

Spieltheorie als Entscheidungswerkzeug für patrouillierende Roboter

Tobias Schmidt

Zusammenfassung—Die Verwendung von mobilen Robotern hat im Lebensalltag des Menschen längst Einzug gefunden. Neben Haushalt oder wissenschaftlichen Untersuchungen kann ein patrouillierender Roboter auch im Bereich der Liegenschafts- und Objektüberwachung eingesetzt werden. Hierbei werden meist spieltheoretische Modelle angewendet, um Strategien für die Wegplanung des Roboters zu finden.

Dieser Ausarbeitung zum Seminar liegt die Arbeit „A Game-Theoretic Approach to Determining Efficient Patrolling Strategies for Mobile Robots“ von Amigoni et al. zugrunde, in der die Autoren einen spieltheoretischen Ansatz für die Ermittlung einer effizienten Strategie für einen patrouillierenden Roboter vorstellen. Zum Verständnis der Arbeit von Amigoni et al. werden in dieser Ausarbeitung relevante, spieltheoretische Grundlagen vermittelt. Außerdem werden die Erkenntnisse und Ergebnisse der Arbeit zusammengefasst und kritisch betrachtet.

I. EINLEITUNG

Mobile Roboter werden für die verschiedensten Aufgaben eingesetzt. Dazu zählen beispielsweise sicherheitskritische Aufgaben wie Atomanlagen-Inspektionen, Katastrophenschutz oder Minensuche. Außerdem unterstützen mobile Roboter Untersuchungen von für den Menschen unerreichbaren Gebieten wie bspw. andere Planeten, Pipelines, etc.. Auch im alltäglichen Leben agieren mittlerweile mobile Roboter und übernehmen unangenehme Tätigkeiten im Haushalt, wie bspw. Staubsaugen. Patrouillen in der Liegenschafts- oder Objektüberwachung sind ebenfalls ein potentiell Einsatzgebiet.

Der klare Vorteil des Einsatzes von mobilen Robotern ist zum einen, dass die Tätigkeit ohne oder nur mit geringer Beaufsichtigung durchgeführt werden kann. Zum anderen haben Roboter im Vergleich zum Menschen in der Regel eine kürzere Reaktionszeit.

Je nach Tätigkeitsbereich ist der Autonomiegrad des eingesetzten Roboters unterschiedlich hoch. Je höher der Autonomiegrad ist, desto öfter muss der Roboter Entscheidungen eigenständig treffen.

In der nachfolgenden Betrachtung sollen die Tätigkeitsbereiche auf solche beschränkt werden, bei denen ein mobiler Roboter verschiedene Objekte bewacht. Es sollen Roboter betrachtet werden, die beispielsweise in einem Museum als Wachposten eingesetzt werden. Im Gegensatz zu einem Roboter, der innerhalb eines Bereiches nach Eindringlingen sucht, soll hier der Roboter eine Umgebung mit unterschiedlich bewerteten Objekten bewachen.

Bei einem solchen patrouillierenden Roboter ist der Erfolg der Missionen von der Wahl der Patrouillenwege abhängig. So muss der Roboter bei der Wegplanung bspw. die Objektbewertung und die ihm bekannten Eindringlingsarten berücksichtigen.

Diese Ausarbeitung zum Seminar beruht auf der Arbeit „A Game-Theoretic Approach to Determining Efficient Patrolling Strategies for Mobile Robots“ [1] von Amigoni et al.. Die Autoren stellen in der Arbeit einen spieltheoretischen Ansatz vor, um eine effiziente Strategie für einen patrouillierenden Roboter zu ermitteln. Sie modellieren ein Spiel mit einem patrouillierenden und einem eindringenden Spieler. Weiterhin formulieren sie für die Berechnung der Strategie des patrouillierenden Roboters zwei mathematische, lineare Probleme und evaluieren ihren Lösungsansatz in Bezug auf unterschiedliche Eigenschaften über Vergleich mit anderen Methoden.

Diese Ausarbeitung soll spieltheoretische Grundlagen vermitteln, welche zum Verständnis der Arbeit von Amigoni et al. beitragen. Außerdem sollen die Erkenntnisse sowie Hintergrundinformationen bzgl. der Arbeit zusammengefasst werden.

In Abschnitt II wird die Problemstellung spezifiziert, die Aufschluss über die gegebenen Informationen und gesuchten Lösungen gibt. Zum Verständnis benötigte spieltheoretische Begriffe und Methoden werden in Abschnitt III erläutert. In Abschnitt IV erfolgt eine genaue Formulierung des Spiels zwischen dem patrouillierenden Roboter und dem Eindringling. Es werden die Annahmen dargestellt, welche Amigoni et al. in ihrer Arbeit getroffen haben, um ihr Modell zu begründen. Anschließend werden in Abschnitt V die Evaluierungsergebnisse der Arbeit von Amigoni et al. präsentiert. Abschließend erfolgt in Abschnitt VI eine kurze Zusammenfassung der vorliegenden Arbeit sowie eine kritische Betrachtung des Ansatzes und der Ergebnisse.

