der Wachsamkeit und Vertrautheit der Personen, der Art des Alarmierungssystems, der Gebäudekomplexität und anderen. Diese Vorgaben basieren jeweils auf wissenschaftlichen Untersuchungen und sollen angewendet werden, wenn keine anderen Werte für das Szenario vorliegen.

Die Richtlinie gibt also vor, wie die Simulation abzulaufen hat und welche Anforderungen sie erfüllen muss. Ein weiterer Teil der Richtlinie gibt eine Anleitung zur Validierung und Verifizierung mit fest vorgegebenen Tests. Dabei wird entgegen der hier vorgestellten Definitionen von Verifizierung und Validierung kein Unterschied zwischen den beiden Begriffen gemacht, sondern beides unter dem Begriff Verifizierung zusammengefasst. Die unterschiedlichen Bereiche der Tests werden somit als Überprüfung der Komponenten, funktionale Verifizierung, qualitative Verifizierung und quantitative Verifizierung angegeben.

Die Tests zur Überprüfung der Komponenten dienen dazu, zu klären, ob die verschiedenen Komponenten der Software wie vorgesehen funktionieren. Es handelt sich also um Verifizierung zur Überprüfung, ob die Software korrekt implementiert wurde. Die funktionale Verifizierung dient dazu zu überprüfen, ob das Modell alle Fähigkeiten dafür besitzt, die in der Richtlinie definierte Form der Entfluchtungsanalyse umzusetzen. Als qualitative Verifizierung wird in der Richtlinie die Überprüfung, ob das vorhergesagte menschliche Verhalten mit den sachkundigen Erwartungen übereinstimmt, bezeichnet. Nach den in dieser Arbeit benutzten Definitionen von Verifizierung und Validierung handelt es sich dabei allerdings um Validierung, da nicht die Korrektheit der Implementierung betrachtet wird, sondern ob das Modell das Verhalten richtig beschreibt und dadurch realistisches Verhalten erzeugt werden kann. Als quantitative Verifizierung wird in der Richtlinie der Vergleich der Vorher-

sagen des Modells mit zuverlässigen Daten aus Evakuierungsübungen bezeichnet. Dies wäre nach den in dieser Arbeit benutzten Definitionen von Verifizierung und Validierung ebenfalls Bestandteil der Validierung und nicht der Verifizierung. Im Gegensatz zu den anderen Bereichen gibt es hierfür allerdings keine vorgegebenen Tests oder Daten. Grund hierfür ist, dass nicht genügend zuverlässige experimentelle Daten für diese Vergleiche vorhanden sind. Aus diesem Grund definiert die Richtlinie die anderen Bereiche als ausreichend für den Verifizierungsprozess. Da die Tests der Richtlinie in der Gesamtheit ausgeführt werden sollten, werden im Folgenden auch die Tests zur Verifizierung vorgestellt. Insgesamt handelt es sich um 14 Tests. Nach der Durchführung der Tests besteht die Möglichkeit das Testdokument auf der Webseite der RiMEA e.V. zu veröffentlichen. Hierdurch wird Außenstehenden, zum Beispiel Behörden, die Möglichkeit gegeben verschiedene Evakuierungssimulationen auf objektive Weise zu vergleichen (Meyer-König 2004).

Die ersten sieben Tests dienen der Überprüfung der korrekten Arbeitsweise der implementierten Komponenten. Im 1. Test wird überprüft, ob eine Person mit einer vordefinierten Gehgeschwindigkeit zwischen 4,5 und 5,1 km/h von dem einen Ende eines Ganges mit 2 m Breite und 40 m Länge zu dem anderen Ende des Ganges in der entsprechenden Zeit gelangt. Die Zeit, die dafür benötigt wird, sollte zwischen 26 und 34 Sekunden liegen. Dieser Test überprüft somit, ob das Festlegen der Geschwindigkeit der Person auch den entsprechenden Einfluss auf ihre Geschwindigkeit in der Simulation hat und ihre Geschwindigkeit während der Ausführung beibehalten wird.

Eine ähnliche Überprüfung führen der 2. und 3. Test aus. In diesen Tests wird jeweils überprüft, ob die vorgegebene Gehgeschwindigkeit einer Person treppauf beziehungsweise treppab beibehalten wird. Die Breite der Treppe ist mit 2 m vorgegeben und die Länge der Treppe entlang der Schräge mit 10 m und es ist zu zeigen, dass das Ende der Treppe in der der Geschwindigkeit entsprechenden Zeit erreicht wird.

Im 4. Test wird der spezifische Fluss durch einen Querschnitt überprüft. Dieser stellt die Anzahl an Personen da, die einen Querschnitt pro Meter lichter Breite und pro Sekunde passieren und wird als Personen/ms angegeben. Der spezifische Fluss ist von der Personendichte, also Personen/m^2 , abhängig. Die Richtlinie gibt eine Formel zum Berechnen des spezifischen Flusses in Abhängigkeit von der Personendichte und ein Fundamentaldiagramm (Abb. 10) zum Vergleich mit den Ergebnissen der Simulation. Außerdem werden die maximalen spezifischen Flüsse für Treppen auf- und abwärts sowie Gänge und Türöffnungen vorgegeben. Der Test fordert nun einen Gang mit 4 m Breite, einer Mindestlänge von 30 m und periodischen Randbedingungen. Dies bedeutet, dass Personen, die den Gang am Ende verlassen, ihn sofort wieder am Anfang betreten. Fluss und Dichte sollen über den gesamten Gang gemittelt werden. Der Test überprüft somit unter anderem, ob sich die Personen bei der Fortbewegung behindern. Da es sich um Individuen handelt und dadurch auch die Gehgeschwindigkeit unterschiedlich ist, sollte es ab gewissen Dichten zu einer

Verringerung bis zum Erliegen des spezifischen Flusses kommen. Bei einer fehlerhaften Implementierung der Kollision von Personen, so dass sie sich zum Beispiel durchqueren können, würde sich der spezifische Fluss nicht entsprechend des Fundamentaldiagramms verhalten.

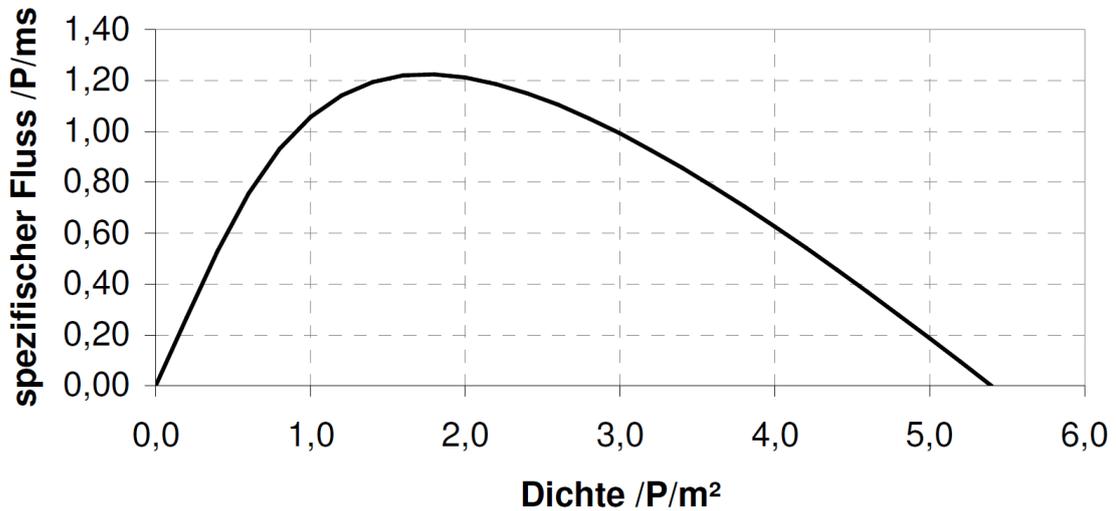


Abbildung 10: Das Fundamentaldiagramm, das den spezifischen Fluss in P/ms zur Dichte in P/m² angibt (Meyer-König et al. 2009)

Der 5. Test befasst sich mit der Reaktionsdauer der Personen. In einem vorgegebenen Raum von 8 m Länge und 5 m Breite mit einem 1 m breiten Ausgang in der Mitte der 5 m langen Wand, befinden sich 10 Personen mit unterschiedlichen Reaktionsdauern, welche gleichverteilt zwischen 10 s und 100 s liegen. Nun soll überprüft werden, ob die Personen auch wirklich nach der ihnen zugewiesenen Reaktionsdauer starten, um den Raum zu verlassen und somit die Reaktionsdauern korrekt implementiert wurden.

Im 6. Test (Abb. 11) wird überprüft, ob die Personen in der Simulation eine Ecke umrunden ohne diese zu durchqueren. Dazu gibt es zwei Gänge mit jeweils 10 m Länge und 2 m Breite. Der zweite Gang setzt rechtwinklig am Ende des ersten Ganges an, so dass aus den beiden Gängen ein Gang mit einer nach links abbiegenden Ecke entsteht. 20 Personen starten in einem 6 m langen Bereich im hintersten Teil des ersten Ganges gleichmäßig verteilt. Bei korrekter Implementierung darf es nicht vorkommen, dass Personen die Ecke durchqueren.

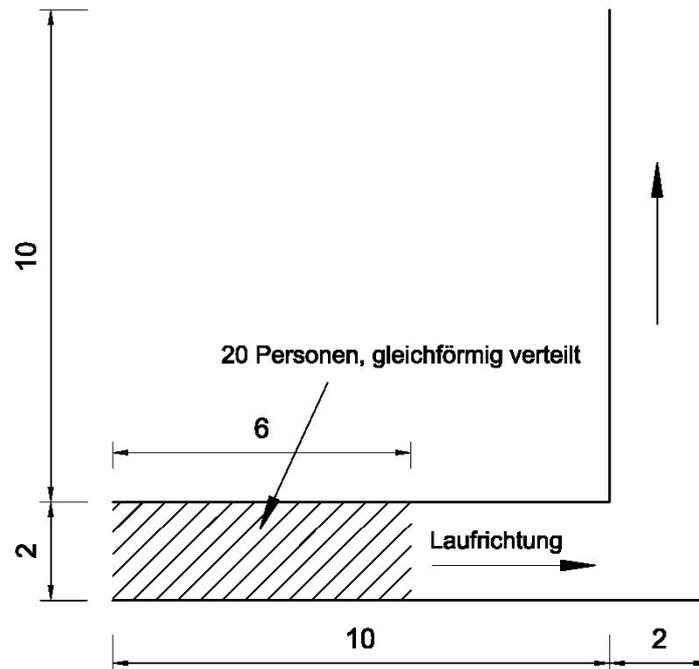


Abbildung 11: Der architektonische Grundriss und die Startverteilung der Personen von Test 6 (Meyer-König et al. 2009)

Im 7. und letzten Test, der der Überprüfung der Komponenten dient, wird die Korrektheit der Zuordnung der demografischen Parameter überprüft. Es sollen die Gehgeschwindigkeiten, die in dem ersten Teil der Richtlinien in einer Tabelle vorgegeben wurden, auf eine aus Erwachsenen bestehende Gruppe verteilt werden. Danach ist zu zeigen, dass die Verteilung mit der Verteilung der Gehgeschwindigkeiten aus der Tabelle vereinbar ist.

Die Tests 1 bis 7 überprüfen die grundlegenden Komponenten der Gruppensimulation auf die korrekte Implementierung. Diese muss gewährleistet sein, damit die anderen Tests durchgeführt werden können. Der folgende Bereich, die funktionale Verifizierung, umfasst in den Richtlinien nur einen Test. Ziel dieses Bereiches ist es zu überprüfen, ob das Modell die benötigten Fähigkeiten besitzt, um alle notwendigen Möglichkeiten der Simulation abzudecken. Diese Überprüfung findet dadurch statt, dass die Entwickler den vollen Bereich

der Möglichkeiten ihres Modells und den daraus folgenden Annahmen darstellen und eine Anleitung zum Gebrauch dieser zur Verfügung stellen. Dies kann, unter der Bedingung, dass das RiMEA-Testdokument zugänglich ist, von Interessenten an der Simulation dafür benutzt werden, herauszufinden, ob die Simulation alles abdeckt, was sie von ihr fordern.

Der 8. Test (Abb. 12) dient der Parameteranalyse, also der Analyse der Auswirkungen der Parameter auf den Ablauf der Simulation. Hierzu ist ein Gebäude bestehend aus Erdgeschoss, erstem und zweiten Stock vorgegeben, die jeweils über Treppen verbunden sind. In den Stockwerken befindet sich eine Vielzahl an Räumen mit jeweils 4 Personen. Um die Auswirkungen der Parameter zu zeigen, wird jeweils nur ein Parameter verändert, die restlichen Parameter werden auf Standardwerte eingestellt. Die Parameter sollen einmal für alle Personen gleich sein und einmal statistisch gleich verteilt um einen festen Mittelwert variiert werden. Die Ergebnisse sollen in Graphen festgehalten werden.

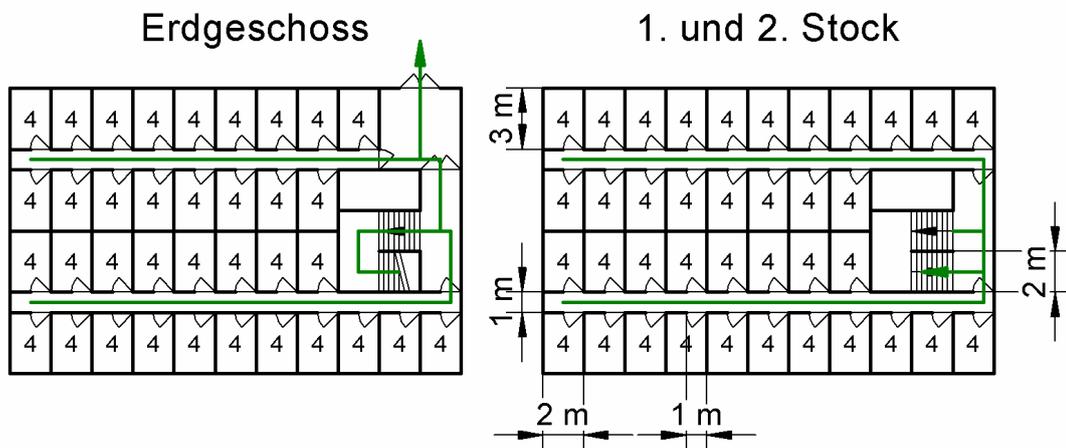


Abbildung 12: Der architektonische Grundriss der verschiedenen Stockwerke und der Rettungsweg von Test 8 (Meyer-König et al. 2009)

Der letzte Bereich, für den Tests vorhanden sind, ist die qualitative Verifizierung. In diesen wird überprüft, ob das simulierte Verhalten den sachkundigen Erwartungen und somit realistischem Verhalten entspricht. Es handelt sich also eigentlich um Tests zur Validierung. Im 9. Test (Abb. 13) verlässt eine große Menschenmenge einen öffentlichen Raum von 30 m x 20 m mit 4 Türen. Jeweils 2 Türen befinden sich an den längeren Seiten mit jeweils einem Abstand von 4 m von der jeweils dichtesten Ecke. In einem Bereich mit 2 m Abstand von allen Wänden werden 1000 Personen gleichmäßig verteilt und die Simulation gestartet. Die Zeit bis die letzte Person den Raum verlassen hat wird aufgezeichnet. Das gleiche wird wiederholt, jedoch sind die Türen auf einer Seite beide nicht passierbar. Die Erwartung ist, dass beim zweiten Durchlauf eine ungefähre Verdopplung der Zeit zum Verlassen des Raumes stattfindet. Dieser Test wird allerdings nicht als Ausschlusskriterium angesehen, weil es noch keine empirische Antwort auf die Frage gibt, inwieweit ein größere Menge an Personen, die sich durch die zwei Ausgängen begeben wollen, zu höheren Dichten und erhöhten Fluss durch die Türen führen kann. Er soll somit nur das Modellverhalten dokumentieren.

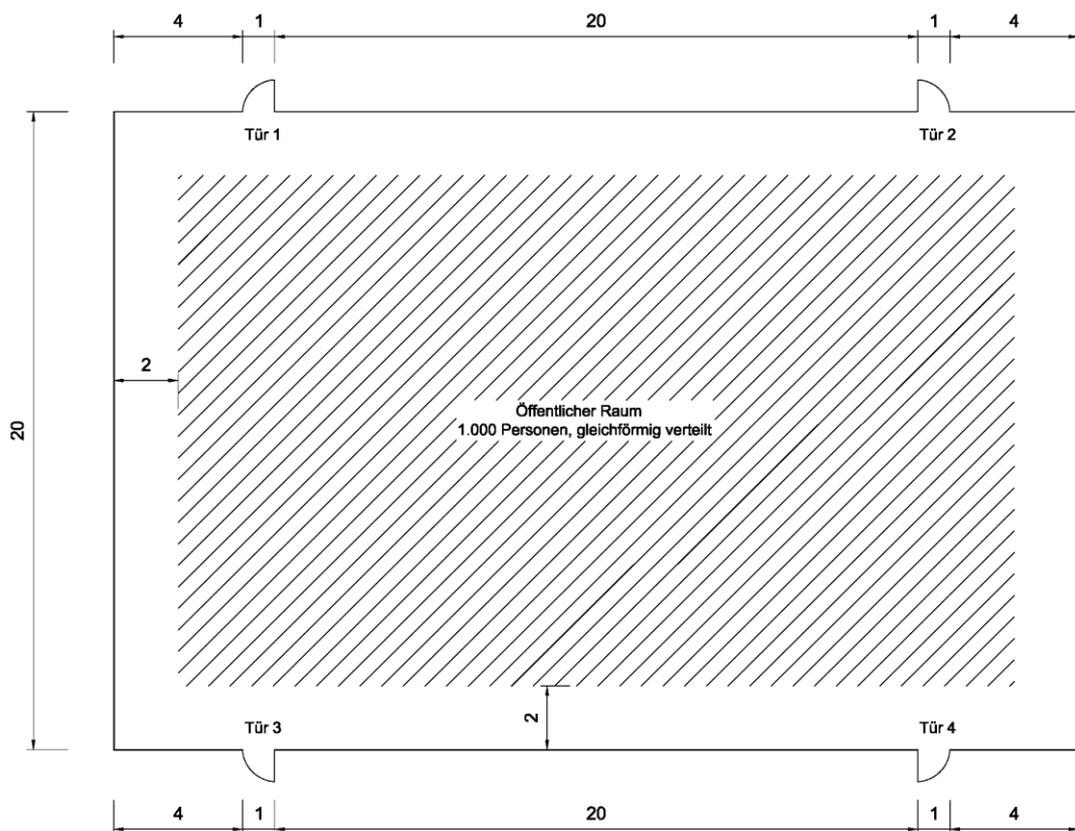


Abbildung 13: Der architektonische Grundriss und die Startverteilung der Personen von Test 9 (Meyer-König et al. 2009)

Im 10. Test (Abb. 14) werden 23 Personen Rettungswege zugewiesen. Die Gehgeschwindigkeiten von Erwachsenen sind auf die Personen verteilt und sie starten mit sofortiger Reaktion aus verschiedenen Räumen. Bei Ausführung der Simulation wird nun erwartet, dass alle Personen den entsprechenden Ausgang nehmen. Dieser Test würde eher in den Bereich der Überprüfung der Komponenten passen, da nur überprüft wird, ob das Zuweisen des Rettungsweges auch zur Nutzung dieses führt, also ob die Implementierung korrekt ist.

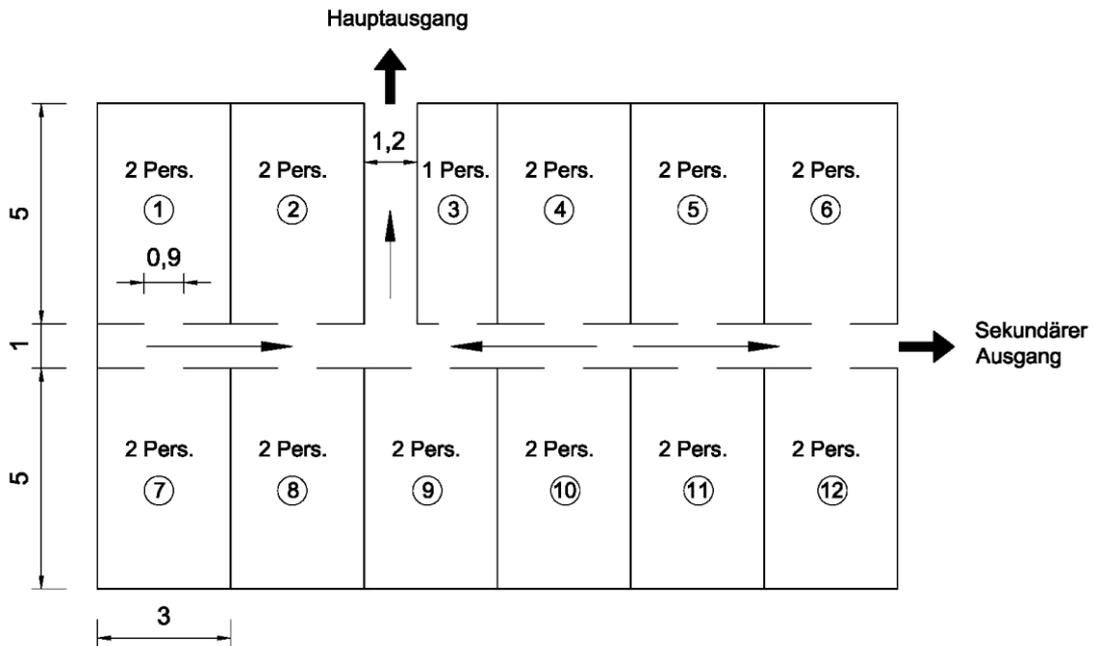


Abbildung 14: Der architektonische Grundriss, die Startverteilung der Personen und die zugewiesenen Rettungswege in Test 10 (Meyer-König et al. 2009)

Im 11. Test (Abb. 15) dagegen geht es um die Wahl des Rettungsweges. Diese wird überprüft indem in einem 30 m x 20 m großen Raum 1000 Personen von links nach rechts mit maximaler Dichte platziert werden, wobei sie mindestens 2 m Abstand von den Wänden haben. An der oberen langen Wand befinden sich 2 Ausgänge. Einer in 2 m Abstand von der Ecke und einer in 8 m Abstand von der Ecke. Bei der Durchführung wird nun davon ausgegangen, dass der nähere Ausgang bevorzugt wird und in diesem Bereich Stauungen auftreten, jedoch einzelne Personen aufgrund der Stauungen auch den weiter entfernten Ausgang benutzen.

