



Patrullering med poker-AI

Systematiskt oförutsägbart patrullering i sjö, mark och cyberrymd

JOEL BRYNIELSSON, MIKA COHEN,
FARZAD KAMRANI, DANIEL OSKARSSON



Joel Brynielsson, Mika Cohen, Farzad Kamrani,
Daniel Oskarsson

Patrullering med poker-AI

Systematiskt oförutsägbar patrullering i sjö, mark och cyberrymd

Titel	Patrullering med poker-AI: Systematiskt oförutsägbart patrullering i sjö, mark och cyberspace
Title	Patrolling with poker AI: Systematically unpredictable patrolling in sea, land and cyberspace
Rapportnr/Report no	FOI-R--5291--SE
Månad/Month	Januari
Utgivningsår/Year	2022
Antal sidor/Pages	17
ISSN	1650-1942
Uppdragsgivare/Client	FOI
Forskningsområde	Ledningsteknologi
Projektnr/Project no	A81206
Godkänd av/Approved by	Lars Höstbeck
Ansvarig avdelning	Försvarteknik

Bild/Cover: ADENVIKEN 2009-06-26. HMS Malmö (i bakgrunden till höger) har funnit en drivande skiff i Adenviken utanför den somaliska staden Bosaasoo. Vid sök ombord på skiffen fann man bland annat raketdrivna granater (RPG), finkalibrig ammunition, knivar och en stege – materiel som antyder på eventuell piratverksamhet. Foto: Mats Nyström/Combat Camera/Försvarsmakten.

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk, vilket bl.a. innebär att citering är tillåten i enlighet med vad som anges i 22 § i nämnd lag. För att använda verket på ett sätt som inte medges direkt av svensk lag krävs särskild överenskommelse.

This work is protected by the Swedish Act on Copyright in Literary and Artistic Works (1960:729). Citation is permitted in accordance with article 22 in said act. Any form of use that goes beyond what is permitted by Swedish copyright law, requires the written permission of FOI.

Sammanfattning

I denna rapport beskrivs operativa AI-baserade beslutsstödsystem som används för att planera systematiskt oförutsägbar patrullering inom olika arenor, systemens underliggande AI-metoder, och experiment som bedrivits med dessa metoder på FOI under 2021.

Nyckelord: Patrullering, Övervakning och spaning, Oförutsägbarhet, Beslutsstöd, AI, Planering under osäkerhet

Summary

This report describes operational AI-based decision support systems used for systematically unpredictable patrolling in different arenas, the systems' underlying AI methods, and experiments with these methods performed at FOI during 2021.

Keywords: Patrolling, Surveillance and reconnaissance, Unpredictability, Decision support, AI, Planning under uncertainty

Innehållsförteckning

1	Inledning	7
2	Oförutsägbar patrullering	9
3	Spelintelligens.....	11
4	Pokerintelligens	13
5	Pokerintelligens på FOI	15
6	Oförutsägbar patrullering med pokerintelligens.....	17

1 Inledning

En viktig del av krigföring är att göra det oväntade. Det fångas i begreppet överraskning.

Militärstrategisk doktrin, 2016

En kustbevakning kan omöjligt övervaka alla områden oavbrutet. Havs- och kustområdena som ska övervakas är för stora och resurserna i termer av personal, utrustning och fartyg alltför begränsade. Inkräktare såsom smugglare, tjuvfiskare och främmande makt kan därför undvika (eller i vart fall minska risken för) upptäckt genom att observera och exploatera mönster i patrulleringen.

Normalt planerar en kustbevakning sin patrullering utifrån erfarenhet och intuition, det vill säga kvalificerade gissningar, om hur inkräktare bäst överraskas. Men sedan snart ett decennium använder kustbevakning i USA ett beslutsstödsystem, PROTECT, för att planera optimalt oförutsägbar patrullering.¹



PROTECT planerar patrullering åt kustbevakningen i USA sedan snart ett decennium. Bild: USCGC Osprey patrols the Puget Sound, Wikimedia Commons, Public domain, Creative Commons Attribution 2.0. Foto: Robert K. Lanier.

Ett flertal liknande beslutsstöd för att planera oförutsägbar patrullering är numera i drift runt om i världen.² De räknas som en av de mest lyckade tillämpningarna av AI under de senaste åren.³

¹ PROTECT – A deployed game theoretic system for strategic security allocation for the United States Coast Guard, Bo An et al., AI Magazine, 2013.

² Stackelberg security games: Looking beyond a decade of success, Arunesh Sinha et al., IJCAI, 2018.