II. PROBLEMSTELLUNG

Die Herausforderungen an die Wissenschaft in der Forschung mit mobilen Robotern bestehen darin, geeignete Modelle zu finden, deren formulierte Probleme sich mit einem vertretbaren Aufwand lösen lassen. Im Bereich der patrouillierenden Roboter beinhalten diese Modelle eine detailliert beschriebene Umgebung, die Beschreibung der teilnehmenden Roboter mitsamt diversen Annahmen über deren Bewegungseigenschaften und Informationsstrukturen.

Gesucht sind folglich

- **Modelle**, welche sowohl die aufgezählten Attribute Umgebung, Bewegungseigenschaften und Informationsstrukturen der teilnehmenden Roboter allumfassend beschreiben und

- **effiziente Strategien** für den patrouillierenden Roboter (oder je nach Definition auch ein Team von solchen), um den oder die eindringenden Roboter zu fassen.

Die Modelle haben in der Regel Wurzeln in der Spieltheorie. Die Umgebung definiert das Spielfeld, die Roboter definieren die teilnehmenden Spieler, die Annahmen über die Eigenschaften der Roboter definieren die Strategien und die Regeln. Letztere können zusätzlich durch Rahmenbedingungen des Modells erweitert werden.

Die Formulierung der Anforderungen an eine Patrouillierungsstrategie erfolgt zumeist mathematisch. Gelöst werden solche mathematischen Probleme durch unterschiedliche Algorithmen, die je nach Formulierung und Anwendung mehr oder weniger effizient sind. Bei unterschiedlichen Modellen kann die Anwendung des gleichen Algorithmus zu verschiedenen Ergebnissen bzgl. der Effizienz führen.

Für die Lösung werden also

- umfassende **spieltheoretische Modelle** genutzt und
- Patrouillierungsstrategien als **mathematische Probleme** formuliert, um sie mit für solche Probleme gültigen Algorithmen lösen zu können.

III. GRUNDLAGEN DER SPIELTHEORIE

Im diesem Abschnitt werden wesentlich relevante spieltheoretische Begriffe erläutert. Außerdem erfolgt eine Beschreibung von verschiedenen Lösungsansätzen in unterschiedlichen Spielen.

Grundsätzlich wird in der Spieltheorie zwischen Spielen in Normalform und Spielen in Extensivform unterschieden.

A. Normalformspiele

Die Normalform wird in der Literatur auch Matrixform [2] oder strategische Form [3] genannt. „Durch Normalformspiele [werden] Konfliktsituationen mit einem Minimum an formalen Konzepten [beschrieben]. Zugfolge, Informationsstand der Spieler über den bisherigen Spielablauf, Zufallszahl usw. werden nicht explizit behandelt, sondern sie gehen alle ein in das Konzept der Strategie und der Auszahlungsfunktion eines Spielers.“ (aus [2, S. 11])

Zur vollständigen formalen Beschreibung eines Spiels in Normalform gehören die Anzahl $n = |I|$ der Spieler, die Spielregeln und die Informationsstruktur, die Menge Σ_i der Strategien σ_i für jeden Spieler i sowie die Auszahlungsfunktionen $H_i : \Sigma \rightarrow \mathbb{R}$ für jeden Spieler i mit $\Sigma := \Sigma_1 \times \dots \times \Sigma_n$. Die Kombination der gewählten Strategien aller n Spieler wird Strategiekonfiguration $\sigma = (\sigma_1, \dots, \sigma_n) \in \Sigma$ genannt. Die Auszahlungsfunktion des Spielers i ist somit durch $H_i(\sigma)$ gegeben.

Weiterhin sind in der Literatur Normalformspiele definiert, bei welchen den Spielern gemischte Strategien gestattet sind. Gemischte Strategien sind Wahrscheinlichkeitsverteilungen über der Menge der Strategien. Zu jeder Strategie σ_j wird bei solchen Spielen eine Wahrscheinlichkeit q_{σ_j} zugeordnet, wobei $0 \leq q_{\sigma_j} \leq 1$ und $\sum_{j=1}^n q_{\sigma_j} = 1$ für alle n Strategien eines Spielers gilt. Der Nutzen—bei Spielen mit gemischten Strategien wird statt von Auszahlung vom

Tabelle I. ÜBERSICHTLICHE DARSTELLUNG DER STRATEGIEN UND AUSZAHLUNGEN BEIM GEFANGENENDILEMMA.

		P_2	
		σ_{21}	σ_{22}
P_1	σ_{11}	2, 2	0, 3
	σ_{12}	3, 0	1, 1

Nutzen gesprochen—berechnet sich entsprechend als Summe der Produkte der Auszahlungen mit den zugehörigen Wahrscheinlichkeiten:

$$u(\cdot) = \sum_{j=1}^n \sigma_j \cdot q_{\sigma_j}$$

Für eine detailliertere Definition und Beispiele wird auf Literatur ([4], [5], [3], [2]) verwiesen.