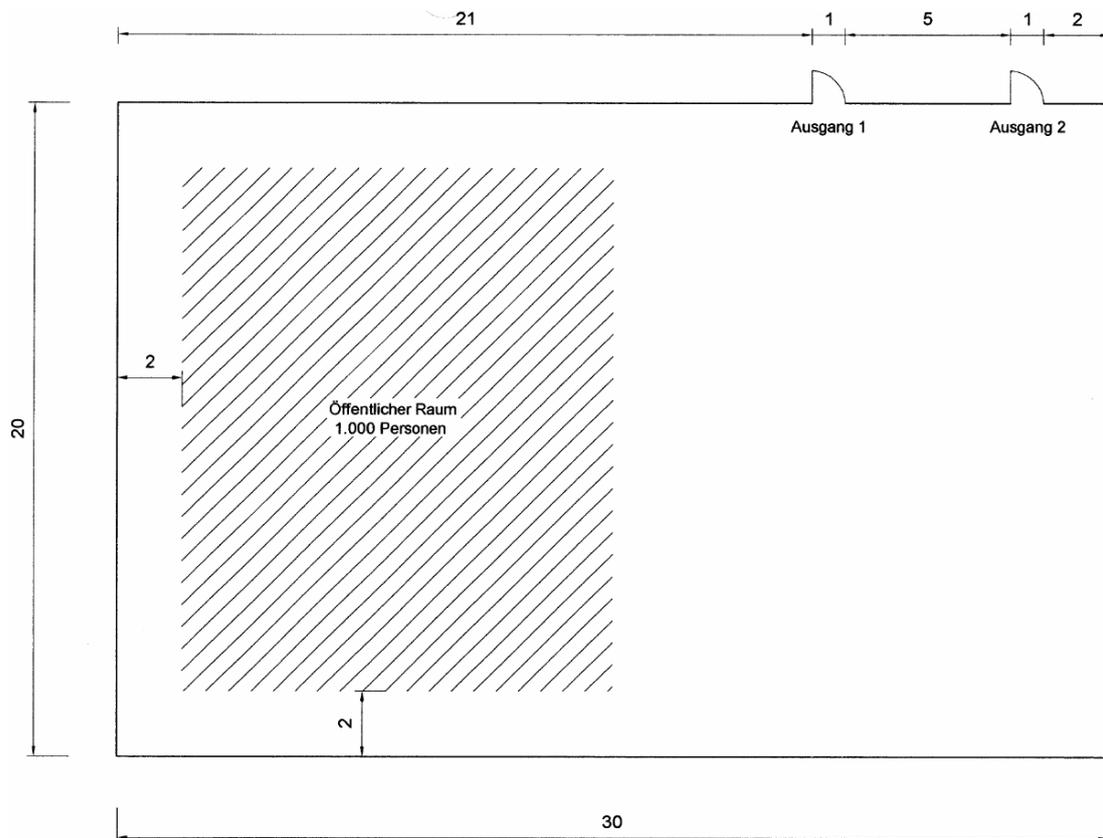


Abbildung 15: Der architektonische Grundriss und die Startverteilung der Personen von Test 11 (Meyer-König et al. 2009)

Test 12 (Abb. 16) überprüft die Auswirkungen von Engstellen. Dazu sind 2 Räume von 10 m x 10 m über einen 5 m langen und 1 m breiten Gang verbunden. Der Ausgang befindet sich im 2. Raum gegenüber der Stelle, an der die Personen den Raum durch den Gang betreten und ist ebenfalls 1 m breit. 150 erwachsene Personen mit unterschiedlichen Gehgeschwindigkeiten wie zuvor definiert haben ihre Startposition im linken Teil des ersten Raumes, der Gang befindet sich in der Mitte des rechten Teils des Raumes. Es wird nun davon ausgegangen, dass es zwar zu Stauungen am Gang kommt, jedoch durch die Begrenzung des Personenstroms durch den Gang keine Stauungen mehr im 2. Raum auftreten. Allerdings ist auch hier bisher keine Entscheidung auf empirischer Basis darüber möglich, ob es nicht doch zu Stauungen im 2. Raum kommen könnte. Aus diesem Grund gilt auch für diesen Test, dass er kein Ausschlusskriterium ist, sondern nur Modellverhalten dokumentieren soll.

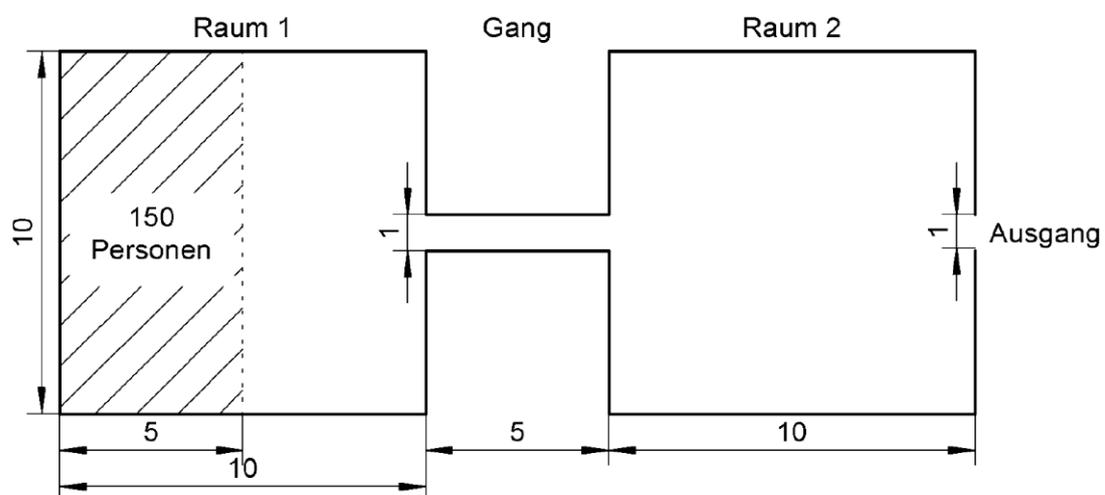


Abbildung 16: Der architektonische Grundriss und die Startverteilung der Personen von Test 12 (Meyer-König et al. 2009)

Test 13 (Abb. 17) befasst sich ebenfalls mit Stauungen. Ein 8 m x 5 m großer Raum ist mit einem 2 m breiten und 12 m langen Gang, der von der Mitte der rechten langen Seite des Raumes abgeht, verbunden. Dieser mündet in einer ebenfalls 2 m breiten und 3 m langen Treppe. Nach dieser Treppe folgt ein weiterer 1 m langer und 2 m breiter Gang, der zum 2 m breiten Ausgang führt. Im Raum befinden sich 150 erwachsene Personen, auf die die unterschiedlichen Gehgeschwindigkeiten verteilt sind. Beim Ablauf der Simulation wird nun erwartet, dass es zuerst zu einem Stau am Ausgang des Raumes kommt und danach bei der Treppe. Der Stau an der Treppe soll über die Zeit zunehmen, da der Fluss über die Treppe geringer als durch den Gang ist. Auch wenn Test 12 und 13 stark in Richtung Verifizierung gehen, können sie noch zur Validierung gezählt werden, da überprüft wird, ob das in der Realität beobachtete Verhalten auftritt.

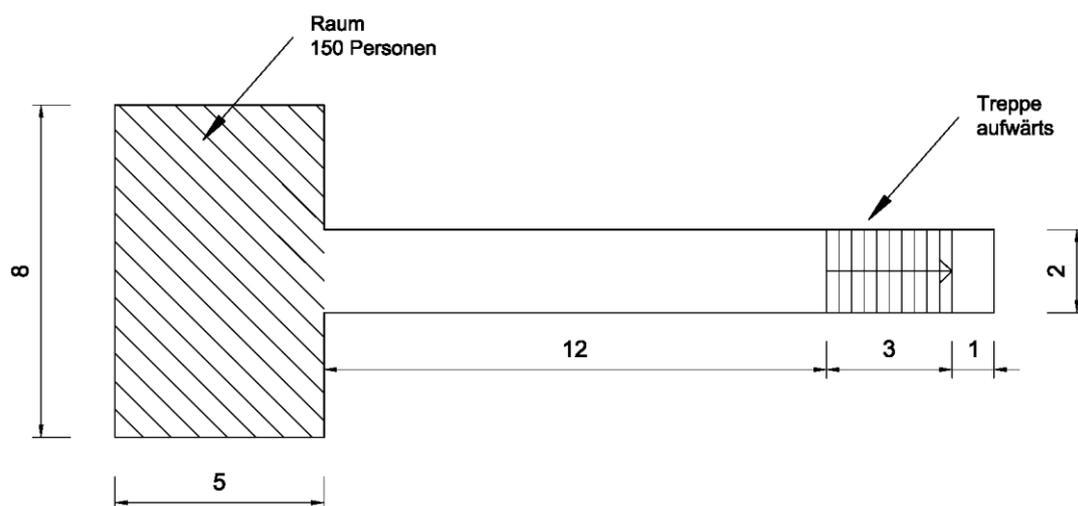


Abbildung 17: Der architektonische Grundriss und die Startverteilung der Personen von Test 13 (Meyer-König et al. 2009)

Der letzte Test, Test 14 (Abb. 18), befasst sich nun schließlich mit der Routenwahl. Die Personen haben die Wahl, ob sie einen kürzeren Weg zum Zielbereich nehmen, der über ein anderes Stockwerk geht, oder ob sie im gleichen Stockwerk bleiben und somit den längeren Weg nehmen. Für diesen Test gibt es keine Erwartungen, er soll nur das Verhalten des Modells dokumentieren.

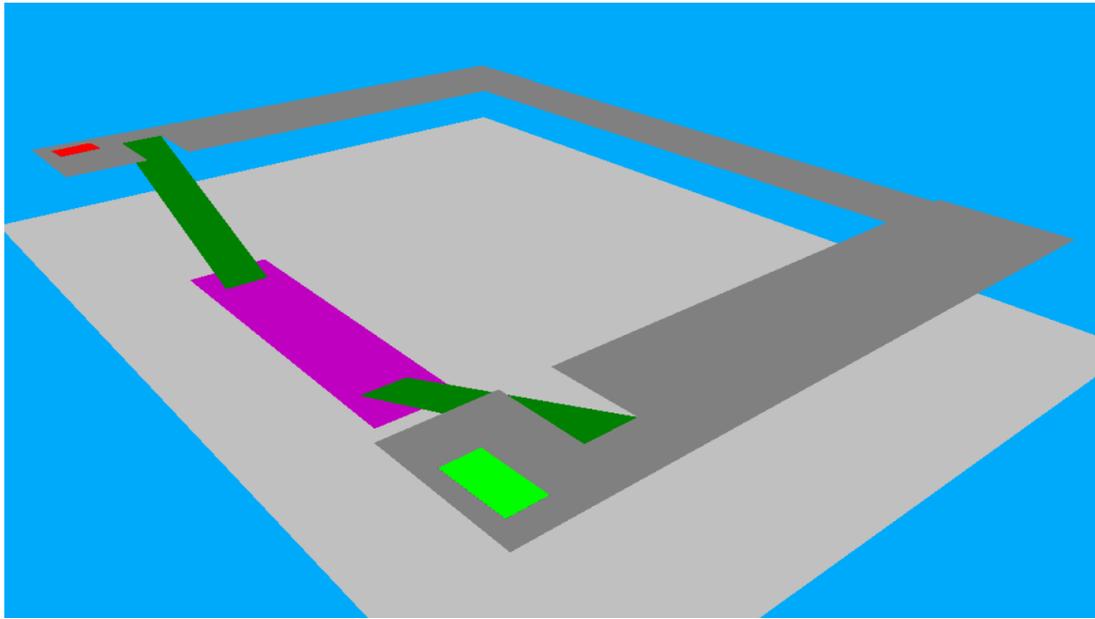


Abbildung 18: Der architektonische Grundriss von Test 14 (Meyer-König et al. 2009)

Der letzte Bereich der in der RiMEA behandelten Validierung und Verifizierung ist die quantitative Verifizierung. Bei dieser sollen die Modellvorhersagen mit zuverlässigen Daten aus Evakuierungsübungen verglichen werden. Hierfür sind keine Daten oder Tests vorgegeben, da noch nicht genügend zuverlässige experimentelle Daten vorhanden sind. Die RiMEA betrachtet daher die vorgestellten Tests als vorerst ausreichend.

Wie in diesem Teil der Arbeit gezeigt wurde, gibt es viele verschiedene Verfahren zur Validierung von Gruppensimulationen. Diese reichen von relativ freien subjektiven Verfahren bis zu sehr konkreten Verfahren mit strikten Vorgehensweisen, Formeln und Tests. Die unterschiedlichen Verfahren besitzen Vor- und Nachteile in verschiedenen Bereichen. So lassen sich psychologische Aspekte und individuelle Entscheidungen insgesamt besser durch Face Validation validieren, auch wenn die Ergebnisse der Validierung subjektiv sind, da es hierfür noch keine konkreten objektiven Verfahren gibt und es an Daten aus Experimenten oder echten Evakuierungen fehlt. Auf der Makroebene hingegen lässt sich die Validierung gut durch objektive Vergleiche von Werten oder durch konkrete Tests durchführen. Insgesamt ist jedoch zu sagen, dass trotz der verschiedenen Ansätze das Fehlen von genügend

zuverlässigen Daten aus Experimenten und echten Evakuierungen weiterhin ein großes Problem der Validierung darstellt, welches gelöst werden muss. Ob sich in Zukunft durch Virtual-Reality-Experimente verlässliche Daten zur Validierung, unter anderem auf der psychologischen Ebene, gewinnen lassen können, muss sich noch zeigen. Aus diesem Grund ist eine Kombination von verschiedenen Verfahren zur Validierung sinnvoll. Hierdurch lassen sich die Vorteile der einzelnen Verfahren nutzen.

3 Erstellung eines Validierungskonzeptes

In diesem Kapitel wird auf Grundlage des im vorherigen Kapitel gezeigten Forschungsstandes ein konkretes Konzept zur Validierung von Gruppensimulationen in den verschiedenen Phasen der Entwicklung erstellt und präsentiert. Hierzu wird eine Auswahl aus den vorgestellten Verfahren auf die Nützlichkeit und den damit verbundenen Aufwand geprüft. Aufgrund dieser Analyse wird anschließend das konkrete Validierungskonzept entworfen. Im letzten Abschnitt wird beschrieben, welche Anforderungen eine Gruppensimulation erfüllen muss, damit dieses Konzept angewandt werden kann.

Ziel des Validierungskonzeptes ist es, zu klären wo die Vor- und Nachteile der vorgestellten Verfahren liegen und wie sie sich kombinieren lassen, um sich gegenseitig zu ergänzen und die Nachteile zu beseitigen. Dazu gehört auch die Betrachtung von verschiedenen Bedingungen, unter denen die Verfahren eingesetzt werden können und die Anforderungen, die die Verfahren an die Simulation sowie an Rahmenbedingungen wie das Vorhandensein von ausführlichen Daten für die Validierung stellen. Außerdem wird geklärt, was mindestens validiert werden sollte und welche Verfahren dafür am besten geeignet sind. Hierfür werden der Ablauf des Entwicklungsprozesses einer Simulation und die für die Phasen geeigneten Verfahren betrachtet. Es wird außerdem unter Berücksichtigung der Art des Projektes geklärt, von wem die Validierung durchgeführt werden sollte. Außerdem wird das entwickelte Konzept auf bisher nicht behebbare Probleme untersucht und ein Ausblick gegeben, wie sich diese später einmal beheben lassen könnten.

3.1 Entwicklungsphasen und die dazugehörige Validierung

Die Entwicklung einer Simulation erfolgt in verschiedenen Phasen, in denen verschiedene Arten der Validität zu gewährleisten sind. Hierfür werden unterschiedliche Verfahren benötigt. Der allgemeine Ablauf der Entwicklung einer Simulation beginnt mit dem Sammeln von Daten zu dem zu modellierenden Problem. Damit aus diesen Daten die Thesen für das konzeptuelle Modell gebildet werden können, muss sichergestellt sein, dass die Daten valide sind, somit ist Datenvalidität gefordert. Bevor das konzeptuelle Modell implementiert wird, muss sichergestellt sein, dass auch dieses korrekt es, in diesem Schritt wird also konzeptuelle Modellvalidierung gefordert. Anhand des implementierten Modells muss die Validität im

Betrieb geprüft werden, so dass letztendlich bewiesen werden kann, dass die Simulation das betrachtete Problem auf eine valide Art und Weise simuliert. Danach ist die Simulation dazu geeignet das Verhalten des Referenzsystems in Hinblick auf den Schwerpunkt des Modells zu simulieren.

Zuerst sollte also die Datenvalidität gewährleistet werden, da nur auf Grundlage von gültigen Daten auch die richtigen Schlüsse und Zusammenhänge gezogen werden können und die Daten die Grundlage des Modells bilden. Insgesamt müssen alle Daten, auch die Daten, die für die Validierung benutzt werden, auf Datenvalidität überprüft werden.

Konzeptuelle Validität und schließlich Validität im Betrieb sind die weiteren sicherzustellen den Arten der Validität. Die hier genannten Arten der Validität beziehen sich auf die Phasen ihrer Anwendung, also beim Datensammeln, Entwickeln des konzeptuellen Modells und der Implementierung des Modells. Es sollte nur mit einer neuen Phase der Entwicklung angefangen werden, wenn die bisherigen Entwicklungsphasen validiert wurden, da sich ansonsten Fehler aus einer vorherigen Phase auf die folgenden Phasen auswirken können. Ist die Validität der Daten nicht sichergestellt, so kann dies zu einem fehlerhaften konzeptuellen Modell und schließlich einer nicht validen Simulation führen. Das gleiche gilt für ein nicht valides konzeptuelles Modell, welches ebenfalls zu einer nicht validen Simulation führt. Die Validierung sollte also während der ganzen Entwicklung durchgeführt werden. Sollten nachträglich Änderungen an dem konzeptuellen Modell oder dem implementierten Modell gemacht werden, so muss dieser Teil und die darauf aufbauenden Teile erneut validiert werden.

Die genauen Schritte bei der Validierung können von der Art der Simulation und des betrachteten Problems abhängen, da diese noch die Validierung von anderen Teilen der Entwicklung erfordern können. Somit handelt es sich bei den hier beschriebenen Schritten um die Teile, die mindestens validiert werden sollten.

Ein weiterer Punkt, der sich bei verschiedenen Projekten unterscheiden kann, ist die Frage wer die Validierung durchführen sollte. Wie gezeigt wurde, ist es besser die Validierung von einer dritten Partei durchführen zu lassen, damit diese auf eine möglichst objektive Art und Weise durchgeführt werden kann. Die Durchführung der Validierung durch eine dritte Partei stellt somit das Optimum dar. Die Durchführung der Validierung durch die Entwickler sollte nur in Ausnahmefällen erfolgen, in denen die Validität der Simulation nicht die höchste Priorität besitzt. Es ist auch möglich, die Validierung durch die Entwickler durchführen zu lassen und diese später durch eine dritte Partei überprüfen zu lassen. Allerdings können zu spät entdeckte Fehler in der Validierung zu großen Änderungen an der Simulation führen, was wiederum zu hohen zusätzlichen Kosten und Aufwand führen kann.

3.2 Auswahl und Bewertung von geeigneten Verfahren

Zur Überprüfung der Datenvalidität gibt es leider nicht viele Möglichkeiten. Beim Sammeln der Daten aus Experimenten und Aufzeichnungen ist sicherzustellen, dass gute Methoden benutzt werden, so dass beim Sammeln keine Fehler auftreten. Das gleiche gilt für das Pflegen der Daten. Des Weiteren sollte nach Ausreißern gesucht werden und diese auf Korrektheit überprüft werden.

Für die konzeptuelle Validierung und die Validierung im Betrieb gibt es mehr Möglichkeiten. Da viele der Verfahren für die Validierung von Simulationen sehr allgemein oder Bestandteil von spezielleren Verfahren für agentenbasierte Simulationen oder Gruppensimulationen sind, wird aus diesem Bereich im Folgenden nur der Degenerierungstest betrachtet. Die Überprüfung der internen Validität und Sensitivitätsanalysen, die auch ein wichtiger Teil der Validierung von Simulationen sind, sind bereits Bestandteil der beiden vorgestellten Verfahren zur Validierung von agentenbasierten Simulationen, die hier im Folgenden ebenfalls betrachtet werden. Des Weiteren werden die verschiedenen Verfahren zur Validierung von Gruppensimulationen betrachtet.

Degenerierungstests dienen dazu, zu überprüfen, ob das Verhalten der Simulation auch nach einem längeren Zeitraum der Ausführung noch korrekt ist. Dabei handelt es sich um einen sehr wichtigen Aspekt, denn es könnten kleine Ungenauigkeiten in der Simulation vorhanden sein, die sich erst nach einer längeren Ausführung zeigen und möglicherweise zu großen Abweichungen in den Ergebnissen führen. Da diese Tests auf dem implementierten Modell ausgeführt werden, ist dies ein Teil der Validierung im Betrieb.

Die vorgestellten Verfahren zur Validierung von agentenbasierten Simulationen von Klügl (Klügl 2008) und Xiang et al. (Xiang et al. 2005) unterscheiden sich unter anderem darin, dass bei Xiang et al. die Validierung bereits bei dem konzeptuellen Modell beginnt, während sie bei Klügl erst mit der Validierung im Betrieb beginnt. Das Verfahren von Xiang et al. besteht aus Face Validation der Thesen des konzeptuellen Modells, gefolgt von Face Validation der visuellen Darstellung der Abläufe der Simulation, interner Validierung und Validierung im Betrieb durch Vergleich von unterschiedlichen implementierten Simulationen. Klügls Verfahren dagegen beginnt mit verschiedenen Arten der Face Validation, und zwar Animationsbewertung und immersiver Bewertung, gefolgt von der aufwendigeren Ausgabebewertung. Darauf folgen Sensitivitätsanalyse, Kalibrierung und Validierung im Betrieb durch statistische Validierung.

Sowohl bei der Face Validation des konzeptuellen Modells, als auch bei der internen Validierung, Sensitivitätsanalyse und Kalibrierung handelt es sich um wichtige Überprüfungen, die Bestandteil eines Validierungskonzepts sein sollten, jedoch nicht alle in beiden Verfahren vorhanden sind. Die Bestandteile der Verfahren sollten also im Validierungskonzept kombiniert werden.