³ Exploring the feasibility and utility of machine learning-assisted command and control: Volume 1, Findings and recommendations, s. 9, RAND, 2021.

I denna rapport diskuteras operativa planeringsstöd för patrullering, underliggande AI-metoder och försök med dessa metoder som bedrivits på FOI under 2021. Rapporten avslutas med förslag på tillämpningar inom spaning och övervakning.

2 Oförutsägbar patrullering

patrol – A detachment of ground, sea, or air forces sent out for the purpose of gathering information or carrying out a destructive, harassing, mopping-up, or security mission.

DoD Dictionary of Military and Associated Terms, 2010

Patrullering är en vanlig uppgift inom försvar och säkerhet. Ubåtar patrullerar hav. Patrullfartyg patrullerar havs- och kustområden. Fotpatruller patrullerar avgränsade markområden. Cybersoldater patrullerar cyberrymden.

Patrulleringens uppgift är att ”verka med små patruller över stora ytor för att tidigt upptäcka och kunna motverka säkerhetshotande verksamhet”.⁴ För att maximera den avskräckande effekten måste allokering, schemaläggning och rutter undvika att vara förutsägbara på ett sätt som kan exploateras av inkräktare. Patrulleringspersonal kan ibland beskriva ett behov av att vara irrationell och ologisk och av att gå på intuition.

Spontant kan det tyckas som att det ”intuitiva” och ”irrationella” beslutsfattandet inte kan vara något för maskiner. Men sedan en tid tillbaka planeras patruller inom olika arenor med beslutsstöd som beräknar optimalt oförutsägbar patrullering.⁵ Beslutsstödsystemet PROTECT, som varit operativt inom den amerikanska kustbevakningen i snart ett decennium, är ett tidigt exempel.⁶ Ett nyare exempel är PAWS, som används för att planera fotpatruller i naturreservat i bl.a. Uganda och nu är på väg att driftsättas på över 800 platser.⁷ Liknande beslutsstöd används i dag runt om i världen för att planera oförutsägbar patrullering inom olika arenor – allt från flygplatser och flygtrafik, hav och vildmark till kollektivtrafik och IT-nätverk.⁸

⁴ <https://www.forsvarsmakten.se/sv/organisation/livgardet/insatsforband/13e-sakerhetsbataljonen>.

⁵ *Exploring the feasibility and utility of machine learning-assisted command and control: Volume 1, Findings and recommendations*, RAND, 2021.

⁶ *PROTECT – A deployed game theoretic system for strategic security allocation for the United States Coast Guard*, Bo An et al., AI Magazine, 2013.

⁷ *PAWS – A deployed game-theoretic application to combat poaching*, Fei Fang et al., AI Magazine, 2017.

⁸ *Stackelberg security games: Looking beyond a decade of success*, Arunesh Sinha et al., IJCAI, 2018.



Beslutsstödet PAWS planerar sedan några år fotpatruller i naturreservat i Uganda, Malaysia och Colombia för att skydda mot tjuvjakt. Bild: Anti-Poaching Unit out on patrol, Public domain. Wikimedia Commons, Creative Commons Attribution 2.0. Foto: Bumi Hills Foundation.

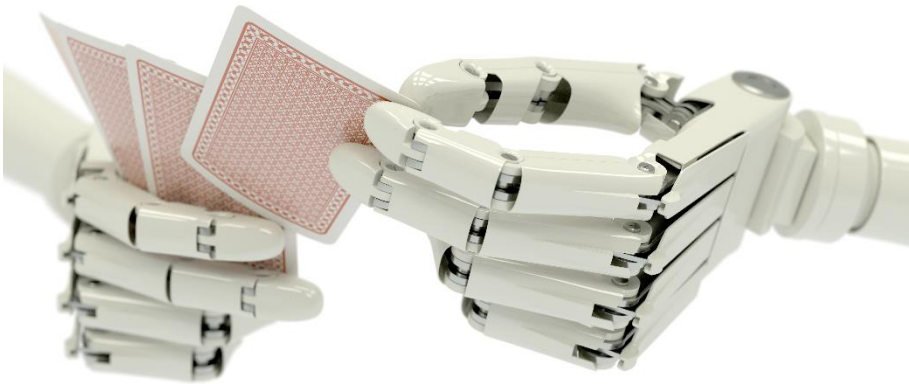
För att göra patrulleringen oförutsägbar – och åstadkomma en maximalt avskräckande överraskningseffekt – tillför beslutsstöden ett element av slumpmässighet i patrulleringen. Men patrullerna agerar inte helt godtyckligt utan patrulleringsbeslut slumpas på ett sätt som maximerar den avskräckande effekten. Med andra ord slumpas beslut på så sätt att fördelen för en inkräktare (t.ex. tjuvjägare) med ett visst angrepp (t.ex. angrepp på område med större densitet vilddjur) motverkas av en risk för upptäckt som står i proportion till det eventuella byte som angreppet skulle kunna resultera i. Därmed uppstår inga exploaterbara mönster i patrulleringen som inkräktaren kan dra nytta av och planera utifrån.