1) *Beispiel Gefangenendilemma*: Das wohl bekannteste Beispiel für ein Normalformspiel ist das Gefangenendilemma.

Die vollständige formale Spielbeschreibung lautet (vgl. [3]):

- **Anzahl der Spieler**: Zwei Verdächtige P_1 und P_2 .
- **Regeln**: Die Verdächtigen sind in unterschiedlichen Räumen und können sich nicht absprechen.
- **Strategien**: Jeder Spieler i hat die Wahl zwischen $\sigma_{i1} :=$ „schweigen“ und $\sigma_{i2} :=$ „gestehen“.
- **Informationsstruktur**:
 - Beide Spieler kennen sowohl die eigenen Strategien als auch die Strategien des anderen.
 - Beide Spieler kennen sowohl die eigene Auszahlungsfunktion als auch die Auszahlungsfunktion des anderen.
 - Beide Spieler wissen, dass der andere Spieler obiges weiß.
 - Kein Spieler kennt die gespielte Strategie des anderen.
- **Auszahlungen**:
 - Schweigen beide, erhalten sie beide eine Auszahlung von 2:

$$H_1(\sigma_{11}, \sigma_{21}) = H_2(\sigma_{11}, \sigma_{21}) = 2.$$

- Gesteht einer und der andere schweigt, erhält der geständige eine Auszahlung von 3 und der schweigende eine Auszahlung von 0:

$$H_1(\sigma_{12}, \sigma_{21}) = H_2(\sigma_{11}, \sigma_{22}) = 3$$

und

$$H_1(\sigma_{11}, \sigma_{22}) = H_2(\sigma_{12}, \sigma_{21}) = 0.$$

- Gestehen beide, erhalten sie beide eine Auszahlung von 1:

$$H_1(\sigma_{12}, \sigma_{22}) = H_2(\sigma_{12}, \sigma_{22}) = 1.$$

Eine übersichtliche Darstellung der Strategien und Auszahlungen des Gefangenendilemmas in Normalform zeigt Tabelle I.

2) *Nash-Gleichgewicht*: Nach [2] ist eine Strategiekonfiguration $\sigma^* = (\sigma_1^*, \dots, \sigma_n^*)$ genau dann ein Nash-Gleichgewicht, wenn für jede Strategie $\sigma_i \in \Sigma_i$ eines jeden Spielers $i \in I$ gilt:

$$H_i(\sigma_{-i}^*, \sigma_i^*) \geq H_i(\sigma_{-i}^*, \sigma_i)$$

Demnach muss jede Strategiekonfiguration überprüft werden, ob für mindestens einen Spieler ein profitableres Ergebnis durch Abweichen zu einer alternativen Strategie möglich ist. Wenn keine solche Abweichung möglich ist, hat kein Spieler einen Anreiz, nicht die entsprechende Strategie zu wählen.

Beim Gefangenendilemma ist das einzige Nash-Gleichgewicht die Strategiekonfiguration $\sigma_{Nash} = (\sigma_{12}, \sigma_{22})$. Sowohl Spieler P_1 als auch Spieler P_2 hätten eine geringere Auszahlung, wenn sie schweigen würden, als wenn sie gestehen.

Für die Definition des Nash-Gleichgewichts bei gemischten Strategien sei an dieser Stelle wiederum auf Literatur wie [4], [5], [3], [2] verwiesen.

B. Extensivformspiele

Die Beschreibung eines Spiels in extensiver Form ermöglicht gegenüber der Beschreibung in strategischer Form (Normalform) die Einbeziehung weiterer Eigenschaften eines Spiels. Spielbeschreibungen in extensiver Form liefern Eigenschaften wie die Zugfolge oder den Informationsstand eines Spielers in einem Zug. Somit können Spieler je nach spezifizierter Informationsstruktur Wissen über vergangene Aktionen der anderen Spieler haben. Die vollständige formale Beschreibung eines Spiels in extensiver Form unterscheidet sich nicht wesentlich von der eines Spiels in Normalform.

Eine vollständige Beschreibung eines Spiels in extensiver Form umfasst nach [3] die Anzahl $n = |I|$ der Spieler, welche Spieler zu welchem Zeitpunkt am Zug sind, die Spielregeln und die Informationsmenge jedes Spielers für jeden Zug, die Menge S_i^k der möglichen Aktionen s_i^k für jeden Spieler i in jedem Zug k , sowie der Auszahlungen u_i^p für jeden Spieler i und alle möglichen Kombinationen p von Entscheidungszügen, wodurch der Spielverlauf bestimmt ist.

Die Darstellung eines Spiels in Extensivform erfolgt zumeist durch einen Spielbaum, der formal als zyklusfreier zusammenhängender mathematischer Graph mit einem eindeutigen Wurzelknoten beschrieben werden kann. Die Knoten entsprechen den Entscheidungssituationen und die Kanten entsprechen den Aktionsmöglichkeiten der Spieler. Die Blätter definieren die Auszahlungen aller Spieler für jeden möglichen Spielverlauf.