Da das hier entwickelte Konzept für die Validierung von Anfang bis Ende der Entwicklung einer Gruppensimulation anwendbar sein soll, muss die Validierung des konzeptuellen Modells ein Bestandteil sein.

Bei der Überprüfung der internen Validität handelt es sich um einen wichtigen Bestandteil, da Gruppensimulationen meistens mit Zufallszahlen arbeiten. Es muss bewiesen werden, dass die Ergebnisse keine zu hohe Variabilität besitzen und es somit nicht an Konsistenz fehlt. Eine Simulation, die bei jedem Durchgang sehr verschiedene Ergebnisse liefert, ist nur schwer für die Vorhersage von Ereignissen nutzbar.

Die Sensitivitätsanalyse wiederum ist daher so wichtig, da sich durch sie Parameter ohne Effekt entfernen lassen. Dies erleichtert die weiteren Tests und reduziert die Komplexität des Modells, was sich positiv auf die Verständlichkeit und Performanz auswirkt.

Die Kalibrierung dient dazu, die richtigen Werte für die Parameter zu finden, so dass das strukturell valide Modell auch valide Ergebnisse liefert. Dies ist vor allem dann wichtig, wenn nicht alle Werte der Parameter eindeutig bekannt sind, weil sie sich im Referenzsystem nicht oder nur sehr schwer beobachten lassen. Die Kalibrierung ist dann die einzige Möglichkeit das ansonsten valide Modell nutzbar zu machen.

Ansonsten sind sich die Verfahren von Klügl und Xiang et al. sehr ähnlich. Bei beiden Verfahren ist die Face Validation ein wichtiger Bestandteil und beide validieren die Abläufe der Simulation auf diese Weise. Klügls Verfahren umfasst allerdings eine größere Auswahl an Methoden zur Face Validation. Vor allem die immersive Betrachtung der Simulation durch die Augen eines Agenten stellt eine gute Möglichkeit zur Beobachtung des Geschehens dar und kann das Verhalten der Agenten besser validieren als eine Vogelperspektive, die allerdings für die Beobachtung des gesamten Ablaufes und von Staus besser geeignet ist. Verschiedene Sichten zur Validierung von verschiedenen Aspekten der Simulation sind also sinnvoll.

Der letzte Unterschied besteht in der Validierung im Betrieb. Klügls Verfahren setzt hierbei auf den Vergleich mit statistischen Daten, während das Verfahren von Xiang et al. einen Vergleich mit anderen Simulationen durchführt. Welches Verfahren besser ist, hängt davon ab, inwiefern sich statistische Daten zu Vergleichszwecken mit dem Referenzsystem sammeln lassen. Ist dies problemlos möglich, so steht einer statistischen Validierung nichts im Wege. Sollte dies allerdings Probleme machen, so ist ein Vergleich mit anderen validen Simulationen vorzuziehen.

Bei den speziellen Verfahren zur Validierung von Gruppensimulationen handelt es sich ausschließlich um Verfahren zur Validierung im Betrieb. Die vorgestellten Verfahren sind der

Vergleich der Simulation mit Evakuierungsübungen, webbasierte Face Validation, Validierung durch Immersion und Virtual Reality, statistische Validierung, mathematische Validierung und die Validierung unter Nutzung einer Richtlinie. Im Folgenden werden für diese Verfahren die Vor- und Nachteile ermittelt sowie die Anforderungen, die diese an die Simulation und die Rahmenbedingungen, wie das Vorhandensein von bestimmten Daten, stellen.

Die Durchführung von Evakuierungsübungen mit dem Ziel des Vergleichs mit der Simulation bietet den Vorteil, dass eine mögliche Situation, die die Simulation simulieren können soll, nachgestellt werden kann. Hierdurch ist es möglich Daten zu Situationen zu erlangen, zu denen es keine Daten von realen Fällen gibt, die für die Validierung der Simulation verwendet werden können. Außerdem besteht die Möglichkeit, die für die Simulation wichtigen Aspekte der Übung zu messen und aufzuzeichnen. So könnten unter anderem die Herzschlagfrequenzen der Teilnehmer gemessen werden oder die Teilnehmer befragt werden, was sie während der Evakuierung gedacht haben und was ihre Entscheidungsfindung beeinflusst hat. Solche Daten hätte man aus realen Situationen nur teilweise. Personen, die an einer echten Evakuierung beteiligt waren, könnten ebenfalls befragt werden, Messungen wären jedoch nicht mehr möglich. Gegen den Vergleich mit Evakuierungsübungen spricht allerdings, dass die Ergebnisse nur schwer mit denen von realen Situationen vergleichbar sind. Die Teilnehmer sind sich bewusst, dass keine Gefahren bestehen und handeln somit ohne Einfluss von Stress und sind auf die Geschehnisse vorbereitet. Dadurch wird sich das Handeln von dem Handeln unter realen Bedingungen unterscheiden. Bei solchen Vergleichen wird zudem meistens nur die Simulation als Ganzes validiert und nicht das Verhalten der einzelnen Teilnehmer mit dem der Agenten verglichen. Außerdem sind die Ergebnisse des Vergleichs subjektiv, weil nicht anhand von festen Bewertungskriterien validiert wird, durch die zum Schluss die Eignung auf eine objektive Weise, zum Beispiel als eine Punktzahl, festgehalten werden kann. Ein weiterer Nachteil ist der hohe Aufwand, der mit der Organisation und Durchführung von solchen Übungen verbunden ist, und dass genügend Freiwillige für die Teilnahme vorhanden sein müssen. Besonders große Evakuierungen und gefährliche Situationen lassen sich auf diese Weise nicht als Übung durchführen.

Tabelle 1: Vergleich der Vor- und Nachteile beim Vergleich der Simulation mit Übungen

Vergleich mit Übungen	
Vorteile	Nachteile
<ul style="list-style-type: none"> + gewünschte Situation nachstellen + Messung und Aufzeichnung wichtiger Aspekte der Situation 	<ul style="list-style-type: none"> - Ergebnisse anders als in realen Situationen - kein Vergleich des Verhaltens von Teilnehmern und Agenten - subjektiver Vergleich - hoher Aufwand - erfordert viele Freiwillige - große und/oder gefährliche Situationen

	nicht als Übung durchführbar
--	------------------------------

Die webbasierte Face Validation bietet den Vorteil, dass keine aufwendige Analyse der Videos erforderlich ist. Es muss nur dafür gesorgt werden, dass simuliertes und beobachtbares Verhalten miteinander vergleichbar ist. Auch auf technische Art und Weise ist das Verfahren vorteilhaft, da entgegen der Anforderungen von Algorithmen, für die die Szenen möglichst vergleichbar dargestellt sein müssen, die Perspektive oder Geschwindigkeit der verglichenen Aufnahmen für die menschlichen Teilnehmer einen geringeren Einfluss auf das Ergebnis der Vergleiche hat. Außerdem lässt sich das Verfahren auch ohne großen Hintergrund an wissenschaftlichen Daten ausführen, da nur vergleichbare Aufnahmen vorhanden sein müssen und die Bewertung durch die Teilnehmer subjektiv erfolgt. Durch die Face Validation lässt sich das Verhalten der Simulation und Agenten auf die Realitätsnähe überprüfen. Ein weiterer Vorteil ist die Umsetzung über das Internet, diese erlaubt, dass die Teilnahme an dem Verfahren mit weniger Aufwand für die Teilnehmer verbunden ist und sich möglicherweise mehr Teilnehmer finden. Nachteilig ist, dass nur kurze Ausschnitte miteinander verglichen werden und kein Vergleich von längeren Situationen erfolgt. Es wird also nur das generelle Verhalten in diesem kurzen Ausschnitt validiert und nicht die gesamte Simulation. Auch eine Validierung des Verhaltens der einzelnen Agenten findet nicht statt. Außerdem wird in dem beschriebenen Verfahren nicht von Experten auf diesem Gebiet validiert, sondern von „normalen“ Menschen. Auch müssen für die Durchführung genügend Teilnehmer vorhanden sein, damit verlässlichere Ergebnisse gewonnen werden können. Die Ergebnisse dieses Verfahrens beruhen immer auf subjektiven Einschätzungen, was die Verlässlichkeit der Ergebnisse fragwürdig erscheinen lässt.

Tabelle 2: Vergleich der Vor- und Nachteile der webbasierten Face Validation

Webbasierte Face Validation	
Vorteile	Nachteile
+ keine aufwendige Analyse der Videos erforderlich	- nur kurze Ausschnitte werden verglichen
+ technisch leicht umsetzbar	- Vergleich nur von generellem Verhalten
+ erfordert nur die Vergleichsaufnahmen	- keine Validierung des Agentenverhaltens
+ Verhalten von Simulation und Agenten auf Realitätsnähe überprüfbar	- keine Verwendung von Experten
+ wenig Aufwand für Teilnehmer	- genügend Teilnehmer für verlässliche Ergebnisse erforderlich
	- subjektive Einschätzungen

Die Validierung durch Immersion und Virtual Reality vereinfacht die Überprüfung des Gesamtverhaltens der Simulation und des Verhaltens von einzelnen Agenten. Dadurch, dass der Teilnehmer an dieser Art von Verfahren sich in der Simulation bewegen kann und alle Geschehnisse so wahrnehmen kann, als ob er wirklich unter den Agenten wäre, fällt es

leichter das Verhalten der Agenten und das Gesamtverhalten zu beurteilen. Dies liegt darin begründet, dass die Wahrnehmung der Agenten durch diese Darstellung besser mit der Wahrnehmung wie sie auch in der Realität ist, verglichen werden kann. Die Nutzung der Immersion hat außerdem den Vorteil, dass sie es in Zukunft erlauben könnte, bei den Teilnehmern das gleiche Verhalten, die gleichen Emotionen sowie die gleichen psychischen und physischen Reaktionen zu beobachten wie bei Personen, die sich tatsächlich in der hier nur simulierten Situation befinden. Dadurch könnten solche Daten auch für Situationen gewonnen werden, für die dies normalerweise sehr schwer oder nicht möglich wäre. Allerdings ist es hierfür zuerst nötig, Simulationen zu haben, die so stark der Wirklichkeit entsprechen, dass dieser Effekt eintritt, was heutzutage noch ein Problem darstellt. Ein weiterer Vorteil ist, dass auch bei diesem Verfahren kein großer Hintergrund an wissenschaftlichen Daten vorhanden sein muss, da die Validierung durch die Teilnehmer über Face Validation erfolgt. Allerdings handelt es sich auch hier um subjektive Meinungen und in dem vorgestellten Fall nicht um Experten, die die Beurteilung übernehmen. Ein weiterer Nachteil ist, dass nicht klar ist, wie verlässlich die durch Beobachtung der Teilnehmer gewonnenen Daten sind, da für absolut verlässliche Daten auch eine absolut realistische Simulation und Immersion gegeben sein müsste. Dies scheint nicht nur technisch problematisch. Auch ist es unwahrscheinlich, dass sich Personen freiwillig in lebens echten Gefahrensituationen befinden wollen. Ohne Angst und Stress sind allerdings keine mit der Realität vergleichbaren Daten zu erlangen. Auch ansonsten ist es nachteilig, dass für diese Art der Validierung Teilnehmer benötigt werden, da es sich dadurch nicht zu beliebigen Zeitpunkten und so oft wie gewünscht, durchgeführt werden kann. Außerdem stellt sie aufgrund der Forderung von Immersion hohe Anforderungen an die Simulation, die für das eigentliche Ergebnis der Simulation, zum Beispiel die Vorhersage von Evakuierungen, keine Rolle spielen. Insgesamt handelt es sich um eine sehr aufwendige Form der Validierung, sowohl in Bezug auf die geforderte Immersion als auch in Bezug auf die Durchführung der Validierung mit Virtual Reality.

Tabelle 3: Vergleich der Vor- und Nachteile der Nutzung von Immersion und Virtual Reality

Nutzung von Immersion und Virtual Reality	
Vorteile	Nachteile
<ul style="list-style-type: none"> + vereinfacht Überprüfung des Verhaltens von Simulation und Agenten + Verhalten, Emotionen und psychische und physische Reaktionen der Teilnehmer beobachtbar oder messbar + keine große Menge an Vergleichsdaten erforderlich 	<ul style="list-style-type: none"> - subjektive Meinungen - keine Verwendung von Experten - Verlässlichkeit der durch Beobachtung der Teilnehmer gewonnenen Daten fragwürdig - Teilnehmer erforderlich - hohe zusätzliche Anforderungen an Simulation - Durchführung aufwendig

Die statistische Validierung bietet den Vorteil, dass sie größtenteils objektiv ist. Nur der festgelegte Wert, der definiert, ob es sich bei einem gemessenen Wert um einen Ausreißer beziehungsweise eine andere Situation handelt, wird subjektiv bestimmt. Durch die Objektivität und die fest vorgeschriebene Berechnung der Werte hat dieses Verfahren außerdem den Vorteil, auch automatisiert durchgeführt werden zu können. Die objektiven Ergebnisse der Validierung können für Vergleiche von Simulationen oder für Optimierung verwendet werden, indem überprüft wird, ob sich die Ergebnisse durch Änderungen verbessern oder verschlechtern. Allerdings erfolgt keine Validierung auf der Agentenebene, sondern es wird nur die Makroebene beachtet, indem die Werte der Agenten gemittelt werden. Dadurch ist ein realistisches Verhalten auf der Agentenebene nicht garantiert. Zudem werden für die Validierung sehr gute Aufnahmen benötigt, aus denen sich viele Daten extrahieren lassen, die der Validierung der Simulation behilflich sind. Da es sich um ein statistisches Verfahren handelt, werden außerdem möglichst viele solche Aufnahmen gebraucht, die auch noch vergleichbare Daten liefern können müssen, damit eine ausreichend große historische Datenmenge zur Verfügung steht. Gerade bei Situationen wie Evakuierungen ist dies problematisch. Auch die Analyse und Aufbereitung der Daten stellt einen großen Aufwand dar. Es findet außerdem keine komplette Validierung des Ablaufes in der Simulation statt, da nur zufällig bestimmte Bilder der Aufnahmen und Simulation betrachtet werden. Komplexe Abläufe lassen sich so daher nicht validieren, sondern nur die Zusammensetzung der betrachteten Menschengruppe.

Tabelle 4: Vergleich der Vor- und Nachteile der statistischen Validierung

Statistische Validierung	
Vorteile	Nachteile
+ größtenteils objektiv	- keine Validierung der Agentenebene
+ kann automatisiert werden	- hohe Anforderungen an Aufnahmen
+ vergleichbare Ergebnisse	- viele vergleichbare Aufnahmen erforderlich
+ Verwendung für Optimierungen	- Analyse der Aufnahmen aufwendig
+ Validierung der Zusammensetzung von Menschengruppen	- Abläufe nicht validierbar

Die mathematische Validierung besitzt ähnliche Vorteile. Auch sie ist objektiv, da feste Berechnungen definiert sind und letztendlich vergleichbare Werte als Ergebnis zur Verfügung stehen. Der Grenzwert wiederum ist ein subjektiver Faktor. Auch die Ergebnisse dieses Verfahrens lassen sich für Vergleiche von Simulationen oder zur Optimierung einsetzen. Aus den festen Berechnungen ergibt sich wieder, dass das Verfahren automatisiert durchgeführt werden kann. Im Gegensatz zum statistischen Verfahren muss für dieses Verfahren bestimmt werden, wie sich vergleichbare Bilder aus Simulation und realer Aufnahme finden lassen. Wie das statistische Verfahren führt dieses Verfahren keine Validierung auf Agentenebene durch, wodurch kein realistisches Verhalten der Agenten garantiert ist. Dadurch,

dass in diesem Verfahren nur die Anzahl an Agenten in verschiedenen Bereichen überprüft wird, werden sehr viel weniger Werte betrachtet, als dies in der statistischen Validierung der Fall ist. Jedoch kann trotzdem durch das Vergleichen der Bilder der Aufnahme und Simulation von gleichen Zeitpunkten ein genauerer Vergleich von Abläufen erfolgen. Auch für dieses Verfahren müssen Aufnahmen für die betrachteten Bereiche vorhanden sein. Allerdings spielt die Vergleichbarkeit der Aufnahmen mit den Bildern der Simulation keine große Rolle, da nur die Anzahl an Personen und Agenten in den Bereichen wichtig ist. Auch wird nur eine Aufnahme benötigt um den Vergleich durchführen zu können. Somit ist dieses Verfahren mit wesentlich weniger Aufwand verbunden. Ein weiteres Problem liegt in der Anpassung der Verteilung in der Simulation, wenn diese zu stark von den Daten aus der Realität abweicht. Dadurch, dass nur gezählt wird wie oft Anpassungen durchgeführt werden mussten und nicht beachtet wird, was in der Simulation nicht richtig lief, gehen wichtige Informationen zur Validierung verloren. Der Vorteil der Vergleichbarkeit wird entkräftet, da sich an der Anzahl der Korrekturen nicht die Schwere der Abweichung feststellen lässt und auch sonst die Probleme nicht direkt erkannt werden.

Tabelle 5: Vergleich der Vor- und Nachteile der mathematischen Validierung

Mathematische Validierung	
Vorteile	Nachteile
+ größtenteils objektiv	- Synchronisierung von Aufnahme und Simulation erforderlich
+ kann automatisiert werden	- keine Validierung der Agentenebene
+ Vergleichbare Ergebnisse	- Agentenwerte werden nicht betrachtet
+ Verwendung für Optimierungen	- Anpassbarkeit der Agentenpositionen in Simulation erforderlich
+ Vergleich von Abläufen	- Grund für Abweichungen zwischen Simulation und Aufnahme und Schwere des Problems nicht ersichtlich
+ geringere Anforderungen an Aufnahmen	
+ nur eine Vergleichsaufnahme erforderlich	
+ geringerer Aufwand	

Die Validierung mit Richtlinien, in diesem Falle die „Richtlinie für Mikroskopische Entfluchtungsanalysen“, bietet den Vorteil, dass es sich ebenfalls um feste Tests handelt, wodurch eine gewisse Objektivität gegeben ist und die Tests nach der Erstellung automatisiert durchgeführt werden können. Die Auswertung erfolgt zwar nicht automatisch, sondern muss von Menschen durchgeführt werden, allerdings sind die Erwartungen an die Ergebnisse der einzelnen Tests vorgegeben. Um die Validierung durchzuführen werden außerdem keine zusätzlichen Daten zu denen aus der Richtlinie benötigt. Die Durchführung der Tests ist schon in frühen Phasen der Entwicklung möglich. Dies liegt allerdings darin begründet, dass die Tests vor allem der Verifizierung der Simulation dienen und die Validierung nur einen geringen Anteil hat. In frühen Phasen ist eine Verifizierung jedoch auch sinnvoll um Fehler in der Implementierung zu finden. Außerdem ermöglichen die Tests das Auffinden

von fehlenden Funktionen, die die Simulation bieten sollte, oder von fehlerhaft implementierten Funktionen. Von Nachteil ist, dass auch in diesem Verfahren keine Validierung auf der Agentenebene erfolgt, sondern nur die Makroebene betrachtet wird. Dies liegt darin begründet, dass die Tests nicht das Verhalten der Agenten berücksichtigen. Außerdem sind nur Tests für die geregelte Evakuierung vorhanden, mögliche Ereignisse wie Feuer oder Wassereintrich werden nicht beachtet.

Tabelle 6: Vergleich der Vor- und Nachteile der Validierung durch die RiMEA

Validierung durch die RiMEA	
Vorteile	Nachteile
+ Objektivität durch feste Tests + kann automatisiert werden + Richtlinie gibt Daten für Evakuierungssimulationen vor + schon in frühen Entwicklungsphasen (teilweise) durchführbar + ermöglicht Auffinden von fehlenden oder fehlerhaften Funktionen	- vor allem Verifizierung statt Validierung - keine Validierung auf Agentenebene - testet nur geregelte Evakuierung

Wie gezeigt wurde, haben die verschiedenen Verfahren alle ihre Vor- und Nachteile. Zusammenfassend kann gesagt werden, dass die subjektiven Verfahren, also Vergleiche mit Übungen, Face Validation und Validierung durch Immersion und Virtual Reality den Nachteil haben, dass sie subjektive Ergebnisse liefern, wodurch die Vergleichbarkeit und Verlässlichkeit der Ergebnisse nicht gewährleistet werden kann. Allerdings ist es gerade durch Face Validation mit Betrachtung und Vergleich von realen Aufnahmen und Simulation und Nutzung von Virtual Reality möglich, das Verhalten auf der Makro- und Mikroebene zu validieren. Die objektiven Verfahren, also statistische Validierung, mathematische Validierung und die Validierung mit der RiMEA ermöglichen zwar eine gewisse Objektivität und dadurch die Vergleichbarkeit von Ergebnissen, sie erlauben allerdings keine direkte Validierung des Verhaltens. Es werden stets nur Werte auf der Makroebene verglichen, somit ist zwar gewährleistet, dass die Simulation im Hinblick auf diese Werte valide ist, allerdings wird nicht gewährleistet, dass das Verhalten in der Simulation auch realistisch ist. Bei agentenbasierten Simulationen wollen wir jedoch auch ein realistisches Verhalten auf der Agentenebene, weil sich sonst die Ergebnisse nicht groß von Simulationen unterscheiden würden, die nur auf der Makroebene arbeiten, wie zum Beispiel Fluss-Modelle. Bei der Validierung von Gruppensimulationen sollten also subjektive und objektive Verfahren kombiniert eingesetzt werden, so dass einerseits eine Validierung des Verhaltens stattfindet und andererseits auch eine Vergleichbarkeit erreicht wird und der Vorteil einer schnellen automatisierten Validierung zum Beispiel nach Änderungen an der Simulation ausgenutzt werden kann. Stehen nicht genügend wissenschaftlich belegte Daten oder Aufzeichnungen zur Validie-

rung durch die objektiven Verfahren zur Verfügung, so ist immerhin die Face Validation trotzdem durchführbar.

3.3 Erstellung eines konkreten Validierungskonzeptes

In den vorherigen Teilen dieses Kapitels wurde gezeigt, dass die verschiedenen Entwicklungsphasen einer Gruppensimulation unterschiedliche Arten der Validierung erfordern. Während der ersten Phase, dem Sammeln von Daten, ist die Datenvalidität zu gewährleisten. In der zweiten Phase, der Bildung des konzeptuellen Modells, muss die konzeptuelle Validität gewährleistet sein. In der dritten Phase, der Implementierung des Modells, ist schließlich die Validität im Betrieb zu überprüfen. In diesem Teil der Arbeit wird nun ein Konzept für die Validierung von Gruppensimulationen, beziehungsweise im speziellen Evaluierungssimulationen, in den einzelnen Phasen erstellt.

