



Beslutsstöd för oförutsägbar aktiv spaning

Sökmönster och positionsskattning för undervattensövervakning med aktiv sonar

Joel Brynielsson, Mika Cohen, Christoffer Limér,
Felix Rydell

Joel Brynielsson, Mika Cohen, Christoffer Limér,
Felix Rydell

Beslutsstöd för oförutsägbar aktiv spaning

Sökmönster och positionsskattning för undervattensövervakning
med aktiv sonar

Titel	Beslutsstöd för oförutsägbar aktiv spaning – Sökmönster och positionsskattning för undervattensövervakning med aktiv sonar
Title	Decision support for unpredictable active surveillance – Search strategies and position estimation for underwater surveillance with active sonar
Rapportnr/Report no	FOI-R--5897--SE
Månad/Month	December
Utgivningsår/Year	2025
Antal sidor/Pages	35
ISSN	1650-1942
Uppdragsgivare/Client	Försvarsmakten
Forskningsområde	Ledningsteknologi
FoT-område	Ledning och MSI
Projektnr/Project no	E38567
Godkänd av/Approved by	Emil Hjalmarson
Ansvarig avdelning	Cyberförsvar och ledningsteknik

Bild/Cover: Jimmy Croona/Försvarsmakten. Spaning med mörkerkikare kring Jungfrufjärden. Försvarsmakten genomför underrättelseoperation i Stockholms skärgård mot misstänkt utländsk undervattenverksamhet.

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 rapporten undersöks i vilken utsträckning ett taktiskt beslutsstöd kan användas för att planera undervattensspaning utan att spaningen blir förutsägbar. Frågan har studerats med en datorprototyp som beräknar sökmönster för aktiva sonarer med hjälp av maskininlärning hämtad från pokerbottar.

Prototypens oförutsägbarhet har utvärderats teoretiskt och empiriskt. Resultatet av studien visar att det inte är möjligt för motståndare, som studerar spaningsinsatser, att förutse och exploatera prototypens förslag på sökmönster.

Slutsatsen av studien är att beslutsstöd baserat på AI för pokerliknande spel kan generera oförutsägbara sökmönster för aktiv spaning.

Nyckelord: beslutsstöd, artificiell intelligens, ubåtsjakt, aktiv spaning, sökmönster, oförutsägbarhet.

Summary

The report examines to what extent a tactical decision support system can be used for planning underwater surveillance without making the surveillance predictable. The issue has been explored using a computer prototype that generates search patterns for active sonars, with the help of machine learning techniques derived from poker bots.

The unpredictability of the prototype has been studied both theoretically and empirically. The results of the study indicate that an adversary, that studies surveillance activity over time, cannot predict and exploit the search patterns suggested by the prototype.

The conclusion from the study is that decision support based on AI for pokerlike games is able to generate unpredictable search patterns for active sonars.

Keywords: decision support, artificial intelligence, anti-submarine warfare, active surveillance, search pattern, unpredictability.

Innehållsförteckning

1	Inledning	7
2	Beslutsproblem	10
	Scenario	11
	Informationsunderläge	12
	Oförutsägbarhet	13
3	Beslutsstöd	14
	Sökmönster	15
	Positionsskattning	20
	Undervattensspaning som poker	24
4	Experimentell verifiering	28
	Spelteoretisk exploaterbarhet	29
	Empirisk exploaterbarhet	30
5	Slutsatser	32
6	Referenser	33

1 Inledning

Det finns tre viktiga faktorer i ubåtsjakt: överraskning, snabbhet och vilseledning.¹

Med moderna, digitala ledningsstödssystem för undervattensspaning kan sjöofficerare planera sökmönster genom att pröva sig fram i en hydroakustik simulering av operationsområdet. I en del ledningsstödssystem finns även beslutsstöd som ger förslag på sökmönster, såsom *Acoustic Mission Planner* (Scuderi, 2021) och *Undersea Warfare Decision Support System* (AN/UYQ-100, 2021) ombord på ubåtsjakthelikoptrar respektive fartyg i USA:s flotta. Den snabbt ökande mängden obemannade plattformar under, på och ovanför vattnet kan komma att medföra ett akut behov av liknande beslutsstöd för planering och omplanering av spaning.