1) *Beispiel Markteintrittsspiel*: Ein bekanntes Szenario, das als Extensivformspiel dargestellt werden kann, ist das Markteintrittsspiel.

Abbildung 1 zeigt den Graph des Spiels. Die vollständige formale Spielbeschreibung dazu lautet (vgl. [3]):

- **Anzahl der Spieler**: Zwei Unternehmen P_1 und P_2 .
- **Zugfolge**: Die Spieler ziehen abwechselnd und P_2 zieht vor P_1 .
- **Spielregeln**:
 - Absprachen sind untersagt.
 - Jeder Spieler zieht maximal ein Mal.
- **Informationsstruktur**:

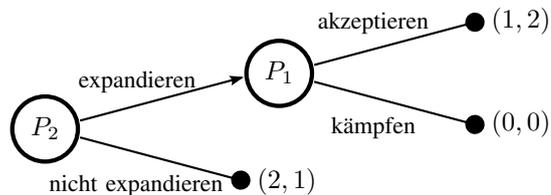


Abbildung 1. Graph des Markteintrittsspiels in extensiver Form mit den Auszahlungen (u_1^p, u_2^p) .

- P_2 kennt die möglichen Aktionen von P_1 ; P_1 kennt die möglichen Aktionen von P_2 .
- P_2 und P_1 kennen sowohl die eigenen als auch die Auszahlung des anderen für jeden Entscheidungspfad.
- P_1 kennt die durchgeführte Aktion von P_2 .
- P_2 weiß, dass P_1 seine Aktion beobachten kann.
- P_1 weiß, dass P_2 weiß, dass P_1 dessen Aktion beobachten kann.
- **Aktionen/Strategien**:
 - P_1 hat die Wahl zwischen $s_{11} :=$ „akzeptieren“ und $s_{12} :=$ „kämpfen“. Er agiert nur, wenn P_2 den Markt betritt.
 - P_2 hat die Wahl zwischen $s_{21} :=$ „expandieren“ und $s_{22} :=$ „nicht expandieren“.
- **Entscheidungspfade**:
 - $p_1 := s_{21}^1 \rightarrow s_{11}^2$
 - $p_2 := s_{21}^1 \rightarrow s_{12}^2$
 - $p_3 := s_{22}^1$
- **Auszahlungen**:
 - Wenn P_2 expandiert und P_1 dies akzeptiert, dann erhält P_1 die Auszahlung $u_1(p_1) = 1$ und P_2 die Auszahlung $u_2(p_1) = 2$.
 - Wenn P_2 expandiert und P_1 sich zum Preiskampf entschließt, dann erhält P_1 die Auszahlung $u_1(p_2) = 0$ und P_2 die Auszahlung $u_2(p_2) = 0$.
 - Wenn P_2 nicht expandiert, dann erhält P_1 die Auszahlung $u_1(p_3) = 2$ und P_2 die Auszahlung $u_2(p_3) = 1$.

2) *Lösungsansatz bei einem Spiel in Extensivform*: Gleichgewichte bei Spielen in extensiver Form können auf unterschiedliche Weisen gefunden werden. Beispielsweise können sie in induzierte Normalformspiele transformiert werden, indem die Mengen der reinen Strategien als Strategiemengen des induzierten Normalformspiels eingesetzt werden. Die Auszahlungsfunktionen dieses Normalformspiels können dann durch eine Zuordnung der Strategiekombinationen in den Entscheidungspfaden des Spielbaums mit den entsprechenden Auszahlungen in den Blättern gefunden werden.

Aus dem in Abbildung 1 illustrierten Spiel resultiert so die induzierte Auszahlungsmatrix in Tabelle II. Die Nash-Gleichgewichte dieses induzierten Normalformspiels sind $s_{N_1} = (s_{11}, s_{21})$ und $s_{N_2} = (s_{12}, s_{22})$. Das zweite Gleichgewicht beruht auf einer unglaublichen Drohung. Die Strategiekonfiguration (s_{12}, s_{22}) bedeutet, dass Unternehmen P_2 nicht expandiert (s_{22}) und Unternehmen P_1 kämpft (s_{12}). Dieser

Tabelle II. DARSTELLUNG DER AUSZAHLUNGSMATRIX DES MARKTEINTRITTSPIELS ALS INDUZIERTES NORMALFORM.

		P_2	
		s_{21}	s_{22}
P_1	s_{11}	1, 2	2, 1
	s_{12}	0, 0	2, 1

Pfad ist jedoch in der Spielbeschreibung nicht definiert.

Wie zu erkennen ist, weist die Lösung eines Extensivformspiels mittels Transformation in ein induziertes Normalformspiel Mängel auf. Es können Gleichgewichte gefunden werden, welche auf undefinierten Tatsachen beruhen. Eine genauere Lösungsmöglichkeit eines Spiels in extensiver Form ist die Rückwärtsinduktion nach von Neumann und Morgenstern [4], [3], [2]. Hier werden Gleichgewichte durch sukzessives Ersetzen von Knoten mit den daran hängenden Auszahlungen der Nash-Gleichgewichte in diesem Teilbaum gefunden. Beim Markteintrittsspiel resultiert nach dem ersten Induktionsschritt der Graph in Abbildung 2. Durch die Wahl der Strategie „akzeptieren“ hat P_1 eine höhere Auszahlung als durch die Wahl der Strategie „kämpfen“. Im Beispiel resultiert aus dem zweiten Induktionsschritt P_2 für P_2 die Strategie „expandieren“, da dieser so eine höhere Auszahlung erwarten kann. Somit ist das einzige Nash-Gleichgewicht der Entscheidungspfad p_1 .