Für die Überprüfung der Datenvalidität gibt es wie erwähnt nur die Möglichkeit, bereits während der Gewinnung der Daten und deren Bearbeitung dafür zu sorgen, dass es zu keinen fehlerhaften Daten kommt. Dazu sind verlässliche Methoden notwendig. Ansonsten kann nur noch nach Ausreißern gesucht und diese auf Korrektheit überprüft werden.

Für die konzeptuelle Validierung müssen die Theorien und Annahmen und die daraus entwickelten Regeln auf ihre Korrektheit überprüft werden und es muss die Darstellung der Problementität auf ihre Sinnhaftigkeit überprüft werden. Hierzu können mathematische Analysen und statistische Methoden auf Daten der Problementität angewendet werden. Allerdings mangelt es gerade im Bereich der Gruppensimulationen oft an einer ausreichenden Menge an Daten, um dies problemlos durchzuführen. Aus diesem Grund sollte eine Face Validation des konzeptuellen Modells erfolgen. Experten aus den betrachteten Forschungsbereichen sollten die aufgestellten Theorien, Annahmen und Regeln auf Korrektheit und die Darstellung der Problementität auf Sinnhaftigkeit überprüfen. Auch wenn diese Validierung subjektiv ist und keine absolute Korrektheit gewährleisten kann, gibt es keine Alternativen hierzu in der Forschung zur Validierung von Gruppensimulationen.

Die Validierung im Betrieb muss überprüfen, ob die Simulation das betrachtete Problem auf valide Art und Weise simuliert und das Ausgabeverhalten eine ausreichende Genauigkeit für den geplanten Zweck des Modells besitzt. Es ist sinnvoll als erstes die interne Validität zu überprüfen, um zu zeigen, dass die Simulation auch wirklich verlässliche Ergebnisse liefert. Sollten die Ergebnisse von wiederholten Läufen der Simulation mit den gleichen Daten aufgrund der stochastischen Eigenschaften zu stark voneinander abweichenden Ergebnissen führen, so ist die Simulation nicht für die Vorhersage von Ereignissen nutzbar und eine weitere Validierung ist nicht sinnvoll, bis dieses Problem behoben wurde. Ansonsten sollte mit der Sensitivitätsanalyse begonnen werden, um ein minimales Modell zu schaffen, was den Aufwand für die nachfolgenden Validierungsschritte und die Ausführung der Simulation verringern kann. Hiernach kann mit der eigentlichen Validierung im Betrieb begonnen werden.

Wie im vorherigen Teil gezeigt wurde, ist es sinnvoll eine Kombination aus subjektiven und objektiven Verfahren zu benutzen. Da alle Verfahren Vorteile besitzen, könnten theoretisch auch alle durchgeführt werden, doch dies würde einen zu großen Aufwand bedeuten. Aus diesem Grund werden nur die Verfahren benutzt, die die verlässlichsten Ergebnisse liefern. Der Vergleich der Simulation mit Übungen erscheint nicht verlässlich genug, da in Übungen keine realistischen Situationen geschaffen und beobachtet werden können. Im Bereich der Verfahren, die Face Validation einsetzen, stehen die Animationsbetrachtung aus der Vogelperspektive und die Betrachtung aus der Egoperspektive eines Agenten unter möglicher Nutzung von Virtual Reality zur Wahl. Diese Verfahren sollten beide eingesetzt werden, da die Animationsbetrachtung aus der Vogelperspektive es sehr gut erlaubt, das allgemeine Gruppenverhalten zu betrachten, während die Nutzung der Egoperspektive vor allem die Betrachtung der einzelnen Agenten vereinfacht. Ist der Agent aus der Egoperspektive steuerbar, so können die Reaktionen der anderen Agenten auf die Aktionen des gesteuerten Agenten besser beobachtet werden. Virtual Reality kann optional für die Egoperspektive eingesetzt werden, um den Testern eine realistischere Wahrnehmung und Interaktion mit der Simulation zu bieten, was zu einer besseren Vergleichbarkeit mit der Realität führen kann. Die Verfahren sollten wenn möglich mit Experten durchgeführt werden, da dies für verlässlichere Ergebnisse sorgt. Der Ansatz der Nutzung von Immersion zur Gewinnung von nutzbaren Daten scheint mit heutiger Technik noch nicht möglich und auch sonst ist es fragwürdig, ob dieser Ansatz wirklich funktionieren kann. Die Virtual Reality sollte daher wirklich nur zur Betrachtung des Verhaltens genutzt werden. Liegen vergleichbare Aufnahmen zu den in der Simulation simulierten Szenarien vor, so können zusätzlich eine Gegenüberstellung und ein Vergleich der Simulation und der Aufnahmen erfolgen.

Auf Seiten der objektiven Verfahren stehen die statistische Validierung, die mathematische Validierung und die Validierung durch die RiMEA zur Verfügung. Hier sollte die mathematische Validierung der statistischen Validierung vorgezogen werden, da die statistische Validierung keine komplexen Abläufe validieren kann. Außerdem stellt die mathematische Validierung geringere Anforderungen in Hinblick auf das Vorhandensein von nutzbaren Aufnahmen. Die statistische Validierung kann jedoch ergänzend durchgeführt werden, um eine zusätzliche Validierung der Zusammensetzung der Gruppe in der Simulation durchzuführen. Das Verhalten der einzelnen Agenten wird von beiden Verfahren nicht berücksichtigt, deshalb sollten die objektiven Verfahren, solange dies der Fall ist, nur eine Ergänzung zu den subjektiven Verfahren darstellen. Beide Verfahren haben jedoch den Vorteil, dass zumindest objektive und vergleichbare Ergebnisse erlangt werden können. Die Gewichtung der Ergebnisse dieser beiden Verfahren sollte jedoch zurzeit noch geringer als die der subjektiven Verfahren sein. Bei der mathematischen Validierung sollten die Zeitpunkte, an denen die Abweichung den Grenzwert überschritten hat, erneut durch Face Validation betrachtet werden, um den Grund für die Abweichung und die Schwere des Problems bestimmen zu können. Die Benutzung der RiMEA ist in frühen Entwicklungsphasen sinnvoll. Die Tests sollten bei der Erstellung der implementierten Simulation durchgeführt werden, um das Vor-

handensein und das Funktionieren von Grundfunktionen der Simulation zu gewährleisten. Degenerierungstests sind sinnvoll, um das Langzeit-Verhalten der Simulation auf Fehler zu überprüfen.

Schlägt die Validierung im Betrieb fehl, obwohl das konzeptuelle Modell valide ist, so kann durch Kalibrierung versucht werden, die Simulationsparameter so anzupassen, dass valide Ergebnisse erzielt werden. Kalibrierung kann auch eingesetzt werden, wenn die objektive Validierung nicht fehlgeschlagen ist, aber noch genauere Ergebnisse der Simulation möglich wären und erwünscht sind. Mit Hilfe der objektiven und der subjektiven Validierung im Betrieb muss bestimmt werden, ob die Kalibrierung erfolgreich war und ob das Agentenverhalten auch nach den Änderungen noch valide ist.

Werden die ausgewählten Verfahren eingesetzt, ist es möglich, das Verhalten auf der Mikro- und Makroebene durch die subjektiven Verfahren zu validieren. Durch die objektiven Verfahren wird außerdem eine Vergleichbarkeit der Ergebnisse gewährleistet. Außerdem können diese automatisch durchgeführt werden, was sehr hilfreich ist, da die Validierung der Simulation bei Änderungen erneut ausgeführt werden sollte.

Die Schritte des Validierungskonzeptes erfolgen also in dieser Reihenfolge:

1. Während des Sammelns von Daten müssen durch verlässliche Methoden zum Sammeln und Pflegen der Daten Fehler vermieden werden und es muss zusätzlich nach Ausreißern gesucht werden, die auf Korrektheit überprüft werden müssen, um valide Daten zu erhalten.
2. Die konzeptuelle Validierung erfolgt nach der Erstellung des konzeptuellen Modells durch Experten, die die Theorien, Annahmen und Regeln auf Korrektheit und die Darstellung der Problementität auf Sinnhaftigkeit überprüfen.
3. Während der Implementierung des Modells kann durch die RiMEA das Vorhandensein und das Funktionieren von Grundfunktionen überprüft werden.
4. Degenerierungstests erlauben, das Langzeit-Verhalten der Simulation auf Fehler zu überprüfen.
5. Um die Zuverlässigkeit der Ergebnisse des Modells sicherzustellen, wird die interne Validität überprüft.
6. Durch Sensitivitätsanalyse wird versucht ein minimales Modell zu schaffen, um die Komplexität des Modells zu verringern.
7. Die Validierung im Betrieb erfolgt mit dem zuvor getesteten und minimierten implementierten Modell und besteht aus zwei Teilen.
 - a) Subjektive Validierung: Es erfolgt Face Validation durch Experten, die sich die laufende Simulation aus der Vogelperspektive und aus der Egoperspektive eines Agenten, der sich in der Umgebung befindet, beobachten. Durch aus der Egoperspektive steuerbare Agenten können die Reaktionen der anderen Agenten auf die Aktionen des gesteuerten Agenten besser beobachtet werden. Vir-

tual Reality kann optional für die Darstellung der Egoperspektive und die Interaktion mit der Simulation eingesetzt werden, um bessere Ergebnisse zu erzielen. Liegen vergleichbare Aufnahme zu den simulierten Szenarien vor, so können diese zusätzlich der Simulation gegenübergestellt und verglichen werden.

- b) Objektive Validierung: Es erfolgt mathematische Validierung zur zusätzlichen Überprüfung der Makroebene. Die Ergebnisse dieses Schrittes sollten als Ergänzung zur subjektiven Validierung gesehen werden. Die mathematische Validierung sollte mit Face Validation verknüpft werden. Die Zeitpunkte mit zu großer Abweichung sollten erneut betrachtet werden, um den Grund für die Abweichung und die Schwere des Problems bestimmen zu können. Wenn möglich kann zusätzlich noch die statistische Validierung durchgeführt werden.
8. Schlägt die Validierung im Betrieb fehl, so kann durch Kalibrierung versucht werden, die Simulationsparameter so anzupassen, dass valide Ergebnisse erzielt werden. Dies kann auch sonst zur Verbesserung der Ergebnisse der Simulation eingesetzt werden. Danach sollten erneut die objektive und die subjektive Validierung im Betrieb durchgeführt werden, um einerseits zu prüfen, ob die Kalibrierung erfolgreich war und andererseits um zu prüfen, ob das Agentenverhalten weiterhin realistisch erscheint.

Das vorgestellte Konzept basiert zu einem großen Teil auf Face Validation. Dies lässt sich zurzeit allerdings wegen einer nicht ausreichenden Menge an Daten, die für die Validierung genutzt werden könnten und dem Problem, dass es zurzeit keine Möglichkeit gibt, das Verhalten sowohl auf Makro- als auch auf Mikroebene auf objektive Art und Weise zu vergleichen, nicht verhindern. Es ist daher wichtig, dass in Zukunft ein Weg gefunden wird, mehr nutzbare Daten zu sammeln und neue Verfahren zum Validieren von Verhalten auf objektive Art und Weise zu entwickeln. Nur dann ist eine noch verlässlichere Validierung möglich.

3.4 Validierung am Beispiel des WALK-Projektes

In diesem Kapitel wird betrachtet, wie das entwickelte Validierungskonzept auf eine konkrete Gruppensimulation angewendet werden kann. Bei dieser handelt es sich um die Gruppensimulation WALK, die an der HAW Hamburg entwickelt wird. Diese wird zuerst vorgestellt und die Besonderheiten werden hervorgehoben. Danach werden der bisherige Entwicklungsstand und die bisher durchgeführte Validierung betrachtet. Als nächstes wird betrachtet, wie das Validierungskonzept auf WALK angewendet werden kann. Dazu werden die Anforderungen und Rahmenbedingungen, die das Validierungskonzept an WALK stellt, verdeutlicht. Es wird überprüft, ob die Anforderungen bereits erfüllt werden. Für nicht erfüllte Anforderungen wird geklärt, wie diese in WALK umgesetzt werden könnten und in welchem Teil von WALK sie zu realisieren sind. Danach wird dargestellt, wie eine Validierung von WALK ausgeführt werden soll. Die Anwendung der einzelnen Schritte in den verschiedenen Entwicklungsphasen wird beschrieben und es wird erklärt wie die Ergebnisse dieser Schritte zu bewerten sind beziehungsweise wozu sie verwendet werden können. Dabei werden die Besonderheiten von WALK bei der Validierung beachtet.

3.4.1 Vorstellung des WALK-Projektes

Bei WALK handelt es sich um eine emotionsbasierte Simulation von Fußgängern in Evakuationszenarien, die an der HAW Hamburg entwickelt wird. Zu WALK gab es bereits Veröffentlichungen, die das Projekt im Allgemeinen (Thiel-Clemen et al. 2011), das Geoinformationssystem (Baldowski 2011) sowie das Multiagentensystem (Thiel 2011) und die Agenten (Roiss 2011) beschreiben. Die Simulation beruht auf einem agentenbasierten Modell und simuliert auf der Nanoebene. Dem zugrunde liegt eine Abspaltung einer noch feineren Ebene von der Mikroebene. Die Mikroebene umfasst Modelle, die auf zellulären Automaten und sozialen Kräften basieren. Auf der Nanoebene befinden sich individuenbasierte Simulationen und agentenbasierte Simulationen. Diese Aufteilung wurde unternommen, da bei den Simulationen der Nanoebene jedes Agentenobjekt seine eigenen Fakten und zugehörigen Methoden besitzt, um mit ihnen zu arbeiten. In mikroskopischen Modellen dagegen folgen alle Individuen den gleichen Regeln und Gesetzen. Das Ziel von WALK ist es, ein verlässliches agentenbasiertes Modell für die Vorhersage von menschlichem Verhalten in kritischen Situationen, unter Berücksichtigung von emotionalem Stress und komplexen Interaktionen zwischen Individuen, zu entwickeln. Der Haupteinsatzzweck des Systems ist zwar die Simulation von Fußgängerbewegungen, allerdings soll die Plattform möglichst unabhängig von der ausgeführten Simulation sein, so dass sie auch Anwendung in anderen Bereichen finden kann.

Die Überlegung hinter WALK basiert darauf, dass in den bisherigen Ansätzen Emotionen nicht beachtet wurden und die Bewegung der Agenten von mechanischen Gesetzen abhängig war, wodurch das Risiko groß war, wichtige Aspekte, die auf die Bewegung Einfluss ha-

ben, zu vernachlässigen. Eine Simulation zum Einfluss von Emotionen auf das Verhalten von Personen hat deutlich gemacht, dass Emotionen großen Einfluss auf die Simulationsergebnisse haben (Thiel-Clemen und Klingenberg 2010). WALK soll daher auch verschiedene Persönlichkeiten und Stimmungslagen für die Agenten besitzen. Auch Unterschiede in der Kultur und Zusammensetzung der Personen sollen beachtet werden, da diese Einflüsse auf den Fluss der Agenten haben. Die Agenten besitzen außerdem bestimmte Ziele und Unterziele, die sie erreichen wollen, sowie Aktionen, die sie zum Erreichen dieser durchführen können. Ziele, Unterziele und Aktionen befinden sich in einer Hierarchie. Neue Unterziele können außerdem dynamisch hinzukommen und werden in die Hierarchie eingebaut. Der Zustand eines Agenten wird in WALK nicht nur durch Zustandsvariablen wie die geografische Position, Ausrichtung und ähnlichem beschrieben, sondern enthält auch individuelle physiologische und psychologische Details des Agenten. Eine weitere Besonderheit ist, dass die Agenten komplexe individuelle Wahrnehmung und Entscheidungsprozesse besitzen. Auch sonst sind WALK-Agenten stark individualisiert und besitzen persönliche Eigenschaften, Alter, Körpergröße und mehr.

Bei den in WALK verwendeten emotionalen Agenten handelt es sich um erweiterte kognitive Agenten. Kognitive Agenten planen ihre Handlungen aufgrund ihrer Ziele und des Nutzens der Handlungen. Bei emotionalen Agenten werden noch zusätzlich die Emotionen simuliert, die einen Einfluss auf das Verhalten des Agenten haben. Hierdurch lässt sich realitätsnäheres Verhalten simulieren, da Emotionen auf die Interaktion zwischen Menschen und die Wahrnehmung der Umgebung einen Einfluss haben. So wird unter anderem durch Stress weniger von der Umgebung wahrgenommen. Ein anderes Beispiel wäre ein Familienvater, der zuerst seine Familie sucht, bevor er das zu evakuierende Gebäude verlässt. Die in WALK verwendeten Agenten verfügen daher über eine eigene Wegfindung, so dass sie sich nicht nur an die vorgegebenen Rettungswege halten können.

Eine weitere Besonderheit von WALK ist, dass dynamische Ereignisse wie die Ausbreitung von Feuer, Rauch oder Wasser simuliert werden sollen. Unter anderem hierzu soll es möglich sein, die Simulation jederzeit zu pausieren und Veränderungen an der Simulation vorzunehmen können. Des Weiteren sollen sowohl Gebäude als auch das sich darum befindende Gelände in der Simulation betrachtet werden. Für diese Punkte wird ein dreidimensionales Geoinformationssystem (GIS) entwickelt, das die räumlichen Daten, also die statischen Geländedaten und die dynamisch simulierten Daten, liefert. Dieses GIS besitzt verschiedene Informationsschichten, die die verschiedenen Informationen wie Gebäudeplan, Feuerausbreitung, Rettungswege und viele weitere umfassen. Hierdurch wird die voneinander unabhängige Betrachtung dieser Informationen ermöglicht.

Da das System zur Simulation von großen Mengen komplexer Agenten eingesetzt werden soll und weitgehend linear skalieren soll, ist ein stark verteiltes System nötig. Das System

basiert daher unter anderem auf verschiedenen Komponenten (Abb. 19), die auf unterschiedlichen Computern verteilt werden können.

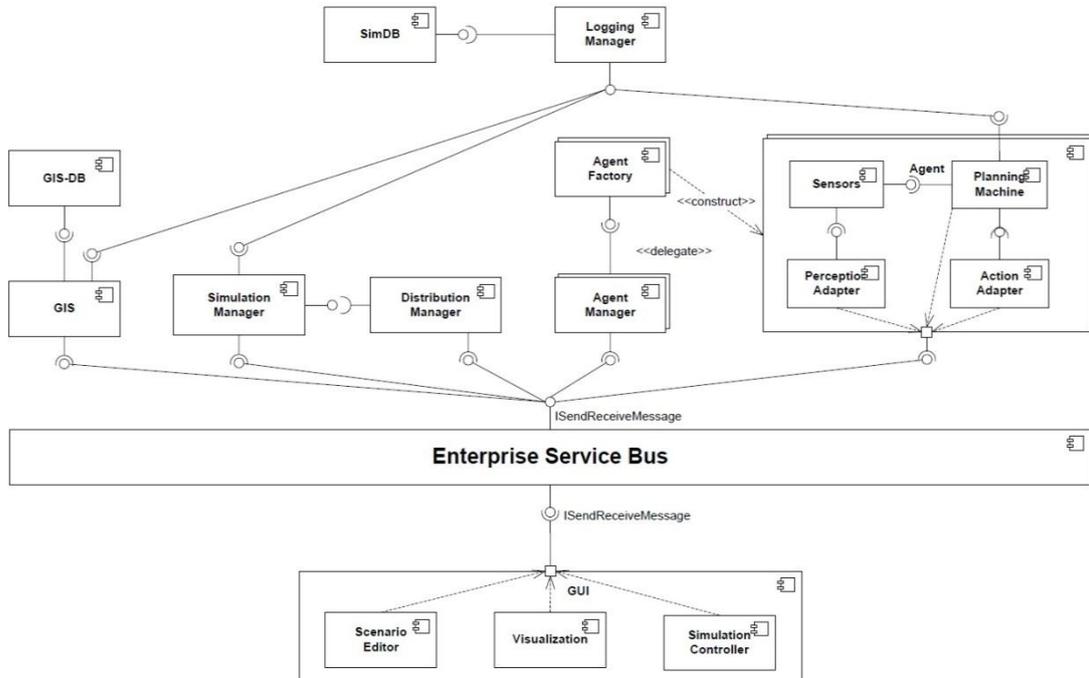


Abbildung 19: Das UML-Komponentendiagramm von WALK (Thiel-Clemen et al. 2011)

Bei dem Enterprise Service Bus handelt es sich um einen Nachrichten-Bus, der für die Kommunikation der meisten Komponenten eingesetzt wird und somit bei der Verteilung der Komponenten auf verschiedenen Computern hilft. Für die Verteilung auf den verschiedenen Computern ist der Distribution Manager zuständig. Der Simulation Manager dient der Definition von Szenarien, also deren Umgebung, der Anzahl an Agenten und deren Startzustand sowie Ereignissen, die während der Simulation auftreten sollen, zum Beispiel ein Feuer. Außerdem initiiert der Simulation Manager die einzelnen Simulationsschritte, da die Simulation schrittweise abläuft. Hierzu teilt er den Agenten mit, dass ein neuer Schritt ausgeführt werden soll und sammelt nach Ende des Simulationsschrittes, das heißt wenn alle Agenten signalisiert haben, dass sie mit ihrem Schritt fertig sind, die Informationen aller Agenten ein, die von dem GIS auf Konflikte überprüft werden. Der Scenario Editor ist die grafische Oberfläche, mit der Szenarien definiert werden können. Die Initialisierung und Kontrolle des Simulationslaufes übernimmt der Simulation Controller und die Darstellung des Laufes wird durch die Visualization übernommen. Die Simulationsläufe werden außerdem mit dem Logging Manager aufgezeichnet.