Acoustic Mission Planner och *Undersea Warfare Decision Support System* föreslår sökmönster som är noggrant anpassade till den specifika ubåtsjaktuppgiften och vattenvolymen. Beslutsstöden identifierar ett sökmönster som är ”akustiskt optimalt” genom att en optimeringsalgoritm prövar sig fram i en hydroakustisk simulering.²

Optimeringen i beslutsstöden gör dessvärre de föreslagna sökmönstren förutsägbara för undervattensfarkosten som jagas – en viss situation (vattenvolym, sökstyrka med mera) ger ett visst bestämt sökmönster. Följer sökstyrkan beslutsstödet rekommendationer i en serie spaningsinsatser kan spaningen exploateras av en adaptiv motståndare som studerar spaningsverksamheten över tid och lär sig hur beslutsstödet ”tänker”.

Det ligger nära till hands att anta att det krävs intuition, psykologi och list för att göra sökmönster oförutsägbara (Hughes, 2025) – förmågor som verkar vara svåra att fånga med en mekanisk algoritm. Men AI

¹ <https://www.forsvarsmakten.se/sv/aktuellt/2018/04/funktionsovning-i-ubatsjakt- visar-att-ubatsjaktformagan-har-starkts/>.

² Sökmönster optimeras i *Acoustic Mission Planner* och *Undersea Warfare Decision Support System* genom en form av tabulär förstärkningsinlärning (Brown S., 1980) respektive genetiska algoritmer (Kierstead & DelBalzo, 2003) givet en modell av vattenvolymen och givet en manuellt specificerad beteendemodell för fiendens undervattensfarkost. Optimeringsmålet, som beror på uppgiften, kan exempelvis vara att maximera kumulativ detektionssannolikhet.

som utvecklats för att spela bluff- och vilseledningsspel har visat sig kunna användas till att generera oförutsägbara rutter för patrullering inom olika arenor. USA:s kustbevakning genererar sedan snart ett decennium tillbaka oförutsägbara rutter för fartygspatruller som skyddar hamnar mot intrångsförsök ovanför ytan med beslutsstödet *PROTECT* (An, et al., 2012) och oförutsägbara rutter för fartygspatruller som skyddar fiskevatten mot tjuvfiske med beslutsstödet *COmPASS* (Fang, Stone, & Tambe, 2015). En mängd naturreservat genererar sedan några år tillbaka oförutsägbara rutter för fotpatruller som skyddar vilda djur mot tjuvjakt med hjälp av beslutsstödet *PAWS* (Fang, et al., 2017). Sjöleder, flygplatser, kollektivtrafik, IT-nätverk m.m. övervakas på liknande vis med datorgenererade oförutsägbara patruller på olika håll (Sinha, et al., 2018).

Beslutsstöden ovan, som genererar oförutsägbara rutter för patrullering inom olika arenor, har beskrivits som en av de mest praktiskt framgångsrika tillämpningarna av AI inom försvar och säkerhet (Walsh, et al., 2021).

Teoretiska arbeten har överfört algoritmerna från okulär spaning på marken och ovanför vattnet till spaning efter undervattensfarkoster (Hew & Yiap, 2018). Men arbetena är begränsade till planering av undervattensspaning med passiva (tysta) sonarer, vilka i dag inte har en rimlig chans att upptäcka små, smygande undervattensfarkoster i Östersjön i realtid (Sadowski, 2025). Att överföra algoritmerna till spaning med aktiva sonarer är problematiskt. När en undervattensfarkost jagas med aktiva sonarer får den information i realtid om var jakten bedrivs och kan kontinuerligt anpassa sin egen fart och riktning därefter, men algoritmerna som används i *PROTECT*, *PAWS* med flera kan inte hantera bluff- och vilseledningsspel i vilka spelare får ny information under spelets gång (Wang, et al., 2019).

Algoritmerna i patrulleringsstöden är över ett decennium gamla. Sedan dess har AI utvecklats i rask takt, och så även AI för bluff- och vilseledningsspel. I dag är algoritmer övermänskligt oförutsägbara i pokerliknande bluff- och vilseledningsspel (Brown & Sandholm, 2019); pokerbottar vinner ofelbart i en turnering med professionella

pokerspelare, och detta även om de mänskliga motspelarna innan turneringen får studera hur pokerboten har agerat i tidigare turneringar.

I denna rapport undersöks om det är möjligt att generera oförutsägbara sökmönster för aktiva sonarer med hjälp av AI hämtad från det senaste inom pokerbottar (Schmid, et al., 2023) (Sokota, et al., 2023):

Går det att anpassa sökmönster till vattenvolym och uppgift, på det sätt som görs i Acoustic Mission Planner, utan att sökmönstergenereringen kan exploateras av en adaptiv motståndare som studerar spaningsinsatser över tid för att lära sig hur algoritmen "tänker"?