Für tiefergehende Informationen zur Spieltheorie wie bspw. der Definition und Lösung von Teilspielen und dem Zusammenhang zu imperfekter Information in Teilspielen sei an dieser Stelle nochmals auf Literatur [4], [5], [3], [2] verwiesen.

C. Informationsstruktur

Von essentieller Bedeutung ist bei Spielen in extensiver Form der Umfang der Information, die einem Spielers bei seinem Zug zur Verfügung steht. Hierbei wird nach [5] zwischen perfekter, imperfekter, vollständiger und unvollständiger Information unterschieden.

Perfekte Information: Wenn einem Spieler im Spielverlauf alle vorausgehenden Züge seiner Mitspieler bekannt sind, dann hat er einen perfekten Informationsstand. Er weiß genau, in welchem Knoten des Spielbaums er sich befindet.

Imperfekte Information: Wenn ein Spieler im Spielverlauf nicht alle Züge seiner Mitspieler beobachten kann, dann hat er einen imperfekten Informationsstand. Kennt der Spieler die Spielstruktur und die Eigenschaften der Mitspieler, kann er nicht genau wissen, in welchem Knoten des Spielbaums er sich befindet.

Vollständige Information: Ein Spiel mit vollständiger Information liegt vor, wenn jeder Spieler über alle Eigenschaften der Mitspieler informiert ist und somit kein Spieler private

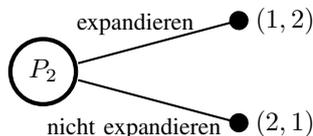


Abbildung 2. Graph des Markteintrittsspiels nach dem ersten Induktionsschritt der Rückwärtsinduktion nach von Neumann und Morgenstern.

Informationen besitzt. Jeder Spieler kennt somit die Auszahlungsfunktionen und Strategien der anderen Spieler.

Unvollständige Information: Ein Spiel mit unvollständiger Information liegt vor, wenn mindestens ein Spieler private Eigenschaften besitzt, welche die Mitspieler nicht kennen.

IV. SPIELTHEORIE BEI PATROUILLIERENDEN ROBOTERN

Nachdem im vorangegangenen Abschnitt die wesentlich relevanten Grundlagen der Spieltheorie vermittelt wurden, zielt dieser Abschnitt auf die Beschreibung des Spiels aus [1] zwischen dem patrouillierenden und einem eindringenden Roboter ab. Außerdem werden der von den Autoren gewählte Lösungsansatz und im darauf folgenden Abschnitt die Ergebnisse erläutert.

Für die Modellierung des Spiels treffen die Autoren von [1] folgende Annahmen:

- Die Zeit ist durch Runden diskretisiert.
- Es gibt n interessante Orte, die patrouilliert werden müssen.
- Der patrouillierende Roboter kann alle diese n Orte von jedem anderen Ort i aus erreichen ($1 \leq i \leq n$).
- Der patrouillierende Roboter ist mittels omnidirektionalem Sensor in der Lage, einen innerhalb des Sichtfeldes befindlichen Eindringling zu erkennen.
- Die interessanten Orte liegen soweit auseinander, dass der patrouillierende Roboter zu einem Zeitpunkt nur an genau einem Ort nach einem möglichen Eindringling Ausschau halten kann.
- Dem patrouillierenden Roboter können unterschiedliche Typen von Eindringlingen bekannt sein. Verschiedene Eindringlingstypen im Modell können als unterschiedliches Interesse an den Orten interpretiert werden.
- Die Umgebung kann sich während der Bewegung des patrouillierenden Roboters ändern.

Weiterhin gilt die folgende Spielbeschreibung:

- **Anzahl der Spieler:** Zwei (patrouillierender Spieler r und eindringender Spieler b).
- **Zugfolge:** Zu Beginn jeder Runde wählen die Spieler gleichzeitig ihre jeweilige Strategie für die Runde (unter Berücksichtigung von Regel 3). r wählt seine Strategie aus Σ_r und b wählt seine Strategie aus Σ_b .
- **Aktionen/Strategien:**
 - $\sigma_i^r \in \Sigma_r$: r wählt aus der Menge der n interessanten Orte einen Ort i aus, zu dem er als nächstes gehen möchte.
 - $\sigma_0^b \in \Sigma_b$: b dringt nicht ein und kann r beobachten.
 - $\sigma_i^b \in \Sigma_b$: b wählt aus der Menge der n interessanten Orte einen Ort i aus, zu dem er eindringen möchte (auch hier gilt $1 \leq i \leq n$); er kann r nicht beobachten.
- **Regeln:**
 - 1) Die Spieler können nicht kommunizieren.
 - 2) b benötigt d Runden für das Erreichen eines Ortes (wobei $1 \leq d \leq n$).
 - 3) Wählt b in der Runde \bar{t} die Strategie σ_i^b („nach Ort i eindringen“), so kann er bis zum Erreichen

des Ortes i in der Runde $\bar{t} + d$ keine andere Strategie wählen. Wählt er die Strategie σ_0^b („warten“), dann kann er (a) r beobachten und (b) in der darauf folgenden Runde $\bar{t} + 1$ eine andere Aktion aus Σ_b wählen.