Auch die Agenten sind auf verschiedene Computer verteilt. Dazu besitzt jeder Computer, auf dem Agenten ausgeführt werden, einen Agent Manager, welcher die einzelnen Agenten verwaltet, startet, stoppt und migriert. Erstellt werden die Agenten durch die Agent Factory. Der Agent Manager ist somit von der eigentlichen Modellierung der Agenten unabhängig, so dass verschiedene Agententypen in einer einzelnen Simulation verwendet werden können. Die Verteilung auf verschiedene Computer erfolgt durch die Aufteilung des gesamten Simulationsbereiches in Sektoren. Jeder Computer des verteilten Systems verwaltet nun seine Sektoren mit den sich darin befindenden Agenten und seine Layer. Außerdem publiziert er Änderungen in den Sektoren durch Nachrichten auf dem Nachrichten-Bus. Die Kommunikation zwischen Komponenten kann direkt sein oder indirekt durch Abonnieren der Nachrichten des Sektors sein. Die Agenten selbst verfügen über Sensoren, die ihnen die Wahrnehmung von Informationen ermöglichen. Der Agent Sensor liefert den von außen sichtbaren Zustand der sichtbaren Agenten, das heißt die Emotionen und Gedanken sind für den Agenten nicht sichtbar. Der Object Sensor liefert Informationen zu allen sichtbaren Objekten und den möglichen Aktionen mit ihnen. Mit dem Temperature Sensor wird die Temperatur an der Position des Agenten wahrgenommen. Die auf den Agenten einwirkenden Kräfte können durch den Force Sensor wahrgenommen werden. Weitere Sensoren sind möglich und nötig. Die Sensoren basieren auf den Informationsschichten des GIS und liefern dem Agenten alle Informationen, die er theoretisch wahrnehmen kann. Die Wahrnehmung des Agenten kann jedoch durch Filter beschränkt werden, so dass Agenten, die zum Beispiel taub oder farbenblind sind, eine andere Wahrnehmung als die anderen Agenten besitzen. Durch den Action Adapter kann der Agent auf eine vordefinierte Menge an Aktionen zugreifen. Aktionen sind Handlungen, die den äußeren Zustand des Agenten oder eines Objektes verändern und somit einen Einfluss auf die Wahrnehmung anderer Agenten haben. Die Agenten haben die Fähigkeit eigenständig zu handeln. Die Planning Machine dient dem Treffen von Entscheidungen auf Basis der Wahrnehmung. Da diese von der Plattform unabhängig ist, lassen sich für eine Simulation verschiedene Modellierungen der Agentenintelligenz benutzen.

3.4.2 Entwicklungs- und Validierungsstand

Die Entwicklung von WALK erfolgt durch Bachelor- und Masterstudenten an der HAW Hamburg unter Leitung von Prof. Dr. Thomas Thiel-Clemen und Prof. Dr. Stefan Sarstedt. Nach Vorarbeit im Bereich der emotionalen Agenten durch die Bachelorarbeit von Arne Klingenberg (Klingenberg 2010) und die gemeinsame Veröffentlichung über die Kombination von zielorientiertem Verhalten und Emotionen in Individuen-orientierten Simulationen von Prof. Dr. Thomas Thiel-Clemen und Arne Klingenberg (Thiel-Clemen und Klingenberg 2010), wurde schließlich im Wintersemester 2010/2011 (das heißt ab 01.09.2010) das Forschungsprojekt WALK ins Leben gerufen. Es handelt sich also um ein Projekt, das sich noch in einem relativ frühen Stadium befindet.

Schwerpunkt der derzeitigen Entwicklung sind das Multiagentensystem und das Geoinformationssystem sowie das generelle Grundsystem. Die Agenten sind dazu in der Lage, sich in einer Umgebung von einem bestimmten Startpunkt zu einem Endpunkt zu bewegen. Hierfür gibt es zurzeit zwei verschiedene Arten der Agenten. Die eine Art arbeitet mit Potentialfeldern, die von Ziel, Wänden, Objekten und anderen Agenten ausgehen. In der Umsetzung dieses Verfahrens in WALK geht vom Ziel ein positives Potentialfeld aus und von den Wänden, Objekten und Agenten jeweils negative Potentialfelder. Der Agent versucht nun sein Potential zu erhöhen, also zu dem Ort mit dem höchsten Potential zu gelangen. Dadurch soll er sich zum Ziel bewegen und er weicht Hindernissen automatisch aus und vermeidet so Kollisionen. Eine Anpassung der Geschwindigkeit zur Kollisionsvermeidung ist noch nicht vorhanden. Allerdings kann dieses Verfahren dazu führen, dass der Agent in Sackgassen läuft. Aus diesem Grund gibt es noch eine alternative Art der Agenten. Diese benutzen einen A*-Algorithmus und bewegen sich entlang der berechneten Route. Für dieses Verfahren muss allerdings noch eine zusätzliche Kollisionsvermeidung mit den mobilen Hindernissen implementiert werden, so dass der Agent zur Kollisionsvermeidung seinen Bewegungsvektor und seine Geschwindigkeit anpasst. Ziel bis zum Ende des Sommersemesters 2011 (31.08.2011) ist die Ausführbarkeit der RiMEA-Testfälle 1 („Beibehalten der vorgegebenen Gehgeschwindigkeit in einem Gang“), 4 („Spezifischer Fluss durch einen Querschnitt“), 9 („Eine Menschenmenge verlässt einen großen öffentlichen Raum“) und 12 („Auswirkung von Engstellen“). Die hierzu benötigten Schnittstellen von WALK und den Komponenten wurden dafür festgelegt und die Architektur des Systems wurde bestimmt. Bei den genannten RiMEA-Testfällen wird nur für den Test 9 eine Entscheidungsfindung der Agenten benötigt, da sie dort zwischen verschiedenen Ausgängen entscheiden müssen. Hierfür sind auf Seite des Agenten jedoch bisher nur das Erkennen der Ausgänge und die Wahl des dichtesten Ausganges geplant. Zur Ausführung der anderen Tests ist auf Seite der Agenten nur eine Wegfindung und Kollisionsvermeidung nötig. Für den Test 4 ist zudem erforderlich, dass das System periodische Randbedingungen ermöglicht, also dass Agenten, die den Gang am Ende des Ganges verlassen, diesen ohne Verzögerung wieder am Anfang betreten. Der aktuelle Entwicklungsstand von WALK umfasst noch keine Agentenintelligenz und keine Emotionen, sondern stellt das Grundgerüst dar. Da sich die Agentenintelligenz unabhängig vom Rest des Systems austauschen lässt, stellt dies kein Problem dar.

Die bisherige Validierung erfolgt wie bereits erwähnt auf Basis der RiMEA, wobei Test 1, 4, 9 und 12 die zurzeit betrachteten Testfälle darstellen. Außer den RiMEA-Testfällen ist noch keine Validierung erfolgt, was unter anderem auf den frühen Entwicklungsstand von WALK zurückzuführen ist.

3.4.3 Anforderungen des Konzeptes

Damit das Validierungskonzept angewendet werden kann, müssen bestimmte Rahmenbedingungen erfüllt sein und die Gruppensimulation muss bestimmte Anforderungen erfüllen. Diese werden für die einzelnen Schritte des Konzeptes erläutert.

Um im ersten Schritt des Konzeptes die Datenvalidität zu gewährleisten, werden neben verlässlichen Methoden auch ausreichend viele geeignete und verlässliche Quellen für die Daten benötigt. Dies können unter anderem Videoaufnahmen sein, die das zu simulierende Verhalten darstellen. Werden für das Projekt keine neuen Daten mehr gesammelt, sondern bereits validierte Daten verwendet, ist dieser Schritt nicht mehr nötig.

Für die optimale Überprüfung des konzeptuellen Modells ist es erforderlich, dass es Experten für die modellierten Gebiete gibt. Diese müssen die Theorien, Annahmen und Regeln des konzeptuellen Modells auf Korrektheit und die Darstellung der Problementität auf Sinnhaftigkeit überprüfen. Dafür ist es nötig, dass das konzeptuelle Modell auf eine für die Experten verständliche Weise dargestellt wird, wobei beachtet werden muss, dass die Experten aus Forschungsbereichen kommen können, die der Informatik nicht nahe stehen. In WALK könnten dies zum Beispiel Psychologen sein, die die Zusammenhänge zwischen verschiedenen Emotionen und Verhaltensweisen betrachten könnten.

Im dritten Schritt werden die RiMEA-Testfälle ausgeführt, um das Vorhandensein und die Korrektheit von Grundfunktionen zu überprüfen. Wenn alle RiMEA-Testfälle ausgeführt werden sollen, müssen bestimmte Besonderheiten in der Evakuierungssimulation erfüllt sein. Auf Agentenebene müssen den Agenten Geschwindigkeiten, Reaktionszeiten und Rettungswege zugewiesen werden können. Die Parameter der Agenten müssen gleich verteilt um einen Mittelwert variiert werden können. Die meisten RiMEA-Testfälle erfordern nur, dass Agenten einem bestimmten Fluchtweg folgen ohne selbst Entscheidungen zu treffen. Jedoch erfordern die Tests 9 („Eine Menschenmenge verlässt einen großen öffentlichen Raum“), 11 („Wahl des Rettungsweges“) und 14 („Routenwahl“) noch die zusätzliche Fähigkeit der Agenten, Entscheidungen zu treffen. Außerdem muss die Simulation neben normalen Räumen und Gängen mit Wänden und Durchgängen auch Treppen und verschiedene Stockwerke modellieren können. Für die Startverteilung der Agenten ist oft gefordert, dass diese gleichförmig in einem Startbereich verteilt positioniert werden, dies muss die Simulation somit auch erfüllen können. Die meisten Tests erfordern außerdem die Messbarkeit der Evakuierungszeit von Agenten. Besondere Anforderungen stellt Test 4 („Spezifischer Fluss durch einen Querschnitt“). Dieser Test erfordert die Messbarkeit des spezifischen Flusses in Personen/ms und der Personendichte in Personen/m². Außerdem wird gefordert, dass die Simulation periodische Randbedingungen unterstützt, also dass Personen, die den Gang am Ende verlassen, diesen ohne Verzögerung am Anfang wieder betreten.

Für die subjektive Validierung im Betrieb werden abermals für eine optimale Validierung Experten aus den modellierten Gebieten benötigt, die die Validierung durchführen. Auf Seiten des Systems wird die Darstellbarkeit einer Vogelperspektive und einer Egoperspektive aus der Sicht eines der Agenten gefordert. Steuerbarkeit eines Agenten in der Simulation und Nutzbarkeit von Virtual-Reality-Ausrüstung für die Steuerung eines Agenten aus der Egoperspektive werden nur gefordert, wenn diese optionalen Punkte in der Validierung benutzt werden sollen. Wenn möglich sollten dann die anderen Agenten auf diesen Agenten reagieren, das heißt unter anderem ausweichen oder zum Beispiel kommunizieren. Dies würde eine Echtzeit-Ausführung der Simulation erfordern.

Die objektive Validierung im Betrieb erfordert zunächst, dass Videoaufnahmen vorhanden sind, die das Verhalten, das simuliert werden soll, darstellen. Diese müssen für die statistische oder mathematische Validierung nutzbar sein. Die statistische Validierung stellt höhere Anforderungen an die Aufnahmen als die mathematische Validierung und fordert außerdem verschiedene Aufnahmen, damit eine ausreichend große historische Datenmenge zur Verfügung stehen kann. Da bei diesem Verfahren möglichst viele Einzelwerte der Personen beziehungsweise Agenten betrachtet werden sollten, wie Positionen, Geschwindigkeiten in bestimmte Richtungen, Ausrichtung und mehr, werden für dieses Verfahren besonders gute Aufnahmen benötigt, da sich diese Werte auf den Einzelbildern der Aufnahmen erkennen lassen müssen. Für die Berechnung der Geschwindigkeit ist es des Weiteren nötig, dass die einzelnen Personen in der Aufnahme auf den verschiedenen Einzelbildern wiedererkannt werden können. Diese Werte müssen durch einen Algorithmus bestimmt werden. Die Werte, die in den Vergleich eingehen, müssen auch in der Simulation bestimmbar sein. Da eine für den Vergleich verwendete Aufnahme vermutlich nur einen Teil der Simulation darstellen kann, vor allem unter der Annahme, dass die Simulation für sehr große Menschenmassen ausgelegt ist, müssen diese Werte auch für einzelne Bereiche der Umgebung in der Simulation bestimmt werden können.

Bei der mathematischen Validierung werden geringere Anforderungen an die Videoaufnahmen gestellt, da nur die Anzahl an Personen beziehungsweise Agenten in den verschiedenen Bereichen der Umgebung verglichen wird. Andere Informationen müssen sich aus den Aufnahmen nicht ableiten lassen. Es muss also ein Algorithmus vorhanden sein, der die Anzahlen der Personen auf den Einzelbildern der Aufnahme in den unterschiedlichen Bereichen bestimmen kann und ein Algorithmus, der in regelmäßigen Abständen die gleiche Bestimmung für die unterschiedlichen Bereiche der Simulation ausführt. Es werden also abermals verschiedene Bereiche in der Simulation gefordert, für die Messungen durchgeführt werden können. Auch hier gilt, dass die Aufnahme wahrscheinlich nicht das ganze Geschehen der Simulation darstellen kann. Da allerdings die Messung in Bereichen bereits vom Verfahren gefordert wird, ändert dies nichts an den Anforderungen. Eine zusätzliche Anforderung dieses Verfahrens ist, dass Verteilungen bei zu großen Abweichungen zwischen Aufnahme und Simulation in der Simulation replizierbar sein müssen, das heißt, dass

die Simulation mit der Verteilung aus der Aufnahme weiterläuft. Auch müssen die Zeitpunkte in der Aufnahme und Simulation bestimmbar sein, die den gleichen Zeitpunkt im realen Szenario darstellen.

3.4.4 Anwendbarkeit des Konzeptes

In diesem Teil wird geklärt, welche der genannten Rahmenbedingungen und Anforderungen an die Simulation im derzeitigen Entwicklungsstand von WALK bereits erfüllt sind. Für nicht erfüllte Rahmenbedingungen und Anforderungen wird geklärt wie sie sich erfüllen lassen, damit eine vollständige Anwendung des Konzeptes in Zukunft möglich ist. Dazu wird gezeigt, was für Funktionen in WALK realisiert werden müssen und wo dies unter Beachtung der derzeitigen Architektur am sinnvollsten wäre.

Die Anforderung zur Gewährleistung der Datenvalidität ist, dass ausreichend viele geeignete und verlässliche Daten zur Verfügung stehen. Da das WALK-Projekt mit der RiMEA arbeitet, die unter anderem viele Daten vorgibt, die zuvor wissenschaftlich ermittelt wurden, ist dies für das Grundsystem zur Simulation von Evakuierungen ohne intelligente Agenten gegeben. Jedoch sollen emotionale Agenten einen wichtigen Bestandteil von WALK darstellen. Deshalb sind auch zu diesem Gebiet Daten zu sammeln. Die Grundlage ist durch Forschung von Arne Klingenberg und Prof. Dr. Thomas Thiel-Clemen bereits gelegt. Weitere Daten werden bei Beginn der Modellierung der emotionalen Agenten benötigt werden.

Für die Validierung des konzeptuellen Modells werden bei WALK Experten für die Bereiche Evakuierung und Psychologie benötigt. In beiden Bereichen sollte es möglich sein, diese zu finden. Da noch kein konkretes konzeptuelles Modell entworfen wurde, ist auch die Anforderung, dieses auf für die Experten verständliche Weise darzustellen kein Problem, da die Gruppen der Leute, die es verstehen sollen bereits bekannt ist.

Die Anforderung der RiMEA, dass Agenten Geschwindigkeiten, Reaktionszeiten und Rettungswege zugewiesen werden können, ist im derzeitigen Entwicklungsstand noch nicht komplett erfüllt. Agenten besitzen in WALK eine AgentDefinition, diese enthält eine Liste von AgentParameters. Diese wiederum haben einen ParameterType, hier Double, Integer, String oder Vector3D, einen Namen und einen Wert. Die Geschwindigkeit geht als AgentParameter mit dem Namen „movingSpeed“ in die Berechnung der Fortbewegung in der move-Methode in der Agenten-Klasse mit ein. Für die Reaktionszeit ließe sich ein AgentParameter mit dem Namen „reactionTime“ erstellen. Bei der Reaktionszeit müsste nur dafür gesorgt werden, dass der Agent bis zum Ablauf dieser Zeit keine Bewegungen in der move-Methode durchführt. Die Zuweisung von Rettungswegen in RiMEA beschreibt eigentlich nur die Zuweisung von bestimmten Ausgängen, im Fall des Tests gibt es einen Haupt- und einen Nebenausgang. Somit müssen den Agenten nur unterschiedliche Werte für den bereits vor-

handenen AgentParameter „targetPosition“ mitgegeben werden, um diese Anforderung zu erfüllen.

Die Anforderung, dass die Parameter den Agenten auf bestimmte Weisen verteilt zugewiesen werden können sollen, ist durch die Klasse AgentParameterGenerator bereits gegeben, die für das Erstellen von AgentParameters benutzt werden kann. In dieser lassen sich Integer-, Double- und Vector3D-Werte zwischen einem minimalen und maximalen Wert unter Angabe der Art der Verteilung generieren. Hierdurch lassen sich zum Beispiel die Geschwindigkeit, Reaktionszeit und Startposition eines Agenten festlegen. Für die Startposition ließen sich der minimale und maximale Wert als Vector3D zum Beispiel als der obere linke Punkt und der untere rechte Punkt definieren, die die rechteckige Fläche festlegen, in der die Agenten auf eine bestimmte Weise verteilt werden sollen.

Für die Ausführung der RiMEA-Testfälle 9, 11 und 14, die von den Agenten die Wahl von Ausgängen und Routen fordern, sind noch kleine Erweiterungen für die WALK-Agenten nötig, da die Agenten noch keine Entscheidungen treffen können, da sie bisher noch keine Intelligenz besitzen. Um die Testfälle jedoch trotzdem zumindest ausführbar zu machen, was die derzeitige Forderung ist, kann die Intelligenz der Agenten auf der simplen Annahme beruhen, dass einfach der dichteste Ausgang gewählt wird. Hierzu muss der Agent die Positionen der Türen kennen. Diese könnte er von vornherein bereits kennen, weil er mit dem Bauplan des Raumes vertraut ist, oder es könnte ein neuer DoorSensor oder ein InteractableObjectSensor für Agenten definiert werden, der wie die anderen Sensoren, bisher AgentSensor und ObstacleSensor, funktioniert, allerdings der Wahrnehmung von Türen, beziehungsweise allen Objekten, mit denen interagiert werden kann, dient. Hierfür müssten auch die Türen als wahrnehmbare Objekte (SimulationObject) modelliert werden, entweder mit einem neuen SimulationObjectType „Door“ oder dem vorhandenen „InteractableObject“ SimulationObjectType. Zurzeit sind Türen nur simple Unterbrechungen der Wand. Da die Agenten später unter anderem mit Türen interagieren können sollen, wird eine Modellierung der Türen als SimulationObject aber sowieso nötig sein. Nach Wahl der dichtesten Tür müsste nur der AgentParameter „targetPosition“ auf diese gesetzt werden. Die Ausführbarkeit des Testfalls 14 stellt auf Seiten der Entscheidungsfindung keine Probleme dar, da die Agenten nach bisherigem Stand versuchen, den kürzesten Weg zum Ziel zu nehmen. Da der Test nur das Modellverhalten dokumentieren soll, ist dies auch zulässig. Allerdings fordern die Testfälle 2, 3, 8, 13 und 14, dass die Simulation Treppen unterstützt. Dies ist im derzeitigen Stand von WALK noch nicht gegeben. Eine Möglichkeit zur Umsetzung dieser Anforderung wäre, sie ebenfalls als „InteractableObject“ zu modellieren. Der Agent müsste mit der Treppe interagieren, um sie zu betreten und würde dabei seine Gehgeschwindigkeit halbieren, wie es von der RiMEA gefordert ist. Wenn der Gebäudegrundriss nur zweidimensional angegeben wird, das heißt andere Stockwerke werden zum Beispiel in einem anderen Bereich der Umgebung dargestellt, so sollte die Position des Agenten bei Verlassen der Treppen so verändert werden, dass er sich in dem Bereich der Umgebung

befindet, der ein höheres oder tieferes Stockwerk darstellt. Hierzu könnte diese Position dem Agenten beim Interagieren mit der Treppe mitgeteilt werden und er könnte beim Verlassen der Treppe seine eigene Position anpassen. Die von der RiMEA geforderten periodischen Randbedingungen ließen sich auf eine ähnliche Weise umsetzen. So könnte sich am Ende des Ganges ein „InteractableObject“ über die ganze Breite des Ganges befinden, bei dessen Interaktion der Agent seine Position so verändert, wie es ihm das „InteractableObject“ mitteilt. Seine neue Position wäre eine Position am Anfang des Ganges abhängig von der Position am Ende des Ganges. Mit beiden genannten Objekten müsste der Agent sofort interagieren, wenn er sie erreicht, so dass keine Verzögerungen entstehen.

Des Weiteren fordern die ersten drei RiMEA-Testfälle die Messbarkeit der Evakuierungszeit der einzelnen Agenten, was im derzeitigen Stand von WALK noch nicht gegeben ist, aber sehr einfach umzusetzen ist. Die Bewegung der Agenten erfolgt in diskreten Zeitschritten. Sie wissen daher für welche Zeitdauer im Szenario die nächste berechnete Bewegung erfolgt. Somit kann der Agent die Zeitdauern der berechneten Bewegungen solange aufaddieren bis er die Zielposition erreicht hat. Dadurch weiß er seine Evakuierungszeit.

Die Messbarkeit des spezifischen Flusses in Personen/ms und der Personendichte in Personen/m² ist für den Testfall 4 erforderlich. Zum Bestimmen des spezifischen Flusses muss es möglich sein, die Anzahl an Personen, die in einer bestimmten Zeit eine Linie mit bekannter Breite überschreiten, zu messen. Hierzu lassen sich die Sektoren von WALK verwenden. So kann der Gang aus zwei Sektoren bestehen und der Fluss kann bestimmt werden, indem für eine Sekunde gemessen wird, wie viele Agenten sich vom ersten in den zweiten Sektor bewegt haben. Da die Breite der Sektoren ebenfalls bekannt ist, lässt sich der spezifische Fluss berechnen. Für die gemittelte Personendichte muss nur die Anzahl aller Agenten in der Simulation durch die von den Agenten betretbare Fläche dividiert werden. Die Messung und Berechnung könnte im GIS erfolgen, da dieses Zugriff auf alle Sektoren und somit auch auf die Layer mit den Agenten hat (Abb. 20). Außerdem sind diese Werte auch außerhalb der Validierung von Nutzen.