Rapporten riktar sig till läsare med intresse för AI-baserade taktiska beslutsstöd. Rapporten behandlar en viss avgränsad uppgift inom undervattensspaning, men maskininlärningsmetoden som presenteras kan vara tillämplig även i andra sammanhang där krigets dimma och adaptiva motståndare nödvändiggör ett oförutsägbart uppträdande.

Rapportens resterande delar är strukturerade enligt följande. Kapitel 2 beskriver det specifika beslutsproblemet inom undervattensspaning som undersöks i rapporten. Kapitel 3 introducerar en prototyp som utvecklats för att stödja beslutsproblemet. Maskininlärningsmetoden beskrivs i detta kapitel översiktligt och läsaren hänvisas till Limér m.fl. (2025) för tekniska detaljer. Kapitel 4 återger hur prototypen har verifierats. Kapitel 5 sammanfattar och presenterar slutsatser.

2 Beslutsproblem

I rapporten undersöks om algoritmer hämtade från pokerbottar kan generera oförutsägbara sökmönster. Undersökningen har avgränsats till följande beslutsproblem:

Hur ska en grupp farkoster med aktiva sonarer bäst söka av en vattenvolym för att hindra en möjlig undervattensfarkost att komma för nära skyddsobjekt i vattenvolymen?

Farkosterna som avses i beslutsproblemet kan vara flygande eller ytgående, bemannade eller obemannade, utrustade med släpsonarer eller doppsonarer. Hotet kan vara bemannat eller obemannat (fjärrstyrt eller autonomt). Skyddsobjekten kan vara orörliga eller rörliga. Beslutsproblemet har dock begränsats till planering av aktiva (sändande) sonarer; passiva sonarer har numera inte en rimlig chans att upptäcka hot i realtid (Sadowski, 2025).

I kapitlet introduceras först ett konkret exempel på beslutsproblemet och därefter diskuteras jaktstyrkans informationsunderläge och behovet av oförutsägbarhet.

Scenario

Följande (pedagogiskt tillrättalagda) scenario tjänar som ett löpande exempel genom rapporten.

Scenario. *Ett misstänkt fartyg från en expeditionsflotta stannar plötsligt mitt ute på havet, vilket är en indikation på att fartyget kan vara på väg att släppa ut en obemannad undervattensfarkost (figur 1). Det finns två infrastrukturanläggningar under vattnet i området som behöver skyddas mot undervattenshot. En sökenhet, i form av en fjärrstyrd ytfarkost med doppsonar,³ skickas till området och anländer efter en timme. Förstärkning förväntas inom sex timmar. Hur ska sökenheten söka av området till dess? Den misstänkta undervattensfarkosten bedöms kunna färdas upp till sju kilometer i timmen. Sökenheten antas hinna med fyra sökningar (dopp med sonar) i timmen med upp till fyra kilometer mellan doppställena. För att göra scenariot enkelt att förstå och följa antas att spaningsområdet är begränsat till det 25 × 25 kilometer stora området som syns i figur 1.*



Figur 1. Scenario. Ett fartyg (fartygsikon) misstänks släppa ut en obemannad undervattensfarkost. Det finns två infrastrukturanläggningar (flaggor) i området som behöver skyddas mot undervattenshot.

³ Obemannade ytfarkoster med doppsonar är en nyare typ av ubåtsjaktförmåga som produceras av Aselan (MIR USV), Elbit (Seaagull USV) med flera.

Informationsunderläge

Vid aktiv spaning är jaktstyrkan i ett betydande informationsunderläge gentemot undervattensfarkosten. Ljudet som en aktiv sonar sänder ut är i allmänhet tillräckligt starkt för att röja den jagande enhetens egen position, även för mycket avlägsna mål.⁴ Undervattensfarkosten kan därmed följa jakten i realtid och manövrera utifrån var sökenheterna befinner sig.

Undervattensfarkosten kan i viss mån även förutse hur sökenheterna kommer att fortsätta från sina nuvarande positioner på basis av hur jakt har bedrivits och observerats historiskt. Fientliga aktörer kan studera jakten över längre tid med kikare, radar, satellit med mera och lära sig statistiska regelbundenheter i hur jagande enheter agerar i olika situationer. Även i en gråzonkonflikt, som i scenariot ovan, kan spaningsverksamheten förväntas bli kartlagd innan ett intrångsförsök genomförs, bland annat genom att insatser provoceras fram.