- 4) r kann die Aktionen von b nicht beobachten, es sein denn, er befindet sich entweder zum Zeitpunkt \bar{t} an demselben Ort wie b oder er betritt während des Eindringens von b einen Ort, welchen b innerhalb der Runden $\bar{t} + d$ betreten hat.
- 5) Das Spiel endet, wenn
 - a) r b gefasst hat. r befindet sich dann entweder am selben Ort wie b oder an einem Ort, den b während des Eindringens—innerhalb der Runden $\bar{t} + d$ —betreten hat.
 - b) b den Zielort i erreicht hat, zu dem er eindringen wollte. r konnte ihn dann nicht fassen.

- **Auszahlungen:** Jeder Ort i hat für den patrouillierenden Spieler r den Wert v_i^r (wobei $v_i^r > 0$) und für den eindringenden Spieler b den Wert v_i^b (wobei $v_i^b > 0$). Wird b gefasst, erhält er zusätzlich die Auszahlung t (wobei $t < 0$) und r erhält zusätzlich die Auszahlung u (wobei $u > 0$). Weiterhin hat r die Kosten $C_{i,j}$ (bspw. als Energiekosten interpretierbar), um ausgehend vom Ort i zum nächsten Ort j zu gelangen ($C_{i,j} > 0$).
- **Informationsstruktur:** Spiel mit vollständiger, imperfekter Information.
 - Dem patrouillierenden Roboter sind alle möglichen Typen von Eindringlingen sowie deren Eigenschaften bekannt.
 - Der patrouillierende Roboter kennt nicht die gespielte Strategie des eindringenden Roboters.
 - Der patrouillierende Roboter ist in der Lage, jede Änderung der Spielumgebung zu registrieren.

Der Gesamtnutzen (Summe der Einzelauszahlungen) des patrouillierenden Spielers berechnet sich aus der Summe der Werte v_i^r der beschützten Orte und u , falls der eindringende Spieler gefasst wurde, abzüglich der Summe der benötigten Kosten $C_{i,j}$. Der Gesamtnutzen des eindringenden Spielers berechnet sich aus der Summe der Werte v_i^b der besuchten Orte, falls er nicht gefasst wird. Andernfalls ist der Nutzen t . Eindringlingstypen unterscheiden sich in verschiedenen Werten v_i^b für einen Ort i .

Amigoni et al. verfolgen in der Arbeit den Ansatz, das Extensivformspiel in Slices einzuteilen. Ein einzelner Slice ist eine Runde, in der sowohl der patrouillierende als auch der eindringende Roboter eine Aktion wählen und durchführen. Genauer betrachtet ist jeder Slice ein Normalformspiel mit zwei Spielern. All diese Normalformspiele in den Slices beinhalten dieselben Strategiemöglichkeiten. Sie unterscheiden sich lediglich in den Auszahlungen. Tabelle III zeigt die Auszahlungsmatrix eines Slices in einer Umgebung mit drei interessanten Orten (1, 2 und 3). Der patrouillierende Roboter befindet sich am Ort l .

Die Lösung eines Slices ist eine gemischte Strategie, mit welcher der patrouillierende Roboter seinen erwarteten Nutzen maximiert. Die Lösung der Patrouillierungsstrategie für das

Tabelle III. BEISPIEL FÜR EINE AUSZAHLUNGSMATRIX DES APPROXIMIERTEN NORMALFORMSPIELS EINES SLICES.

		r		
		σ_1^r	σ_2^r	σ_3^r
b	σ_0^b	$0, v_1^r - C_{l,1}$	$0, v_2^r - C_{l,2}$	$0, v_3^r - C_{l,3}$
	σ_1^b	$t, v_1^r + u - C_{l,1}$	$v_1^b, v_2^r - C_{l,2}$	$v_1^b, v_3^r - C_{l,3}$
	σ_2^b	$v_2^b, v_1^r - C_{l,1}$	$t, v_2^r + u - C_{l,2}$	$v_2^b, v_3^r - C_{l,3}$
	σ_3^b	$v_3^b, v_1^r + u - C_{l,1}$	$v_3^b, v_2^r - C_{l,2}$	$t, v_3^r + u - C_{l,3}$

gesamte Spiel wird somit durch die Sequenz der Lösungsstrategien der einzelnen Normalformspiele approximiert.