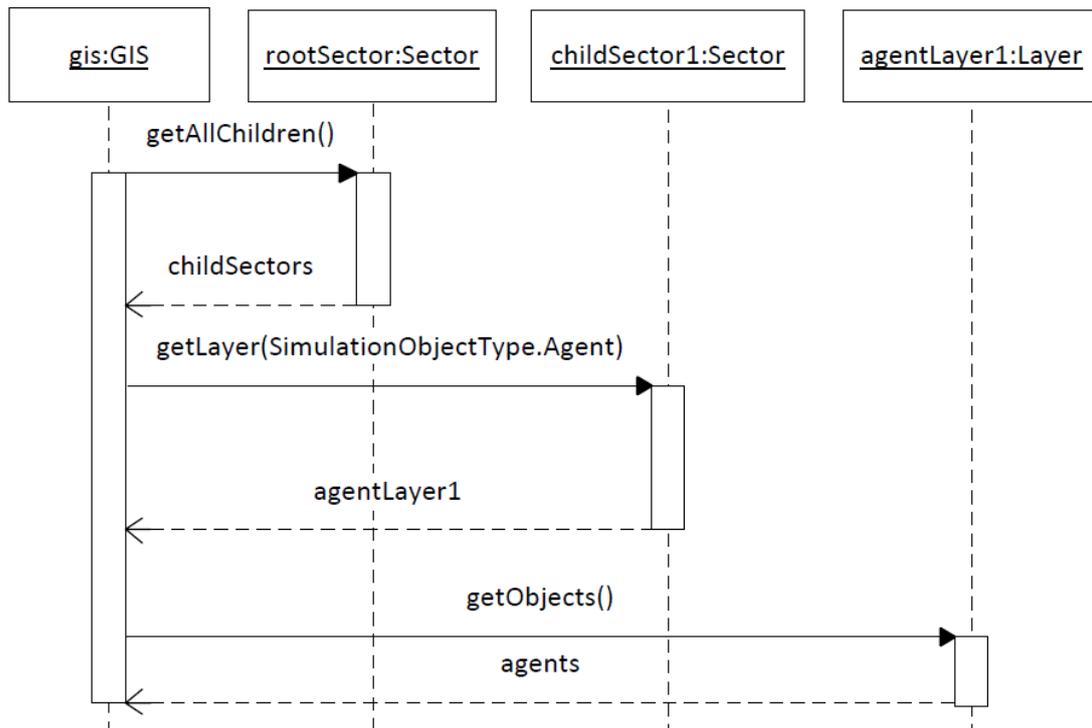


Abbildung 20: UML-Sequenzdiagramm zum Zugriff vom GIS auf die Agenten

Die Forderung einer Vogelperspektive für die subjektive Validierung im Betrieb ist in WALK bereits erfüllt. Eine Egoperspektive aus der Sicht eines Agenten ist noch nicht vorhanden. Es lassen sich jedoch bereits verschiedene Kamerapositionen wählen, die in der Klasse `JMEVisualization` definiert sind. Diese Klasse stellt das Grundgerüst der 3D-Darstellung der Simulation und Steuerung der Kamera durch die `jMonkeyEngine 3.0`, einer Java Spiele- und OpenGL-Engine, dar. Hier müsste nur eine neue Kameraposition definiert werden, die sich immer auf Höhe des „Kopfes“ des Agenten und an der Position des Agenten befindet und in die aktuelle Blickrichtung beziehungsweise Laufrichtung des Agenten zeigt. Die optionale Forderung der Steuerbarkeit eines Agenten stellt in WALK ein Problem dar, da die Berechnung der Simulation nicht in Echtzeit geschehen soll. Begründet ist dies darin, dass die Berechnung der Simulation mit einer großen Menge von Agenten mit komplexer Intelligenz sehr viel Rechenzeit benötigt. Aus diesem Grund soll die Simulation berechnet und aufgezeichnet werden und später mit Visualisierung abgespielt werden. Hierbei ist dann allerdings keine Steuerung eines Agenten mehr möglich, da die anderen Agenten ihn nicht wahrnehmen würden, da sie sich nur auf die vorher berechnete Weise verhalten. Es wäre allerdings trotzdem möglich, eine Kameraposition zu implementieren, die den gleichen Gesetzen wie die Agenten unterliegt, das heißt sie würde sich immer auf Höhe des „Agentenkopfes“ befinden und nicht durch Wände und andere Hindernisse bewegt werden können. Diese Kamerasteuerung inklusive Kollisionserkennung würde in der `JMEVisualization`

Klasse kein Problem darstellen, da alle Hindernisse in ihr gezeichnet werden und somit bekannt sind. Außerdem unterstützt die jMonkeyEngine 3.0 aufgrund der Eigenschaft, vor allem für Spieleprogrammierung entworfen worden zu sein, bereits Funktionen wie Kollisionserkennung. Diese könnte durch einfache Bounding Boxes um die gezeichneten Objekte erreicht werden. Für die Agentensteuerung unter Nutzung von Virtual Reality gilt das gleiche bereits beschriebene Problem, dass die Simulation nicht in Echtzeit berechnet wird. Es könnte also nur eine Kamerasteuerung durch Virtual-Reality-Ausrüstung erfolgen. Hierfür müsste eine Unterstützung von Sensoren zum Erkennen der Blickrichtung und Erkennen der Laufbewegung implementiert werden. Ein weiterer wichtiger Punkt für die Face Validation ist, dass es möglich sein muss, die aufgezeichnete Simulation beim Betrachten anhalten zu können, sowie die Möglichkeit des „Vor-“ und „Zurückspulens“ zu haben, damit bestimmte Stellen mehrmals und gegebenenfalls aus verschiedenen Perspektiven betrachtet werden können. Im Moment erfolgt die Berechnung und Darstellung der Simulation noch in Echtzeit. In Zukunft soll jedoch eine Aufzeichnung der Simulation über einen Logging Manager erfolgen. Für die Wiedergabe müsste somit eine Art Simulation Player geschaffen werden, der die Logs einliest und dazu eine Visualisierung erstellt. Dazu müssten in den Log-Dateien die Beschreibung des Szenarios zum Start sowie die sich jeweils ändernden Werte wie zum Beispiel Agentenpositionen gespeichert werden. Die Visualisierung könnte wie die derzeitige Visualisierung mit der jMonkeyEngine erfolgen. Diese würde die im Log beschriebenen Strukturen darstellen und in regelmäßigen Abständen jeweils den nächsten Teil des Logs einlesen und darstellen. „Vor-“ und „Zurückspulen“ würde kein Problem darstellen, da einfach nur zu früheren oder späteren Stellen im Log gesprungen werden müsste. Beim Pausieren müsste nur der Algorithmus unterbrochen werden.

Für die objektive Validierung im Betrieb müssen noch verwendbare Videoaufnahmen beschafft werden. Diese könnten zum Beispiel von Überwachungskameras kommen. Da die Verfahren eine feste Kameraposition erfordern, wären Aufnahmen wie diese besonders geeignet. Aufnahmen, die zum Beispiel von Personen, die sich in der Situation befunden haben, gemacht wurden, lassen sich für die automatischen Vergleiche nicht verwenden, da sie nicht immer die Erkennung und Bestimmung der Ausrichtung und Bewegungen der gleichen Personen erlauben. Dies liegt an der Position, aus der die Aufnahmen gemacht wurden, und der Tatsache, dass die Aufnahmen keine festen Perspektiven besitzen, sondern die Kameras vermutlich frei bewegt werden. Für die objektive Validierung sollte deshalb versucht werden Aufnahmen von Überwachungskameras zu bekommen oder andere Aufnahmen, die von einer festen Position aufgenommen wurden und immer einen Blick auf den gleichen Bereich erlauben. Ein weiterer noch nicht erfüllter Punkt ist ein Algorithmus beziehungsweise eine Software, die genügend Werte aus den Einzelbildern der Videoaufnahme berechnen kann, so dass diese für den Vergleich mit der Simulation verwendet werden können. Im einfachsten Fall, der mathematischen Validierung, müssten nur die Personen in einzelnen Bereichen des Bildes erkannt und gezählt werden. Eine Wiedererkennung in dem nächsten Einzelbild ist nicht erforderlich. Die statistische Validierung dagegen for-

dert die Wiedererkennung der Personen, so dass unter anderem Positionsveränderungen zur Berechnung der Geschwindigkeit erkannt werden können. Zusätzlich sollte die Software dazu in der Lage sein, möglichst viele andere für den Vergleich sinnvolle Werte wie die Ausrichtung der Personen und die Dichte um Personen zu bestimmen. Die Ergebnisse müssen verlässlich sein, so dass sie auch für die Validierung geeignet sind. Die Umsetzung solch einer Software ist nicht trivial. Es sollte daher in Betracht gezogen werden, eine fertige Software hierfür zu verwenden, da ansonsten abseits der eigentlichen Entwicklung der Gruppensimulation ein weiteres komplexes Projekt zu entwickeln ist. Für die Berechnung dieser Werte aus den Bildern der Simulation ist keine zusätzliche Software nötig, da die Berechnung intern stattfinden kann. Die Simulation muss hierzu in der Lage sein, die geforderten Werte für den betrachteten Bereich zu berechnen. Da in WALK bereits das Konzept der Sektoren vorhanden ist, könnte dieser Bereich durch einen oder mehrere Sektoren beschrieben werden. Die Bestimmung der durchschnittlichen X- und Y-Position, durchschnittlichen Geschwindigkeit in X- und Y-Richtung und durchschnittlichen Ausrichtung stellt kein Problem dar, da jeder Agent diese Werte bereits besitzt und nur noch ein Durchschnitt für den oder die Sektoren gebildet werden muss. Die Berechnung könnte wieder im GIS stattfinden, da auch diese Werte außerhalb der Validierung sinnvoll sind. Der Vergleich der Werte der Aufnahmen und Simulation muss und kann nicht in Echtzeit erfolgen. Die Werte könnten daher jeweils in einer Datei, zum Beispiel im leicht einlesbaren XML-Format, gespeichert werden und mit einem weiteren Algorithmus verglichen werden. Dieser müsste zuerst die Dateien einlesen. Um die Komplexität der folgenden Rechnung zu verringern, sollte eine Hauptkomponentenanalyse durchgeführt werden. Danach muss er die simpliziale Tiefe berechnen und anhand eines Grenzwertes bestimmen, ob der gerade betrachtete Simulationslauf zu der historischen Datenmenge passt oder nicht und zum Schluss das Verhältnis von Läufen der Simulation, die über dem Grenzwert lagen zu der Gesamtanzahl an Läufen für die für die Validierung ausgeführt wurden, in Prozent angeben (Abb. 21). Diese Aufgaben sind vom Rest des Systems unabhängig, daher kann für sie eine neue Klasse geschrieben werden, die nur die Methoden zur Erfüllung der Aufgaben besitzt und keine Abhängigkeit vom System hat.

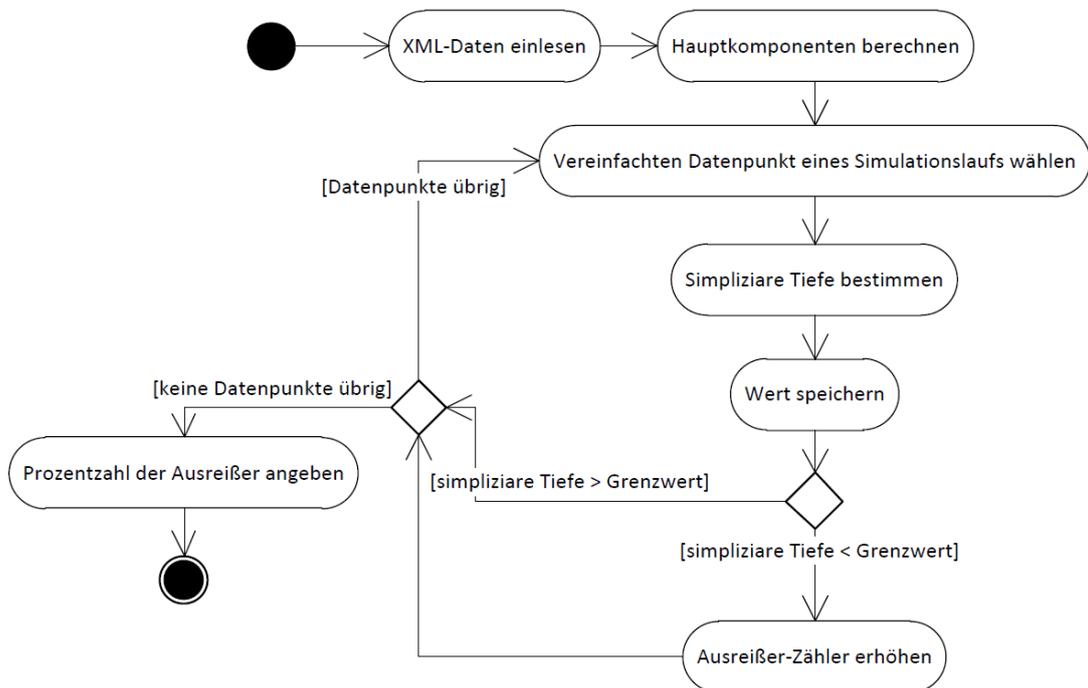


Abbildung 21: UML-Aktivitätsdiagramm zum Ablauf der Berechnungen bei der statistischen Validierung

Die mathematische Validierung stellt zwar geringere Anforderungen an die Qualität der Videoaufnahmen, jedoch stellt sie durch die Forderung der Replizierbarkeit der Verteilungen aus der Aufnahme in der Simulation bei zu großen Abweichungen höhere Anforderungen an WALK. Die Verteilungen für die Aufnahme können bereits im Voraus außerhalb des Systems berechnet werden. Auch hier könnten sie in einem leicht einlesbaren Format wie XML gespeichert werden. Für die Berechnung der Verteilungen der Aufnahme ist eine Software nötig, die die Personen in der Aufnahme erkennen und somit die Personen in verschiedenen Bereichen des Bildes zählen kann. Der Vergleich mit den Verteilungen der Simulation muss jedoch während der Berechnung der Simulation erfolgen. Hierzu wäre es möglich, das System nur so lange rechnen zu lassen, bis der zu vergleichende Zeitpunkt in der Simulation erreicht ist. Danach wird die Simulation gestoppt und die Distanz der Verteilungen wird mit Hilfe der zuvor beschriebenen Gleichung berechnet und mit einem vorher definierten Grenzwert, der die maximal zulässige Distanz angibt, verglichen. Wird der Grenzwert nicht überschritten, so kann die Simulation einfach weiterlaufen und bis zum nächsten zu vergleichenden Zeitpunkt weiterrechnen. Bei zu großer Abweichung muss die Verteilung der Aufnahme in der Simulation repliziert werden und danach die Simulation weiterlaufen (Abb. 22).

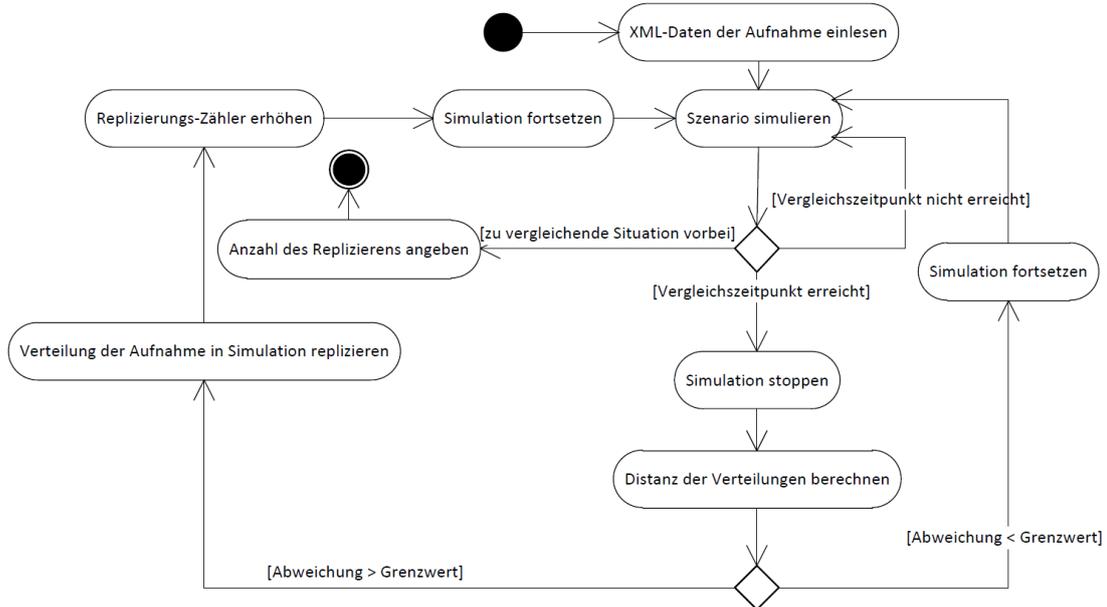


Abbildung 22: UML-Aktivitätsdiagramm zum Ablauf der mathematischen Validierung

Zur Replizierung der Verteilung der Personen aus der Aufnahme in der Simulation könnten einfach die Positionen der Agenten so verändert werden, dass die Verteilungen wieder übereinstimmen. Allerdings ist die Position der Agenten zurzeit nur durch sie selbst änderbar. Eine Alternative hierzu wäre, die aktuelle Simulation zu beenden und eine neue Simulation mit den richtigen Verteilungen zu starten. In der AgentFactory würden dann über `createAgent` Agenten mit einem veränderten AgentParameter „startPosition“ erstellt werden. Für die Berechnung der Verteilungen muss die Simulation verschiedene Bereiche besitzen, in denen die Anzahl der Agenten gemessen werden kann. Hierfür wäre abermals die Verwendung der vorhandenen Sektoren möglich. Sowohl die Aufnahme als auch die Simulation könnten jeweils auf die gleiche Weise in ausreichend viele Bereiche aufgeteilt werden, für die die Messungen stattfinden. Die Anzahl dieser Bereiche ist von der geforderten Genauigkeit der Übereinstimmung abhängig. Mehr und damit kleinere Bereiche, für die gemessen wird, erlauben höhere Genauigkeit. Da der Vergleich der Werte außerhalb der Validierung nicht sinnvoll ist, sollte hierfür eine neue Klasse, die nur für die Validierung verwendet wird, geschrieben werden. Diese könnte über das `CommunicationSystem` auf das `GIS` und den `SimulationManager` und ihre öffentlichen Methoden zugreifen. Das `GIS` und der `SimulationManager` könnten zuerst über die Methode `pauseEnvironment` die Simulation unterbrechen und dann die nötigen Informationen der Sektoren liefern. Die Validierungs-Klasse würde dann die Verteilung der Simulation berechnen. Die Methoden `pauseEnvironment` und `resumeEnvironment` sind sowohl für das `GIS` als auch für den `SimulationManager` geplant, aber noch nicht umgesetzt. Diese müssten dafür sorgen, dass die dynamischen Ereignisse und die Agenten gestoppt und wieder gestartet werden. Dazu bräuchten

unter anderem die Agenten noch eine einfache boolesche Variable, die angibt ob sie aktiv sind oder nicht. Für den Vergleich der Verteilungen in der Validierungs-Klasse müsste nur die Distanz berechnet werden und mit dem Grenzwert verglichen werden. Ist die Distanz geringer als der Grenzwert, so kann die Simulation über die Methode `resumeEnvironment` des GIS und SimulationManager weiterlaufen gelassen werden (Abb. 23).

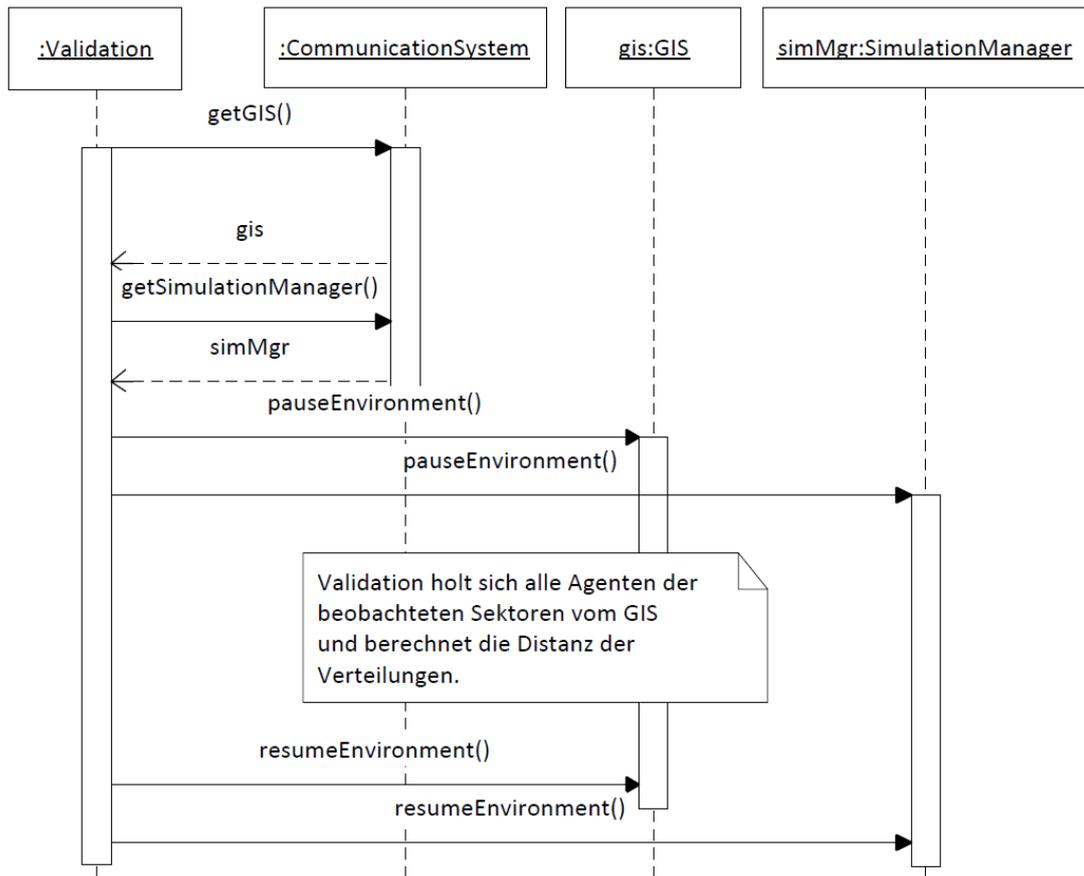


Abbildung 23: UML-Sequenzdiagramm zur Kommunikation zwischen Validierungs-Klasse, CommunicationSystem, GIS und SimulationManager

Bei Überschreitung des Grenzwertes würde die erwähnte Anpassung der Agentenpositionen erfolgen und danach die Simulation mit `resumeEnvironment` weiterlaufen gelassen werden oder die alte Simulation beendet werden und eine neue Simulation mit der Verteilung aus der Aufnahme gestartet werden. Eine weitere Anforderung ist, dass für die Vergleiche die gleichen Zeitpunkte des Szenarios in Aufnahme und Simulation bestimmt werden müssen. In der Simulation muss die reale Zeit, die im Szenario vergangen ist, bereits bekannt sein, damit unter anderem Evakuierungszeiten und die aktuell vergangene Zeit im Szenario bestimmt werden können. Die reale Zeit im Szenario sollte in Zukunft der Simula-

tionExecutionThread des SimulationManagers liefern, was im Moment noch nicht der Fall ist. Läuft die Aufnahme mit realer Geschwindigkeit ab oder gibt es eine Möglichkeit die Geschwindigkeit der Zeit zum Beispiel durch eine bei Überwachungskameras übliche Zeitanzeige zu bestimmen, so gibt es keine Probleme, die entsprechende Simulationszeit zu einer Aufnahmezeit zu erhalten. Wird die Simulation nun zum Beispiel mit der gleichen Situation wie zu Anfang der Aufnahme gestartet, ist es möglich die gleichen Szenariozeitpunkte in Simulation und Aufnahme zu bestimmen. Ansonsten sollte das erste Verfahren zum Bestimmen dieser Zeitpunkte aus dem Teil zur mathematischen Validierung angewendet werden. Für dieses Verfahren wird das Wissen über die Anzahl aller Agenten in der Simulation benötigt beziehungsweise die Anzahl der Agenten in den beobachteten Sektoren. Diese Information steht dem GIS zur Verfügung. Dieser Wert muss in gleichen Zeitabständen in der Simulation abgerufen und gespeichert werden können. Dies könnte unter anderem im SimulationExecutionThread in der run-Methode erfolgen. Durch das CommunicationSystem hat auch dieser Zugriff auf das GIS.