Undervattensfarkosten har sammanfattningsvis ett betydande informationsöverläge. Undervattensfarkosten vet var sökenheterna har varit hitintills under jakten, och var de är nu och troligen är närmast, och kan anpassa sin egen fortsatta rörelse utifrån detta.

⁴ Till detta kommer att en operatör av en fjärrstyrd undervattensfarkost i dag även kan följa sökenheternas rörelser ovanför vattenytan med kikare, radar, satellit med mera.

Oförutsägbarhet

Vid fri ubåtsjakt som syftar till att eliminera ett hot genom bekämpning ska generellt största möjliga överraskningseffekt eftersträvas. Vid genomförande av fri ubåtsjakt som syftar till att förhindra, alternativt försvåra, fientlig/främmande undervattensverksamhet ska även uthållighet vägas in.

Överraskning erhålls genom dold ingång i spaningsområdet, passiv spaning eller genom upprepad, kortvarig och opredikterbar aktiv spaning. God överraskningseffekt erhålls genom sjöoperativ helikopter-rote som oupptäckt påbörjar samtidig spaning för att därefter snabbt förflytta sig och inta nya spaningslägen inom det aktuella spaningsområdet.⁵

Aktiv spaning behöver ofta vara oförutsägbar. Men att spana oförutsägbart är icke-trivialt och kräver erfarenhet, intuition, psykologi och list (Hughes, 2025). Att flytta aktiva sonarer på måfå inom osäkerhetsområdet är sällan effektivt.

Sonarernas förmåga att detektera målet (undervattensfarkosten) varierar under spaningen beroende på avståndet till målet, den omgivande bottenpografin, anslagsvinkeln med vilken sonarens ljud träffar målet, djupskiktet som målet befinner sig i, farkosternas fart, bakgrundsbrus med mera. Sökmönstrets oförutsägbarhet behöver därför anpassas till akustiken i vattenvolymen och till de olika sätt som målet kan agera på för att undgå upptäckt.

⁵ Reglemente Taktik för marina operationer – Del 2.

3 Beslutsstöd

Detta kapitel presenterar en prototyp till ett beslutsstöd för aktiv spaning, *Deep Acoustic Mission Planner* (DAMP). Prototypen beräknar sökmönster och positionsskattning för aktiv spaning utifrån en hydroakustisk modell av vattenvolymen, liksom *Acoustic Mission Planner* gör ombord på ubåtsjakthelikoptrar i USA:s flotta. Men sökmönsterberäkningen görs på ett sätt som – om prototypen fungerar som avsett – inte kan exploateras av en adaptiv motståndare.

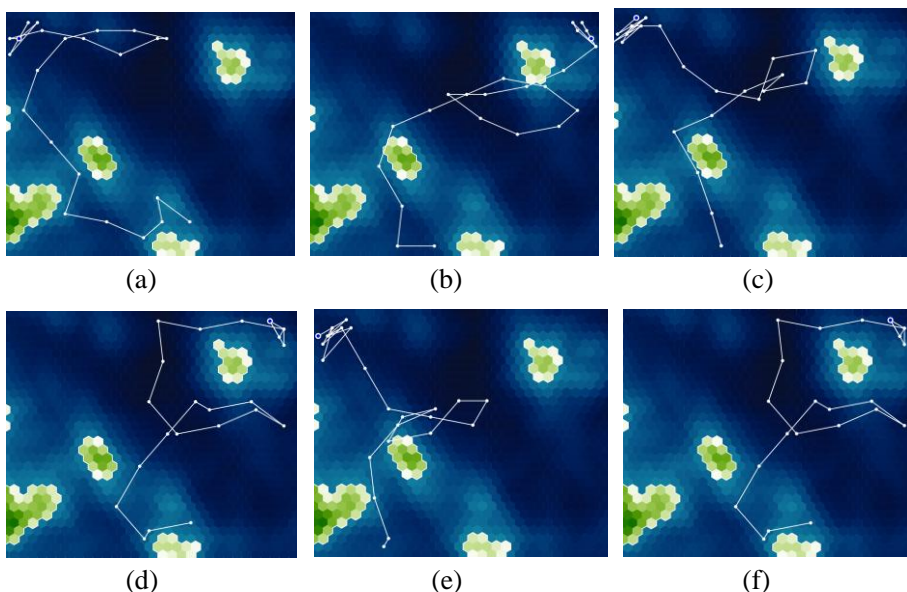
Kapitlet introducerar sökmönsterberäkningen och positionsskattningen och beskriver därefter översiktligt den underliggande maskininlärningen.