Für die Suche nach einer Lösungsstrategie innerhalb eines Slices formulieren die Autoren von [1] zwei mathematische Probleme, welche sich mit Algorithmen für lineare mathematische Optimierungsprobleme lösen lassen. Das eine als multilineares (multi-linear programming – multi-LP) und das andere als ganzzahliges lineares Optimierungsproblem (mixed integer linear programming – MILP) formuliert. Aufgrund der Komplexität und der Aussagekraft für diese Arbeit wird an dieser Stelle für eine allgemeine Erläuterung solcher Probleme auf [6], [7] sowie für die detaillierte Formulierung auf [7], [8], [1] verwiesen.

Die Resultate der Algorithmen sind äquivalent. Sie liefern beide eine Menge γ_i (mit $i = 1, 2, \dots, n$), wobei $\gamma_i \in [0, 1]$ die Wahrscheinlichkeit angibt, mit welcher sich der patrouillierende Roboter entscheidet, den Ort i zu besuchen ([1, S. 502], frei übersetzt).

V. ERGEBNISSE VON AMIGONI ET AL.

Die Autoren von [1] evaluieren ihren Lösungsansatz unter folgenden Gesichtspunkten:

- 1) Laufzeit der Berechnung einer Lösung für einen Slice.
- 2) Performanz der Berechnung bzgl. Abdeckung der patrouillierten Orte, Unvorhersehbarkeit der Patrouillierungsstrategie und Durchschnittsnutzen des patrouillierenden Roboters bei einem Slice in unterschiedlichen Umgebungen (linear, ring- und sternförmig)
- 3) Umgang mit dynamischen Umgebungsänderungen

Aufgrund des begrenzten Umfangs wird für eine ausführliche Beschreibung der Evaluierung auf die zugrunde liegende Arbeit verwiesen.

Die Ergebnisse in Tabelle IV zeigen, dass bei der Evaluierung zu Punkt 1 die Laufzeit der Berechnung mit zunehmender Anzahl der interessanten Orte n zunimmt. Die multi-LP-Formulierung skaliert hierbei deutlich besser. Bei gleichbleibender Anzahl der interessanten Orte und Erweiterung der Anzahl möglicher Eindringlingstypen skaliert die MILP-Formulierung wesentlich besser.

Die Ergebnisse in Tabelle V zeigen, dass bei der Evaluierung zu Punkt 2 der Durchschnittsnutzen größer ist als mit den verglichenen Lösungsansätzen, unabhängig von der Umgebung. Die Abdeckung der patrouillierten Orte und die Unvorhersehbarkeit der Strategie sind akzeptabel. Verglichen wird (a) mit dem Ansatz, jede Strategie mit der Wahrscheinlichkeit $1/n$ zu wählen, (b) die Wahrscheinlichkeit proportional zum festgelegten Wert des Ortes abzüglich der benötigten

Tabelle IV. LAUFZEIT DER BERECHNUNG EINER LÖSUNG FÜR EINEN SLICE (IN S) [1].

multi-LP					
n (# Orte)	1 Typ	2 Typen	3 Typen	4 Typen	5 Typen
3	0,068	0,342	1,278	6,854	62,972
5	0,079	0,796	13,227	544,481	-
10	0,280	16,095	-	-	-
15	0,509	146,665	-	-	-
25	1,085	-	-	-	-
50	3,702	-	-	-	-
100	16,096	-	-	-	-

MILP					
n (# Orte)	1 Typ	2 Typen	3 Typen	4 Typen	5 Typen
3	0,006	0,015	0,031	0,068	0,138
5	0,014	0,033	0,103	0,357	1,128
10	0,029	0,223	1,615	4,320	-
15	0,074	1,351	8,128	-	-
20	0,280	4,706	14,834	-	-
25	0,752	12,955	-	-	-
50	11,220	-	-	-	-
100	215,209	-	-	-	-

Tabelle V. PERFORMANZ DER PATROUILLIERUNGSSTRATEGIE [1].

Lineare Umgebung			
Methode	# nie besuchter Orte	Entropie	Nutzen
multi-LP	15,4	2,081	78,95
(a)	18,7	3,219	52,124
(b)	12,2	3,014	69,528
(c)	0,5	3,027	48,028

Ringförmige Umgebung			
Methode	# nie besuchter Orte	Entropie	Nutzen
multi-LP	15,4	2,117	82,911
(a)	18,1	3,219	56,149
(b)	10,6	3,019	66,450
(c)	1,3	3,033	50,494

Sternförmige Umgebung			
Methode	# nie besuchter Orte	Entropie	Nutzen
multi-LP	14,9	2,169	75,014
(a)	19	3,219	34,831
(b)	16,4	3,007	55,714
(c)	1,2	3,031	40,444

Kosten für die Erreichung dieses Ortes zu wählen und (c) die Wahrscheinlichkeit für jeden Ort zufällig zu wählen.