In diesem Unterkapitel wurde gezeigt, dass im derzeitigen Entwicklungsstand von WALK noch nicht das komplette Validierungskonzept anwendbar ist. Hierfür fehlen einerseits noch ein konzeptuelles Modell, das die Thesen, Annahmen und Regeln der Simulation beschreibt und validiert werden soll und andererseits noch einige Funktionen in der Implementierung des Systems, die für die Anwendbarkeit des Konzeptes erforderlich sind. Für diese Funktionen wurde gezeigt, wie sie umgesetzt werden können. Der wichtigste Punkt ist allerdings, dass eine ausführliche Validierung im Betrieb ohne das Vorhandensein der geplanten Agentenintelligenz noch nicht wirklich sinnvoll ist, da das Agentenverhalten noch nicht dem geplanten Agentenverhalten entspricht. Hingegen ist die Überprüfung der Grundfunktionalität durch die RiMEA-Testfälle, die zu einem großen Teil auch einer Verifizierung entspricht, bei dem bisherigen Vorgehen in der Entwicklung sinnvoll. In der fortschreitenden Entwicklung von WALK werden nach und nach alle RiMEA-Testfälle durchführbar und erfüllbar sein. Nach Erstellung des konzeptuellen Modells kann mit der Validierung dieses begonnen werden. War diese erfolgreich, so ist es an der Zeit das Modell umzusetzen und das System im Betrieb zu validieren.

3.4.5 Anwendung des Konzeptes auf WALK

In diesem Teil wird unter Berücksichtigung des bisherigen Vorgehens bei der Entwicklung und Validierung von WALK betrachtet, wie das Validierungskonzept am besten anzuwenden ist. Es wird erläutert, in welcher Reihenfolge die einzelnen Schritte auszuführen sind, wie deren Ausführung in WALK konkret aussehen soll, welche Teile der Entwicklung zuvor abgeschlossen sein sollen und wie die Ergebnisse der Validierung bewertet werden sollen.

Bei der Entwicklung von WALK wurde nicht mit der Erstellung eines konzeptuellen Modells begonnen, das die Theorien, Annahmen und Regeln beschreibt, sondern mit der Konzepti-

on einer flexiblen und verteilten Architektur und deren Umsetzung. Die Grundfunktionen einer Evakuierungssimulation werden zuerst umgesetzt, wobei die RiMEA und ihre Testfälle einen guten Überblick über den Fortschritt und die geforderten Funktionen geben. Durch die Flexibilität, die es erlaubt die Agentenintelligenz unabhängig vom Rest des Systems zu modellieren, ist dies in diesem Fall auch sinnvoll. So kann zuerst ein in den Grundfunktionen valides System fertiggestellt werden, das später mit unterschiedlichen Agenten laufen kann. Die Validierung und Verifizierung dieses Systems erfolgt durch die RiMEA-Testfälle. Die Durchführung läuft entsprechend der Vorgaben der Richtlinie ab. Ein valides Grundsystem ist bei Bestehen aller Testfälle vorhanden.

Durch die beschriebene Flexibilität ist es möglich, das Sammeln von Daten für das konzeptuelle Modell der Agentenintelligenz und die darauf folgende Erstellung dieses Modells unabhängig von der Entwicklung des Grundsystems durchzuführen. Die Datenvalidität ist beim Sammeln von Daten wie im Konzept beschrieben sicherzustellen. Verlässliche Methoden zum Sammeln und Pflegen der Daten sind daher nötig. Die gesammelten Daten sollten zudem auf Ausreißer untersucht werden. Werden Ausreißer gefunden, dann müssen diese auf Korrektheit überprüft werden. Das konzeptuelle Modell sollte nur auf Basis von Daten erstellt werden, die als valide und somit verlässlich angesehen werden.

Wurde das konzeptuelle Modell erstellt, so hat die konzeptuelle Validierung zu erfolgen. Diese sollte für WALK mit Hilfe von Experten aus dem Bereich der Psychologie und Evakuierung durchgeführt werden. Diese haben einerseits tieferes Wissen aus den für WALK wichtigen Fachbereichen und sind andererseits vom Entwicklungsteam unabhängig und erlauben damit eine objektivere Bewertung des konzeptuellen Modells. Die Experten müssen die Thesen, Annahmen und Regeln, welche das Verhalten der Agenten spezifizieren, überprüfen. Vor allem der Einfluss von psychologischen Faktoren und Emotionen auf das Verhalten der Agenten ist für WALK zu betrachten. Diese Validierung erfolgt durch Face Validation, indem die Experten das konzeptuelle Modell betrachten. Dieses muss daher für die Experten verständlich dargestellt werden. Werden während der Validierung Fehler gefunden, so sind diese zu beheben. Erst mit einem validierten konzeptuellen Modell sollte mit der Implementierung begonnen werden. Ansonsten könnten zu spät gefundene Fehler dafür sorgen, dass ein unbrauchbares implementiertes Modell erstellt wurde, was einen hohen unnützen Arbeitsaufwand bedeutet.

Wurde das konzeptuelle Modell umgesetzt, so sollten erneut die RiMEA-Testfälle ausgeführt werden. Dies dient dazu, zu überprüfen, ob auch nach der Implementierung der individuellen Agentenintelligenz das erwartete Verhalten auf Makroebene eintritt. Auch lassen sich so Fehler in der Implementierung entdecken. Da die RiMEA-Testfälle bereits vorher umgesetzt wurden, stellt dies keinen großen Mehraufwand dar. Zusätzlich sind auch Degenerierungstests einzusetzen. Eine Möglichkeit hierfür wäre, sich eine große Menge Agenten für eine sehr lange Zeit durch eine große und komplexe Umgebung bewegen zu lassen. Zu

beobachten wäre unter anderem, ob die Agenten in der Umgebung stecken bleiben, in einer Schleife hängen und immer die gleichen Aktionen ausführen und somit nicht weiterkommen oder anderes unerwartetes Verhalten zeigen. Werden Fehler gefunden, so ist die Ursache zu beheben, bevor mit der Validierung fortgefahren wird.

Bevor die aufwändigen Verfahren zur Validierung im Betrieb angewendet werden, sollte zuerst die interne Validität überprüft werden. Dazu werden vielen Simulationsläufe mit dem gleichen Szenario und den gleichen Agenten gestartet. Nun wird überprüft wie weit sich die verschiedenen Ergebnisse der Simulationsläufe voneinander unterscheiden. Hierfür ist es sinnvoll sich die minimale, mittlere und maximale Evakuierungszeit und die statistische Verteilung der Evakuierungszeiten anzusehen. Da in WALK auch Gefahrenquellen wie Feuer, Rauch und Wasser simuliert werden sollen, ist auch eine Betrachtung der Anzahl der Verletzten und Toten sinnvoll. Weichen diese in den verschiedenen Simulationsläufen stark voneinander ab, so ist das Modell nicht für Vorhersagen geeignet. Vor allem für Evakuierungssimulationen ist dies aber nötig, da mit ihnen Vorhersagen zum Ablauf von Evakuierungen getroffen werden sollen. Auch für WALK hat dieser Test eine besondere Wichtigkeit, da gezeigt werden muss, dass auch mit Einbringung von Emotionen in eine Evakuierungssimulation noch verlässliche und nutzbare Ergebnisse erzielt werden.

Auch der Sensitivitätsanalyse kommt in WALK eine besondere Bedeutung zu. Durch die komplexe Modellierung der Agenten und das Erweitern der Agenten um Emotionen in WALK erfordert die Berechnung jedes Simulationsschrittes einen großen Rechenaufwand. Es ist daher sinnvoll zu überprüfen, welche Parameter einen messbaren Einfluss auf den Ablauf der Simulation haben. Vor allem ob die Emotionen einen Einfluss auf den Ausgang der Simulation haben, ist zu zeigen. Dies wurde bereits in einer früheren Arbeit untersucht (Thiel-Clemen und Klingenberg 2010). Da es sich bei WALK jedoch um ein komplexeres Projekt handelt, sollte dies erneut durchgeführt werden. Es muss für WALK bewiesen werden, dass die Emotionen tatsächlich einen merkbaren und damit nicht zu vernachlässigenden Einfluss auf das Ergebnis einer Evakuierungssimulation haben. Die Überprüfung erfolgt durch Anpassen der verschiedenen Parameter und Vergleichen der Ergebnisse der Simulation. Die Überprüfungen sollten für möglichst viele realistische Werte der Parameter erfolgen. Da es nicht möglich sein wird alle verschiedenen Werte der Parameter zu testen, sollten die Parameter so gewählt werden, dass die wichtigsten Arten von Agenten abgedeckt werden. In Bezug auf die Emotionen wäre es zum Beispiel sinnvoll, starke Ausprägungen verschiedener Emotionen zu wählen, um zu zeigen, dass diese unterschiedliche Einflüsse haben. Für jeden Parameterwert und jede Kombination von Parameterwerten müssen mehrere Durchläufe durchgeführt werden, um Unterschiede, die nur durch stochastische Eigenschaften entstanden sind, nicht als Resultat der Parameteränderung zu interpretieren. Wird für einen Parameter festgestellt, dass er keinen Einfluss auf die Simulation hat, ist er zu entfernen. Dadurch wird das Modell minimiert, so dass die Komplexität sinkt. Die zu betrachtenden Werte der Simulation sollten auch hier die minimale, mittlere und maximale

Evakuierungszeit, die Verteilung der Evakuierungszeiten sowie die Anzahl der Verletzten und Toten sein. Eine genaue Betrachtung des Verhaltens der einzelnen Agenten sollte erst später geschehen, da in diesem Schritt nur die Sinnhaftigkeit der einzelnen Parameter für die Nutzung des Systems für die Simulation von Evakuierungen interessiert. Hierfür ist es prinzipiell egal was in der Simulation passiert, so lange verlässliche und realistische Vorhersagen getroffen werden können.

Die eigentliche Validierung im Betrieb wird erneut mit Hilfe von Experten aus den Bereichen Psychologie und Evakuierung durchgeführt. Damit diese eine Face Validation durchführen können, sollten ihnen verschiedene Szenarien gezeigt werden, in denen unterschiedliches Verhalten der Agenten zu beobachten ist. Wegen der Erweiterung der Evakuierungssimulation um Emotionen sollten vor allem auch Agenten mit unterschiedlichen Emotionen in den Szenarien vorhanden sein, so dass das Verhalten der Agenten in Bezug auf ihre Emotionen validiert werden kann. Für jedes Szenario sollte jeweils eine Validierung aus der Vogelperspektive, eine Validierung aus der Egoperspektive eines Agenten und eine Validierung mit einer frei beweglichen Kamera, die von den Experten gesteuert werden kann und sich wie ein Agent bewegt, erfolgen. Durch die Vogelperspektive können die Experten die Makroebene validieren, da sie ihnen einen Überblick über das ganze Geschehen und das generelle Verhalten der Agenten gibt. Durch die Egoperspektive eines Agenten wiederum können die Experten gut das Verhalten des Agenten validieren, aus dessen Sicht sie das Geschehen sehen. Sie können sehen was der Agent sieht und überprüfen, ob das Verhalten des Agenten dazu passend ist. Den Experten sollten außerdem die Parameter wie Emotionen sowie die Sensorwerte mitgeteilt werden. So könnte unrealistisches Verhalten besser erkannt werden. Ein Beispiel hierfür wäre, dass sich ein Agent, obwohl er eine bessere Wahl hätte, in Gefahr begibt, indem er zum Beispiel in Richtung des Rauches und Feuers läuft, von dessen Existenz er weiß. Durch die Egoperspektive ist das Gesehene besser auf die eigenen Erfahrungen der Experten übertragbar. Die freibewegliche Kamera dient wiederum der Validierung des Verhaltens einzelner Agenten und der Gruppe. Die Experten können sich wie ein Agent frei durch die Umgebung bewegen und das Verhalten einzelner Agenten oder das Gruppenverhalten beobachten. Die Egoperspektive dieser Sicht erleichtert das Validieren insofern, dass die Sicht abermals eine Übertragung des Gesehenen auf reale Situationen erlaubt. Im Gegensatz zur Egoperspektive eines Agenten gibt sie dem Experten jedoch mehr Freiraum und erlaubt ihm, sich frei durch das Szenario zu bewegen und nach unrealistischen Begebenheiten zu suchen. Aus jeder Sicht sollte jeweils der ganze Ablauf der Szenarien betrachtet werden, jedoch sollte es auch möglich sein die Perspektive zeitweise zu ändern, falls etwas Auffälliges wahrgenommen wird, dass aus der aktuellen Perspektive nicht richtig erkannt werden kann. Auch sollte den Experten die Möglichkeit gegeben werden, die Simulation anzuhalten und das Geschehen „vor-“ und „zurückzuspulen“, um Situationen erneut betrachten zu können. Wird unrealistisches beziehungsweise unerwartetes Verhalten entdeckt, ist dieses zu beheben. Stehen mit den simulierten Szenarien vergleichbare Aufnahmen zur Verfügung, so können diese noch zusätzlich den simulierten Szenarien

gegenübergestellt werden, um sie zu vergleichen. Hierzu können unterschiedliche Bildschirme eingesetzt werden, von denen einer die Aufnahme zeigt und der andere die Simulation. Nun muss geprüft werden, ob das Verhalten der Agenten aus der Aufnahme auch in der Simulation wiederzufinden ist. Wird das Verhalten nicht wiedererkannt, muss geklärt werden, ob es sich um einen Fehler im Modell handelt oder ob dies durch andere Randbedingungen erklärbar ist. Handelt es sich um einen Fehler, muss er behoben werden.

Für die objektive Validierung im Betrieb müssen zuerst für die simulierten Szenarien passende Aufnahmen besorgt werden, die auch den Anforderungen der Verfahren entsprechen. Ist dies nicht möglich, so muss auf diesen Teil der Validierung verzichtet werden. Die mathematische Validierung sollte das bevorzugte Verfahren darstellen. Die statistische Validierung kann zusätzlich ausgeführt werden, wenn dafür genügend nutzbare Aufnahmen zur Verfügung stehen. Zuerst muss die Videoaufnahme mit einer Videoanalyse-Software analysiert werden. Für die mathematische Validierung muss die Anzahl der Personen in unterschiedlichen Bereichen der Aufnahme bestimmt werden. Die Bereiche der Aufnahme, die betrachtet werden, müssen so gewählt werden, dass sie sich durch Sektoren in der Simulation abdecken lassen. Für die statistische Validierung muss die Videoanalyse-Software für das ganze Bild verschiedene Werte bestimmen können wie die Positionen, Geschwindigkeit und Ausrichtung der Agenten. Diese werden gemittelt. Für die mathematische Validierung erfolgt die Analyse für jedes Einzelbild der Aufnahme beziehungsweise des Teils der Aufnahme, der verwendet werden soll. Bei der statistischen Validierung wird dies nur für eine zufällig bestimmte Menge an Bildern von verschiedenen Aufnahmen durchgeführt. Bei beiden Verfahren sollten die Ergebnisse der Analyse in eine XML-Datei gespeichert werden, um eine weitere Verwendung der Daten zu vereinfachen. Das Szenario der Aufnahme muss nun in der Simulation nachgestellt werden. Die Startverteilung in dem Bereich, zu dem die Aufnahme Informationen gibt, sollte so wie in der Aufnahme gewählt werden. Auch die Agenten müssen gemäß dem Wissen über das Szenario gewählt werden. Es ist möglich, dass nicht alle wichtigen Informationen zum Szenario aus der Aufnahme sichtbar sind. Dazu gehören unter anderem Informationen zur Architektur der Umgebung und den Agenten außerhalb des Bildes. Für das Verhalten in der Aufnahme, wie eine Fluchtbewegung der Agenten in eine bestimmte Richtung, kann unter anderem der Grund nicht sichtbar sein. Es müssen daher alle nötigen Informationen zum Szenario gesammelt und in der Simulation nachgestellt werden. Für den Vergleich müssen in der Simulation die Sektoren bestimmt werden, die den von der Aufnahme beobachteten Bereich abdecken. Die Positionen und Größen der Sektoren müssen so gewählt werden, dass sie den in der Aufnahme betrachteten Bereich so genau wie möglich in der Simulation abdecken und es möglich ist, die Werte der Sektoren mit den Werten der Bereiche der Aufnahme zu vergleichen. Bevor mit dem Vergleich begonnen werden kann, müssen für die mathematische Validierung die Zeitabstände, in denen in der Simulation gemessen wird, so gewählt werden, dass die verglichenen Bilder der Aufnahme und der Simulation jeweils den gleichen Zeitpunkt des Szenarios darstellen. Läuft die Aufnahme mit realer Geschwindigkeit ab oder gibt es eine Möglichkeit

die Geschwindigkeit der Zeit in der Aufnahme anders zu bestimmen, so müssen nur die Zeitabstände zwischen den Einzelbildern bestimmt werden. Da die Simulation die real im Szenario vergangene Zeit kennt, muss nur immer wenn der bestimmte Zeitabstand der Szenariozeit vergangen ist, die Bestimmung der Messwerte erfolgen. Ansonsten sollte zur Bestimmung dieser Zeitabstände das erste im Teil zur mathematischen Validierung vorgestellte Verfahren benutzt werden, bei dem durch Betrachtung der Anzahl der Personen und Agenten, die den betrachteten Teil der Umgebung zu verschiedenen Zeitpunkten verlassen haben, eine Bestimmung der Zeitabstände stattfindet.

Bevor mit der mathematischen Validierung begonnen werden kann, muss ein Grenzwert bestimmt werden, der die erlaubte Abweichung zwischen den Verteilungen der Aufnahme und den Verteilungen der Simulation bestimmt. Dieser kann zwischen 0 und 1 liegen. Je kleiner dieser Grenzwert ist, desto genauer müssen die Verteilungen übereinstimmen. Für die Ausführung der mathematischen Validierung wird nun die XML-Analysedatei der Aufnahme geladen und die Simulation wird gestartet. Nach dem bestimmten Zeitabstand wird die Simulation pausiert und die Verteilungen werden berechnet. Diese werden nun mit Hilfe eines Algorithmus mit den Verteilungen aus der Aufnahme verglichen. Liegt das Ergebnis dieser Berechnung unterhalb des Grenzwertes, so wird die Simulation wieder für den bestimmten Zeitabstand weiterlaufen gelassen. Ansonsten muss die Verteilung der Aufnahme in der Simulation repliziert werden. Danach fährt der Algorithmus mit den Vergleichen fort. Die Validierung ist beendet, wenn alle Verteilungen der Aufnahmen mit denen der Simulation verglichen wurden. Danach wird betrachtet, wie oft das Replizieren durchgeführt werden musste. Je geringer die Anzahl dieses Vorganges ist, desto genauer ist die Simulation für das Szenario und den Grenzwert. Die Validierung kann noch für andere Grenzwerte durchgeführt werden. Danach lässt sich die Fläche unter der Kurve, die die Anzahl des Replizierens zum Grenzwert darstellt, betrachten. Diese Fläche kann als Maß für die Genauigkeit der Simulation angesehen werden. Das Ergebnis dieser Validierung kann nun für verschiedene Zwecke eingesetzt werden. Es könnten die Zeitpunkte im Ablauf der Simulation betrachtet werden, bei denen es zu zu großen Abweichungen kam. Dadurch kann es möglich sein, Fehler im Verhalten der Simulation zu entdecken und die Schwere des Problems ist bestimmbar. Außerdem kann eine Kalibrierung durchgeführt werden. Hierzu werden die Parameter der Simulation beziehungsweise Agenten verändert und es wird erneute mathematische Validierung durchgeführt. Das Ergebnis der erneuten Validierung kann mit dem vorherigen Ergebnis verglichen werden. Dies kann automatisiert mit kleinen Änderungen an den Parametern durchgeführt werden. Werden bessere Werte für Parameter gefunden, so sollten diese gespeichert werden. Hierdurch ist es möglich ohne großen personellen Aufwand bessere Werte für die Parameter zu bestimmen. Allerdings sollte das Ergebnis der Kalibrierung erneut mit Face Validation betrachtet werden, so dass die Simulation nicht nur auf die mathematische Validierung zugeschnitten ist, sondern auch weiterhin realistische Ergebnisse liefert.