3 Spelintelligens

The analysis of a more realistic poker game than our very simple model should be quite an interesting affair.

John Forbes Nash, 1951

PROTECT, PAWS och de andra patrulleringsstöden ovan räknas som en av de mest lyckade tillämpningarna av AI under de senaste åren.⁹ Patrullerna planeras med *algoritmisk spelteori*, mer populärt *spelintelligens*, som är ett delområde inom AI som överlappar med (delområdena) maskininlärning och multiagentsystem.



Mycket av forskningen inom algoritmisk spelteori rör spel med privat information. Målet har länge varit att bemästra poker. Bild: Shutterstock, id: 148169480.

Inom algoritmisk spelteori utvecklas artificiella spelare (agenter) som interagerar optimalt med varandra i flerpersonerspel (datorspel för två eller fler spelare). I tillämpningar (bortom underhållningsspel) är speltillståndet ("spelbrädet") ofta delvis eller helt dolt för spelarna, såsom i patrulleringsfallet ovan, vilket ger spelarna skilda lägesbilder (s.k. privat information).¹⁰ Mycket av forskningen

⁹ *Exploring the feasibility and utility of machine learning-assisted command and control: Volume 1, Findings and recommendations*, s. 9, RAND, 2021.

¹⁰ Jämför krigsspel: ända sedan krigsspelen uppstod på 1800-talet som en form av schack med skärm mellan spelarna har delvis dolda spelbräden ("krigsdimma", "osäkerhet") setts som fundamentalt inom krigsspel (*Wargaming Handbook*, UK Ministry of Defence, 2017).

inom algoritmisk spelteori rör därför spel med ett delvis dolt speltillstånd.¹¹ De skilda lägesbilderna innebär att en spelare delvis måste slumpa sina drag för att inte signalera för mycket av sin lägesbild och sin handlingsplan till sin motspelare; en pokerspelares bud bör t.ex. inte allt för entydigt spegla spelarens kort. Om en spelare alltid agerar på samma sätt (givet sin lägesbild) kan motspelarna härleda spelarens lägesbild ur spelarens agerande och därmed även härleda hur spelaren kommer att agera härnäst.

Ända sedan Nash m.fl. initierade spelteorin på 1950-talet har just poker – eller närmare bestämt, en agent som överträffar människor i poker – utgjort en slags helig gral. Målet har verkat avlägset, om ens överhuvudtaget nåbart. Algoritmer har endast löst poker i extremt förenklad form med kortleken reducerad till ett fåtal kort och det maximala antalet spelronder reducerat till bara några få budgivningsronder. Mänsklig pokerlist har förefallit ligga bortom algoritmers förmåga.

¹¹ *The state of solving large incomplete-information games, and application to poker*, Tuomas Sandholm, AI Magazine, s. 13, december 2010.

4 Pokerintelligens

And that's why there's never going to be a computer that can play world-class poker. It's a people game.

Doyle Brunson, Super/System, 1979

För fem år sedan förlorade Lee Sedol, legendarisk 18-faldig världsmästare i Go, mot AlphaGo, en AI från DeepMind.¹² Händelsen väckte stor uppmärksamhet – att bemästra Go sågs av många som det yttersta inom mänsklig intelligens och kreativitet. Pentagon skapade raskt en ny enhet, Defence Innovation Unit, med syftet att nyttja den nya formen av AI.



*AlphaGo överraskade en hel värld med att uppnå övermänsklig spelstyrka i Go.
Bild: AlphaGo the movie (2017), Moxie Pictures, Press kit.*

Go är ett spel med publikt spelbräde, men i tillämpningar av algoritmisk spelteori är speltillståndet ofta delvis dolt. Skulle det vara möjligt att uppnå övermänsklig spelstyrka även i spel med delvis dolda speltillstånd, såsom poker? Många var fortfarande skeptiska, AlphaGo till trots. Det psykologiska och listiga tänkandet i poker kring bluff, vilseledning och överraskning kan förefalla mer avlägset för algoritmer än det mer mekaniska tänkandet i Go om möjliga sekvenser av drag och motdrag.