Bei der Evaluierung zu Punkt 3 wird gezeigt, dass der Lösungsansatz gut auf Änderungen der Umgebung während des Spiels reagieren kann. Hierbei wurde in einer linearen Umgebung mit drei interessanten Orten die Bewegungskosten $C_{l,1}$ und $C_{l,2}$ nach hundert Schritten unterschiedlich stark erhöht. Wie Tabelle VI zeigt, wird vor der Änderung Ort 1 aufgrund des höchsten Wertes am häufigsten besucht (mit Bewegungskosten $C_{l,1} = C_{l,2} = C_{l,3} = 1$). Nach der Änderung der Kosten zu den Orten 1 und 2 wird Ort 2 am häufigsten besucht. Ort 1 wird aber weiterhin oft besucht, da der festgelegte Wert nach wie vor hoch ist im Vergleich zu den anderen.

Tabelle VI. REAKTION AUF UMGEBUNGSÄNDERUNG [1].

Ort i	v_i^*	# Besuche vor Änderung	# Besuche nach Änderung
1	90	47,9	171,2
2	9	8,5	0
3	30	43,6	228,8

VI. FAZIT

Amigoni et al. haben in ihrer Arbeit [1] einen Ansatz zur Bestimmung einer Patrouillierungsstrategie für einen mobilen Roboter gezeigt, der als Aufgabe die Überwachung von Orten mit unterschiedlichem Wert hat. Sie haben gezeigt, dass ihr Ansatz im Vergleich mit anderen Ansätzen effizienter ist. Ihr Modell basiert auf einem Spiel in extensiver Form, welches durch die Einteilung in Slices und wiederholte Optimierung des Normalformspiels in diesen Slices näherungsweise gelöst werden kann. Die vorliegende Seminararbeit hat spieltheoretische Grundlagen vermittelt, welche zum Verständnis der Arbeit von Amigoni et al. beitragen. Es wurde der Unterschied zwischen Spielen in Normalform und in extensiver Form betrachtet und unterschiedliche Lösungen erläutert. Weiterhin wurden unterschiedliche Informationsstrukturen erläutert.

Basierend auf der Annahme, dass sich die Umgebung während des Spiels dynamisch ändern kann, war zu erwarten, dass der schrittweise Lösungsansatz der Autoren einen größeren Nutzen hat als die verglichenen Ansätze. Amigoni et al. schlussfolgern aus dem größeren Nutzen die Effizienz ihres Modells. Da sie keine Annahmen über die Dauer einer einzelnen Runde machen, ist diese Aussage korrekt. Sobald jedoch eine solche Rahmenbedingung definiert wird, muss die Effizienz des Modells auch in Hinblick auf die Laufzeit der Berechnung überprüft werden. Ein Vergleich der Berechnungsdauer zwischen ihrem Ansatz und den Methoden bei den Performanzvergleichen wäre hierfür interessant gewesen.

Eine Möglichkeit zur Beschleunigung der Berechnung könnte die Methode der dynamischen Programmierung nach Bellman [9] sein. Hierbei wird ein Optimierungsproblem durch optimales Lösen von Teilproblemen gelöst.

LITERATUR

- [1] F. Amigoni, N. Gatti, und A. Ippedito, „A Game-Theoretic Approach to Determining Efficient Patrolling Strategies for Mobile Robots,” in *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 02*, Serie WI-IAT '08. Washington, DC, USA: IEEE Computer Society, 2008, S. 500–503. [Online]. <http://dx.doi.org/10.1109/WIIAT.2008.324>
- [2] S. K. Berninghaus, K.-M. Ehrhart, W. Güth, S. K. Berninghaus, K.-M. Ehrhart, und W. Güth, *Strategische Spiele*, Serie Springer-Lehrbuch. Springer Berlin Heidelberg, 2010. [Online]. <http://dx.doi.org/10.1007/978-3-642-11651-3>
- [3] M. J. Osborne, *Introduction to Game Theory: International Edition*. Oxford University Press, 2009.
- [4] D. Fudenberg und J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.
- [5] A. Wellenreuther, „Spieltheorie,” Vorlesungsskript, 2008.
- [6] V. Conitzer und T. Sandholm, „Computing the optimal strategy to commit to,” in *Proceedings of the 7th ACM conference on Electronic commerce*, Serie EC '06. New York, NY, USA: ACM, 2006, S. 82–90. [Online]. <http://doi.acm.org/10.1145/1134707.1134717>
- [7] P. Paruchuri, J. P. Pearce, M. Tambe, F. Ordonez, und S. Kraus, „An efficient heuristic approach for security against multiple adversaries,” in *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, Serie AAMAS '07. New York, NY, USA: ACM, 2007, S. 181:1–181:8. [Online]. <http://doi.acm.org/10.1145/1329125.1329344>
- [8] P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordonez, und S. Kraus, „Playing games for security: an efficient exact algorithm for solving Bayesian Stackelberg games,” in *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 2*, Serie AAMAS '08. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2008, S. 895–902. [Online]. <http://dl.acm.org/citation.cfm?id=1402298.1402348>
- [9] R. Bellman, *Dynamic Programming*. Princeton, New Jersey: Princeton University Press, 1957.