Die statistische Validierung sollte als optionaler Zusatz zur bisherigen Validierung aufgefasst werden. Sie ist nur möglich, wenn eine große Anzahl an passenden Aufnahmen vorhanden ist. Im Gegensatz zur mathematischen Validierung lassen sich keine Abläufe mit ihr validieren. Ihr Zweck ist es, verschiedene Zusammensetzungen von Gruppen zu erkennen. Es kann also mit ihr versucht werden, die richtige Zusammensetzung einer Gruppe zu finden, wenn ein bestimmtes Szenario nachgestellt werden soll. Nach Bestimmung der historischen Datenmenge durch Videoanalyse von verschiedenen Aufnahmen müssen die gleichen Werte in der Simulation bestimmt werden. Hierzu müssen auch hier für eine beliebig bestimmte Menge an Bildern diese Werte bestimmt und schließlich gemittelt werden, so dass ein Simulationslauf einen mehrdimensionalen Datenpunkt ergibt. Die Bestimmung sollte für die zum Bereich der Aufnahmen passenden Sektoren durchgeführt werden. Die Bestimmung der Werte erfolgt innerhalb des Systems und nicht über Videoanalyse. Die Werte müssen danach nur noch in einer XML-Datei gespeichert werden. Dies sollte aufgrund der Zufälligkeit für eine Menge an Simulationsläufen erfolgen, die groß genug ist, um zu einem verlässlichen Ergebnis zu kommen. Um die Komplexität der Berechnung der simplizierten Tiefe zu verringern, sollte eine Hauptkomponentenanalyse durchgeführt werden. Hierfür stehen viele freie und kostenlose statistischen Programme und Quellcode zur Verfügung (Wikipedia 2011). Nun muss nur noch mit Hilfe eines Algorithmus die simplizierte Tiefe der vereinfachten Datenpunkte der Simulationsläufe in der historischen Datenmenge bestimmt werden. Dies könnte zum Beispiel durch die Zuhilfenahme der freien Programmiersprache „R“ erfolgen. Diese unterstützt die Berechnung der simplizierten Tiefe (Agostinelli und Romanazzi 2009). Die Ergebnisse dieser Berechnung werden zusätzlich mit einem vorher zu definierenden Grenzwert verglichen, der angibt, ob ein Datenpunkt noch als Datenpunkt des gleichen Szenarios angesehen wird. Je höher dieser Grenzwert ist, desto geringere Abweichungen werden zugelassen. Schließlich kann angegeben werden, wie viele Prozent der Datenpunkte der Simulationsläufe als Datenpunkt des gleichen Szenarios angesehen wurden. Je höher diese Prozentzahl ist, desto höher ist die Übereinstimmung. Auch dieses Ergebnis kann zu Optimierungszwecken eingesetzt werden, indem die Agentenparameter angepasst werden und das ganze Validierungsverfahren erneut durchgeführt wird. Auch hier ließe sich dieses automatisiert durchführen. Ebenfalls gilt, dass nach der Kalibrierung erneut Face Validation durchgeführt werden sollte, damit auch das Verhalten auf der Mikroebene valide bleibt.

Nach Durchführung dieser Schritte ist die Gruppensimulation so gut validiert, wie es derzeit möglich ist. Eine objektivere Validierung wäre nur möglich, wenn mehr Daten zur Validierung zur Verfügung stehen würden und neue objektive Verfahren entwickelt werden, welche das Validieren der Gruppensimulation auf allen Ebenen ermöglichen. Durch die durchgeführte Kombination der einzelnen Verfahren werden jedoch die Vorteile der Verfahren vereint und die meisten Nachteile beseitigt.

4 Schlussbetrachtung

Dieses Kapitel fasst die wichtigsten Erkenntnisse der Arbeit zusammen und zieht aus diesen ein Fazit. Des Weiteren wird ein Ausblick über weitere in diesem Bereich zu tätige Forschungen gegeben, die nötig sind, um eine noch verlässlichere Validierung zu ermöglichen.

4.1 Zusammenfassung der Ergebnisse

In dieser Arbeit wurde gezeigt, dass die Validierung von Gruppensimulationen eine wichtige Rolle im Entwicklungsprozess spielt. Nur durch Validierung ist es möglich zu beweisen, dass die Simulation für den geplanten Einsatzzweck auch tatsächlich die geforderte Genauigkeit besitzt. Dies ist auch die Voraussetzung dafür, dass die Simulation zur Vorhersage von realen Ereignissen benutzt werden kann, da sonst die Korrektheit der Vorhersagen nicht gewährleistet werden kann. Die Validierung von agentenbasierten Gruppensimulationen stellt trotz der Wichtigkeit dieser noch immer ein großes Problem dar. Begründet liegt dies vor allem im Fehlen von ausreichend detaillierten Daten, die für die Validierung der Makro- als auch der Mikroebene verwendet werden könnten. Aus diesem Grund wurde eine Vielzahl von Verfahren entwickelt, die trotz der Probleme eine bestmögliche Validierung erlauben sollen. Verschiedene Entwicklungsphasen erfordern unterschiedliche Validierungen, so dass auch in den verschiedenen Phasen unterschiedliche Validierungsverfahren benötigt werden. Generell gilt, dass in jeder Entwicklungsphase Validierung durchgeführt werden sollte und erst mit dem nächsten Schritt fortgefahren werden sollte, wenn das Ergebnis des vorherigen Schrittes validiert wurde. Dabei ist es besser, die Validierung von einer dritten Partei durchführen zu lassen, da diese eine objektivere Sichtweise hat als die Entwickler. Beim Betrachten der Verfahren zur Validierung wurden nicht nur die speziellen Verfahren zur Validierung von Gruppensimulationen betrachtet, sondern auch allgemeinere Verfahren für Simulationen und agentenbasierte Simulationen. Dabei hat sich gezeigt, dass zwar viele dieser allgemeinen Verfahren auch in den spezielleren Verfahren wiederzufinden sind, jedoch Verfahren wie die Degenerierungstests und Kalibrierung nicht berücksichtigt werden, obwohl die Anwendung dieser während der Validierung sinnvoll ist. Die Verfahren wurden beschrieben und auf ihre Brauchbarkeit untersucht. Dabei hat sich gezeigt, dass für eine komplette Validierung der Gruppensimulation mehrere Validierungsschritte und Validierungsverfahren während der Entwicklung nötig sind. Des Weiteren erlaubt keines der vorhandenen Verfahren eine Validierung aller Ebenen einer Gruppensimulation auf eine objektive Weise. Die objektiven Verfahren ermöglichen bisher noch keine Validierung auf der

Mikroebene. Dies liegt unter anderem darin begründet, dass noch nicht genügend wissenschaftlich bestimmte und verlässliche Daten vorhanden sind, die hierfür benötigt werden. Die Validierung dieser Ebene kann also nur auf subjektiver Basis durch Menschen erfolgen. Diese müssen sich sowohl das zugrundeliegende konzeptuelle Modell ansehen und überprüfen als auch verschiedenen Simulationsläufe, um auf Basis ihrer Erfahrungen und ihres Wissens über die Validität der Simulation zu entscheiden. Dieses Betrachten und Bewerten erfolgt jedoch subjektiv und kann somit von Mensch zu Mensch anders ausfallen. Um verlässlichere Ergebnisse zu gewinnen, sind daher auch die objektiven Verfahren wichtig, da diese zumindest eine objektive Validierung auf der Makroebene ermöglichen und somit die Validierung nicht nur auf subjektiven Einschätzungen beruht. Durch die objektiven Verfahren fällt es außerdem sehr viel leichter verschiedene Simulationen zu vergleichen. Dadurch wird es möglich, durch Kalibrierung die Simulation zu verändern und das Ergebnis dieser Änderung mit vorherigen Ergebnissen zu vergleichen, wodurch eine Optimierung der Simulation vorgenommen werden kann. Außerdem besteht die Möglichkeit verschiedene Modelle auf objektive Weise zu vergleichen und somit das Modell zu bestimmen, das zumindest auf der Makroebene bessere Ergebnisse liefert. Auf Basis dieser Überlegungen wurde eine Kombination von verschiedenen Verfahren für die verschiedenen Entwicklungsschritte einer Gruppensimulation in einem Validierungskonzept beschrieben. Dass dieses Konzept auch wirklich anwendbar ist und wie dieses Konzept anwendbar ist, wurde anhand eines Beispiels gezeigt. Das Konzept beschreibt die einzelnen auszuführenden Schritte für die Validierung und welche Voraussetzungen für die Durchführung der einzelnen Schritte vorhanden sein müssen. Es kann somit als Anleitung zur kompletten Validierung einer Gruppensimulation verwendet werden und erleichtert die Validierung.

4.2 Fazit und Ausblick

Die Arbeit hat gezeigt, wie man trotz der Probleme, die es bei der Validierung von Gruppensimulationen zurzeit gibt, eine möglichst verlässliche Validierung durchführen kann. Es wurde jedoch auch deutlich, dass die komplette Validierung ohne menschliche Bewertung von Teilen der Simulation nicht auf allen Ebenen möglich ist. Dies bringt eine gewisse Subjektivität in die Ergebnisse der Validierung, was einen negativen Einfluss auf die Verlässlichkeit und Vergleichbarkeit der Ergebnisse hat.

Um noch verlässlichere Ergebnisse bei der Validierung zu erzielen, ist es notwendig, dass mehr für die Validierung der Mikro- und Makroebene nutzbare Daten vorhanden sind. Hierfür müssen unter anderem neuen Verfahren entwickelt werden, die es erlauben mehr Daten zu den Individuen in den zu simulierenden Szenarien zu sammeln. Dies ist die Voraussetzung für neue objektive Verfahren, die auch eine Validierung von individuellen Agenten durchführen können. Das Ziel muss eine Validierung sein, die nur vollkommen objektive Ergebnisse für alle Ebenen liefert. Dadurch wird die Verlässlichkeit der Ergebnisse der Validierung gesteigert. Dies erlaubt bessere Vergleichbarkeit von verschiedenen Simulationen

und vereinfacht die Optimierung der Simulation, so dass auch diese verlässlichere Ergebnisse liefert. Die Verlässlichkeit der Ergebnisse der Simulation sollte die höchste Priorität haben.

5 Literaturverzeichnis

Balci 1994

BALCI, Osman: Validation, verification, and testing techniques throughout the life cycle of a simulation study. In: TEW, Jeffrey D. (Hrsg.) ; MANIVANNAN, Mani S. (Hrsg.) ; SADOWSKI, Deborah A. (Hrsg.): *Proceedings of the 1994 Winter Simulation Conference*. San Diego, CA : Society for Computer Simulation International, 1994. ISBN 0-7803-2109-X, S. 215-220. – Online verfügbar unter: <http://portal.acm.org/citation.cfm?id=193201.194018> Abruf: 2011-06-19

Balci 1997

BALCI, Osman: Verification, validation and accreditation of simulation models. In: ANDRADÓTTIR, Sigrún (Hrsg.) ; HEALY, Kevin J. (Hrsg.) ; WITHERS, David H. (Hrsg.): *Proceedings of the 1997 Winter Simulation Conference*. Washington, DC : IEEE Computer Society, 1997. ISBN 0-7803-4278-X, S. 135-141. – Online verfügbar unter: <http://dx.doi.org/10.1145/268437.268462> Abruf: 2011-05-31

Baldowski 2011

BALDOWSKI, Mariusz: Entwicklung eines 3D-Geoinformationssystem für Gefahrensituationen im In- und Outdoorbereich im Rahmen von WALK. In: WITTMANN, Jochen (Hrsg.), WOHLGE-MUTH, Volker (Hrsg.): *Simulation in Umwelt- und Geowissenschaften : Workshop Berlin 2011*. Aachen : Shaker Verlag, 2011. ISBN 978-3-8440-0284-3, S. 113-122.

Banerjee und Kraemer 2010

BANERJEE, Bikramjit ; KRAEMER, Landon: Validation of Agent Based Crowd Egress Simulation (Extended Abstract). In: LUCK, Michael (Hrsg.) ; SEN, Sandip (Hrsg.): *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*. Bd. 1. Richland, SC : International Foundation for Autonomous Agents and Multiagent Systems, 2010. ISBN 978-0-9826571-1-9, S. 1551-1552. – Online verfügbar unter: <http://portal.acm.org/citation.cfm?id=1838476> Abruf: 2011-07-11

Burr et al. 2004a

BURR, Michael A. ; RAFALIN, Eynat ; SOUVAINE, Diane L.: Simplicial depth : An improved definition, analysis, and efficiency for the finite sample case. In: *Proceedings of the 16th Canadian Conference on Computational Geometry*. Montreal, Québec : The Canadian Conference on Computational Geometry, 2004. S. 136-139. – Online verfügbar unter: <http://www.cccg.ca/proceedings/2004/49.pdf> Abruf: 2011-07-17

Burr et al. 2004b

BURR, Michael A. ; RAFALIN, Eynat ; SOUVAINE, Diane L.: Simplicial depth : *An improved definition, analysis, and efficiency for the finite sample case*. 2004. Online verfügbar unter: http://www.cs.tufts.edu/research/geometry/depth/simplicial_cccg04.pps Abruf: 2011-07-17

Fridman und Kaminka 2010

FRIDMAN, Natalie ; KAMINKA, Gal A.: Modeling pedestrian crowd behavior based on a cognitive model of social comparison theory. In: *Computational & Mathematical Organization Theory* 16 (2010), Nr. 4, S. 348-372. – Online verfügbar unter: <http://dx.doi.org/10.1007/s10588-010-9082-2> Abruf: 2011-07-11

Härdle und Simar 2004

HÄRDLE, Wolfgang ; SIMAR, Léopold: *Applied Multivariate Statistical Analysis (Härdle, Simar)* ISBN: 3-540-03079-4. Stand: 2004-08-26 <http://pascal.upf.edu/am/dades/h-s-mva-book/mvahtmlframe149.html> Abruf: 2011-08-18

Pelechano et al. 2008

PELECHANO, Nuria ; STOCKER, Catherine ; ALLBECK, Jan ; BADLER, Norman: Being a Part of the Crowd: Towards Validating VR Crowds Using Presence. In: PADGHAM, Lin (Hrsg.) ; PARKES, David C. (Hrsg.) ; MÜLLER, Jörg (Hrsg.): *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*. Bd. 1. Richland, SC : International Foundation for Autonomous Agents and Multiagent Systems, 2008. ISBN 978-0-9817381-0-9, S. 136-142. – Online verfügbar unter: <http://portal.acm.org/citation.cfm?id=1402383.1402407> Abruf: 2011-05-31

Pelechano und Malkawi 2008

PELECHANO, Nuria ; MALKAWI, Ali: Evacuation simulation models: Challenges in modeling high rise building evacuation with cellular automata approaches. In: *Automation in Construction* 17 (2008), Nr. 4, S. 337-385. – Online verfügbar unter: <http://dx.doi.org/10.1016/j.autcon.2007.06.005> Abruf: 2011-06-16

Klingenberg 2010

KLINGENBERG, Arne: *Prototypische Entwicklung eines emotionalen Agenten auf der Basis des Goal Oriented Action Plannings*. Hochschule für Angewandte Wissenschaften Hamburg, Bachelor-Thesis, 2010. – Online verfügbar unter: <http://opus.haw-hamburg.de/volltexte/2010/964/> Abruf: 2011-08-07

Klügl 2008

KLÜGL, Franziska: A Validation Methodology for Agent-Based Simulations. In: WAINWRIGHT, Roger L. (Hrsg.) ; HADDAD, Hisham M. (Hrsg.): *Proceedings of the 2008 ACM symposium on Applied computing*. New York, NY : Association for Computing Machinery, 2008. ISBN 978-1-59593-753-7, S. 39-43. – Online verfügbar unter: <http://dx.doi.org/10.1145/1363686.1363696> Abruf: 2011-06-12

Klüpfel 2003

KLÜPFEL, Hubert L.: *A Cellular Automaton Model for Crowd Movement and Egress Simulation*. Universität Duisburg-Essen, Dissertation, 2003. – Online verfügbar unter: http://deposit.ddb.de/cgi-bin/dokserv?idn=96883180x&dok_var=d1&dok_ext=pdf&filename=96883180x.pdf Abruf: 2011-07-14

Malone et al. 2008

MALONE, Linda C. ; KAUP, D. J. ; OLESON, Rex ; ROSA, Mario ; JENTSCH, Florian ; CLARKE, T. L. ; FAULKNER, Jennifer ; JAGGIE, Renee: Validation of Crowd Simulations. In: COOK, D. (Hrsg.) ; TAYLOR, K. (Hrsg.): *Proceedings of the 2008 Summer Computer Simulation Conference*. San Diego, CA : The Society for Modeling and Simulation, 2008. ISBN 1-56555-320-9, S. 363-370. – Online verfügbar unter: http://www.math.ucf.edu/~kaup/kaup_files/Pub237.pdf Abruf: 2011-07-11

Massive Software 2011

MASSIVE SOFTWARE (Hrsg.): *[Massive Software :: Film]* Stand: 2011-06-14 <http://massivesoftware.com/film.html> Abruf: 2011-06-14

Meyer-König 2004

MEYER-KÖNIG, Tim: *RiMEA-Projekt – Analyse*. RiMEA-Projekt, 2004. – Online verfügbar unter: <http://www.rimea.de/downloads/analysen/analyse.doc> Abruf: 2011-07-21

Meyer-König et al. 2009

MEYER-KÖNIG, Tim ; MOROGE, Christian ; SCHRECKENBERG, Michael ; SCHWENDIMANN, Martin ; WALDAU, Nathalie: *Richtlinie für Mikroskopische Entfluchtungsanalysen : Version: 2.2.1 08. Juni 2009*. RiMEA-Projekt, 2009. – Online verfügbar unter: <http://www.rimea.de/downloads/r2.2.1.pdf> Abruf: 2011-05-31

Agostinelli und Romanazzi 2009

AGOSTINELLI, Claudio ; ROMANAZZI, Mario: *Package 'localdepth' : April 19, 2009*. The R Project for Statistical Computing, 2009. – Online verfügbar unter: <http://cran.r-project.org/web/packages/localdepth/localdepth.pdf> Abruf: 2011-08-24

Roiss 2011

ROISS, Patrick: Modellierung von Agenten mit Hilfe emotionaler Verhaltensmodelle. In: WITTMANN, Jochen (Hrsg.), WOHLGEMUTH, Volker (Hrsg.): *Simulation in Umwelt- und Geowissenschaften : Workshop Berlin 2011*. Aachen : Shaker Verlag, 2011. ISBN 978-3-8440-0284-3, S. 123-128.

Sargent 2008

SARGENT, Robert G.: Verification and validation of simulation models. In: MASON, Scott J. (Hrsg.) ; HILL, Ray R. (Hrsg.) ; MÖNCH, Lars (Hrsg.): *Proceedings of the 2008 Winter Simulation Conference*. Piscataway, NJ : IEEE, 2008. ISBN 978-1-4244-2708-6, S. 157-169. – Online verfügbar unter: <http://portal.acm.org/citation.cfm?id=1516744.1516780> Abruf: 2011-06-12

Schlesinger et al. 1979

SCHLESINGER, Stewart ; CROSBIE, Roy E. ; GAGNÉ, Roland E. ; INNIS, George S. ; LALWANI, C. S. ; LOCH, Joseph ; SYLVESTER, Richard J. ; WRIGHT, Richard D. ; KHEIR, Naim ; BARTOS, Dale: Terminology for model credibility. In: *Simulation* 32 (1979), Nr. 3, S. 103-104. – Online verfügbar unter: <http://dx.doi.org/10.1177%2f003754977903200304> Abruf: 2011-05-31

Thalmann und Musse 2007

THALMANN, Daniel ; MUSSE, Soraia R.: *Crowd Simulation*. 1. Aufl. London: Springer-Verlag London Limited, 2007. – ISBN 978-1-84628-824-1

Thiel 2011

THIEL, Christian: Realisierung eines Multiagentensystems zur Simulation von Fußgängerströmen. In: WITTMANN, Jochen (Hrsg.), WOHLGEMUTH, Volker (Hrsg.): *Simulation in Umwelt- und Geowissenschaften : Workshop Berlin 2011*. Aachen : Shaker Verlag, 2011. ISBN 978-3-8440-0284-3, S. 129-138.

Thiel-Clemen et al. 2011

THIEL-CLEMEN, Thomas ; KÖSTER, Gerta ; SARSTEDT, Stefan: WALK – Emotion-based pedestrian movement simulation in evacuation scenarios. In: WITTMANN, Jochen (Hrsg.), WOHLGEMUTH, Volker (Hrsg.): *Simulation in Umwelt- und Geowissenschaften : Workshop Berlin 2011*. Aachen : Shaker Verlag, 2011. ISBN 978-3-8440-0284-3, S. 103-112.

Thiel-Clemen und Klingenberg 2010

THIEL-CLEMEN, Thomas ; KLINGENBERG, Arne: Kombination von zielorientiertem Verhalten und Emotionen in Individuen-orientierten Simulationen. In: WITTMANN, Jochen (Hrsg.) ; MARETIS, Dimitris K. (Hrsg.): *Simulation in Umwelt- und Geowissenschaften : Workshop Osnabrück 2010*. Aachen : Shaker Verlag, 2010. ISBN 978-3-8322-9367-3, S. 71-80.

Wikipedia 2011

WIKIPEDIA (Hrsg.): *Principal component analysis* Stand: 2011-08-24 [http://en.wikipedia.org/wiki/Principal component analysis#Software.2Fsource code](http://en.wikipedia.org/wiki/Principal_component_analysis#Software.2Fsource_code) Abruf: 2011-08-24

Xiang et al. 2005

XIANG, Xiaorong ; KENNEDY, Ryan ; MADEY, Gregory: Verification and Validation of Agent-based Scientific Simulation Models. In: YILMAZ, Levent (Hrsg.): *2005 Agent-Directed Simulation Symposium*. San Diego, CA : The Society for Modeling and Simulation International, 2005. ISBN 1-56555-291, S. 47-55. – Online verfügbar unter: http://www.nd.edu/~nom/Papers/ADS019_Xiang.pdf Abruf: 2011-05-31

Zeigler et al. 2000

ZEIGLER, Bernard P. ; PRAEHOFER, Herbert ; KIM, Tag G.: *Theory of Modeling and Simulation : Integrating Discrete Event and Continous Complex Dynamic Systems*. 2. Aufl. San Diego, CA : Elsevier Science, 2000. – ISBN 0-12-778455-1

Zhou et al. 2010

ZHOU, Suiping ; CHEN, Dan ; CAI, Wentong ; LUO, Linbo ; LOW, Malcolm Y. H. ; TIAN, Feng ; TAY, Victor S.-H. ; ONG, Darren W. S. ; HAMILTON, Benjamin D.: Crowd Modeling and Simulation Technologies. In: *ACM Transactions on Modeling and Computer Simulation* 20 (2010), Nr. 4, S. 20:1-20:35. – Online verfügbar unter: <http://dx.doi.org/10.1145/1842722.1842725> Abruf: 2011-05-31

Versicherung über Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe.

Hamburg, den _____