¹² *Mastering the game of Go with deep neural networks and tree search*, David Silver et al., Nature, januari 2016.

Men bara några år senare, 2019, uppnådde en AI från Facebook, Pluribus, övermännisklig spelstyrka i Texas hold'em för sex spelare, den mest utmanande tävlingsformen av poker.¹³ Poker kunde därmed anses löst.

DeepMinds AlphaGo och Facebooks Pluribus räknas som två historiska milstolpar inom AI – hittills de enda två mottagarna av det prestigefyllda priset *Marvin Minsky Medal for Outstanding Achievements in AI*. De lär sig spela Go respektive poker genom en likartad process av försök-och-misstag i massivt simulerade spel: AI:n spelar kontinuerligt mot sig själv – i praktiken parallellt på många datorkärnor samtidigt – varvid den oavbrutet uppdaterar sin värdering av drag utifrån det observerade utfallet. Värderingen, och därmed spelstyrkan, blir successivt allt bättre, för att till slut – efter många partier – bli övermännisklig.

DeepMind förenklade och generaliserade sedermera AlphaGo till en generell algoritm, AlphaZero, som utan annan input än själva spelreglerna lär sig ett godtyckligt spel med öppet spelbräde,¹⁴ en vidareutveckling som i sig räknas som ännu en milstolpe inom AI.¹⁵ I slutet av 2020 kombinerade Facebook AlphaZero med Pluribus till en allmän algoritm för spel med delvis slutet spelbräde (privat information).¹⁶ Den nya algoritmen, ReBeL, utgör i dag den kanske kraftfullaste formen av AI för poker.

Facebook publicerade en implementation av ReBeL för Liars Dice, en enklare form av poker med tärningar i stället för spelkort.¹⁷ Men för att hindra otillåten användning i onlinepoker har Facebook inte offentliggjort implementationen för poker.¹⁸ Att överföra den publicerade implementationen från Liars Dice till poker anses icke-trivialt.¹⁹

¹³ *Superhuman AI for multiplayer poker*, Noam Brown et al., Science, september 2019.

¹⁴ *A general reinforcement learning algorithm that masters chess, shogi and Go through self-play*, David Silver et al., Science, december 2018.

¹⁵ *Exploring the feasibility and utility of machine learning-assisted command and control: Volume 1, Findings and recommendations*, s. 8, RAND, 2021.

¹⁶ *Combining deep reinforcement learning and search for imperfect-information games*, Noam Brown et al., NEURIPS, 2020.

¹⁷ <https://github.com/facebookresearch/rebel>.

¹⁸ ”The most immediate risk posed by this work is its potential for cheating in recreational games such as poker ... Partly for this reason, we have decided not to release the code for poker.”, *ibid*.

¹⁹ ”However, the chances of anyone using ReBel to play online poker are bleak as the researchers have taken precautions against it.”, <https://pokerguru.in/poker-news/facebook-artificial-intelligence-rebel-ai-poker-300720/>.

5 Pokerintelligens på FOI

In chess, the heart of the game is control of the center. In the Game of the Generals, it is control of your opponent's mind.

Sofronio H. Pasola Jr.

Under 2021 har FOI överfört ReBeL till två klassiska krigsspel, *Stratego* och *Game of the Generals*.²⁰ Detta utgör den första rapporterade reproduktionen av ReBeL utöver Facebooks egen utveckling. Det utgör även den första användningen av ReBeL för annat än poker och (pokerlika) Liars Dice.



Stratego och Game of the Generals kan ses som en kombination av schack och poker. Liksom i schack manövrerar spelarna enheter över ett rutnät; när en enhet beger sig in på en ruta med en motståndarenhet slås endera pjäsen ut enligt en given utslagstabell. Men liksom i poker ser spelaren inte identiteten (valören) på motståndarens enheter. Bild: Shutterstock, id: 1822592534.

²⁰ *AI for an imperfect-information wargame with self-play reinforcement learning*, Filip Ryblad, Examensarbete, KTH, 2021 och *Stratego using deep reinforcement learning and search*, Anton Falk, Examensarbete, KTH, 2021.

Stratego och Game of the Generals tillhör en genre av liknande blockspel (eng.: block wargames).²¹ Så här beskrivs Game of the Generals av spelutvecklaren:²²

Spelet simulerar arméer som försöker övermanna, vilseleda, överflygla, utmanövrera och förstöra varandra ... Spelet simulerar 'krigsdimma' eftersom identiteterna på motspelarens pjäser är dolda och bara kan gissas utifrån deras position och rörelse eller från utfall i sammandrabbningar.

Stratego och Game of the Generals valdes som tillämpningsfall av flera skäl. Spelen är lättillgängliga och relativt populära, men det saknas ändå AI:n som når bortom nybörjarnivå. Spelen är därtill lätta att implementera, men samtidigt komplexa nog att rymma t.ex. patrullering av olika slag som taktiska delproblem.

På grund av begränsade beräkningsresurser har ReBeL-agenten hittills endast fått lära sig Stratego och Game of the Generals på mindre spelbräden med 5×5 rutor. Redan med dessa mindre spelbräden utgör dock spelen en intellektuell utmaning med ofta långa partier med 50–100 drag och fler möjliga speltillstånd (beslutssituationer) än exempelvis i Go.²³ Under 2022 planeras träningen flytta till en kraftfull beräkningsmaskin med förhoppning om att ReBeL ska kunna lära sig bemästra 8×8-bräden, vilket är en vanlig brädstorlek då människor tävlar i Stratego (s.k. "Stratego Barrage").

²¹ https://en.wikipedia.org/wiki/Block_wargame. En väsentlig skillnad mellan spelen är graden av krigsdimma. I Stratego röjs pjäsernas identitet vid en sammandrabbning: när två spelbrickor drabbar samman i en ruta vänder spelarna på dem (så att identiteterna blir publika) innan spelarna gemensamt dömer av striden. I Game of the Generals, å andra sidan, avdöms sammandrabbningen av en neutral domare (t.ex. spelmotorn i digitala versioner) utan att spelbrickornas identitet röjs (annat än indirekt via spelarnas slutledningsförmåga).

²² https://en.wikipedia.org/wiki/Game_of_the_Generals.

²³ I Go utgörs en beslutssituation av det aktuella spelbrädet. Men i Stratego och Game of the Generals utgörs en beslutssituation av hela sekvensen av observationer som spelaren gjort från spelstart fram tills nu (på grund av den privata informationen), vilket medför att antalet beslutssituationer växer exponentiellt med antalet drag i spelet. Även på små bräden som 5×5 pågår Game of the Generals ofta i många drag (50–100) om spelarna är någorlunda skickliga.

6 Oförutsägbar patrullering med pokerintelligens

Maybe most of, let's say, military planning, or most of business strategy, will actually be done strategically using computational game theory, that's what I would like to see as a next five or ten year goal.

Tuomas Sandholm, 2019

Pluribus, ReBeL och annan övermännisklig poker-AI från Facebook tillämpas nu inom Pentagon.²⁴ Arbetet leds av Tuomas Sandholm, en förgrundsgestalt inom algoritmisk spelteori som initierat ett antal framgångsrika tillämpningar, bl.a. patrullplanering. Detaljerna i tillämpningen är hemliga. Men en tillämpning inom patrullering ligger nära till hands. Systematiskt oförutsägbar patrullering är numera en beprövad och uppskattad tillämpning av algoritmisk spelteori.

Algoritmisk spelteori har internationellt visat sig ge praktisk nytta och fungera väl som beslutsstödskomponent i människa-maskinsystem som patrullerar olika arenor – allt från flygplatser och flygtrafik, hamnar, hav och vildmark till kollektivtrafik och it-nätverk. Även i Sverige bör algoritmisk spelteori såsom poker-AI kunna resultera i bättre beslutsstöd inom försvar och säkerhet.



Patrullering i olika former är en vanlig uppgift inom Försvarsmakten. Bild: HMS Visby/Försvarsmakten.

²⁴ <https://www.aalto.fi/en/news/tuomas-sandholm-applies-game-theory-in-his-companies-and-encourages-everyone-to-study-ai>.

FOI är en huvudsakligen uppdragsfinansierad myndighet under Försvarsdepartementet. Kärnverksamheten är forskning, metod- och teknikutveckling till nytta för försvar och säkerhet. Organisationen har cirka 1000 anställda varav ungefär 800 är forskare. Detta gör organisationen till Sveriges största forskningsinstitut. FOI ger kunderna tillgång till ledande expertis inom ett stort antal tillämpningsområden såsom säkerhetspolitiska studier och analyser inom försvar och säkerhet, bedömning av olika typer av hot, system för ledning och hantering av kriser, skydd mot och hantering av farliga ämnen, IT-säkerhet och nya sensorers möjligheter.



FOI
Totalförsvarets forskningsinstitut
164 90 Stockholm

Tel: 08-55 50 30 00
Fax: 08-55 50 31 00

www.foi.se