Prototypen försöker generera sökmönster som är anpassade till vattenvolymen och målsättningen på samma sätt som beslutsstöden hos USA:s flotta, men som också är oförutsägbara.

Oförutsägbarhet

DAMP skapar osäkerhet för undervattensfarkosten genom randomisering. Vägpunkterna i sökmönstret slumpas fram, men inte godtyckligt (med uniform slump), utan på ett avvägt sätt mellan ”effektiva, men väntade” beslut, såsom att röra sig mot vatten som inte söktes av tidigare under jakten, och ”mindre effektiva, men oväntade” beslut, såsom att vända tillbaka mot vatten som nyligen söktes av. Följande exempel illustrerar hur vägpunkter slumpas på ett avvägt sätt.

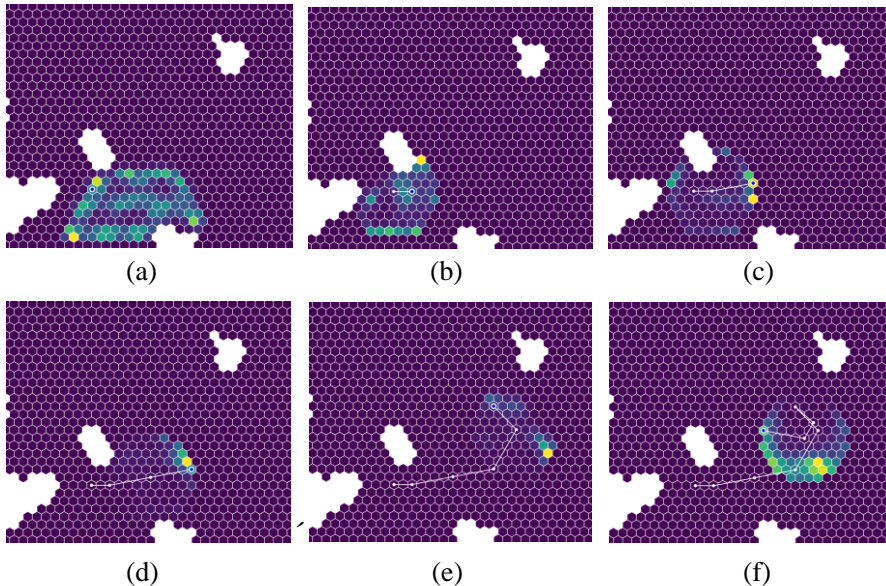
Exempel 1 (fortsättning). *Figur 3 visar sökmönster som DAMP slumpar fram när det tillfrågas sex gånger. Sökmönstren uppvisar en gemensam övergripande stridsidé: att söka av från söder (där undervattensfarkosten sjösattes) mot skyddsobjekten i nordväst och nordost, att ibland växla mellan öster och väster (men ibland påbörja ett sådant sidbyte för att strax därefter avbryta det), att några gånger vända tillbaka mot ett område som nyligen söktes av med mera.*



Figur 3. Randomiserade sökmönster för scenariot i kapitel 2.

Randomiseringen ändras allteftersom under jakten beroende på hur sökenheterna har rört sig hitintills. Exemplet nedan illustrerar.

Exempel 1 (fortsättning). *Figur 4 visar hur sannolikhetsfördelningen över vägpunkter kan förändras över tid i scenariot. Första vägpunkten (a) slumpas någorlunda uniformt inom osäkerhetsområdet (som ges av undervattensfarkostens bedömda maxfart), med lite mer sannolikhetsmassa mot perimetern och västerut. Andra vägpunkten (b) slumpas inom en cirkel runt den första, med större sannolikhet för positioner i söder och på större avstånd och med mindre sannolikhet för positioner i nordväst. Beslutet om tredje vägpunkten (c) är (grovt sett) ett val mellan att fortsätta österut eller att vända om, där mer sannolikhet ges till att fortsätta. Fjärde vägpunkten (d) slumpas med tyngdpunkt i nordost. Beslutet om sjätte vägpunkten (e) är (återigen lite förenklat) ett val mellan nordväst och sydost, där mer sannolikhet ges till det senare. Nionde vägpunkten (f) slumpas i en ungefärlig halvcirkel runt åttonde vägpunkten, med tyngdpunkt söderut.*

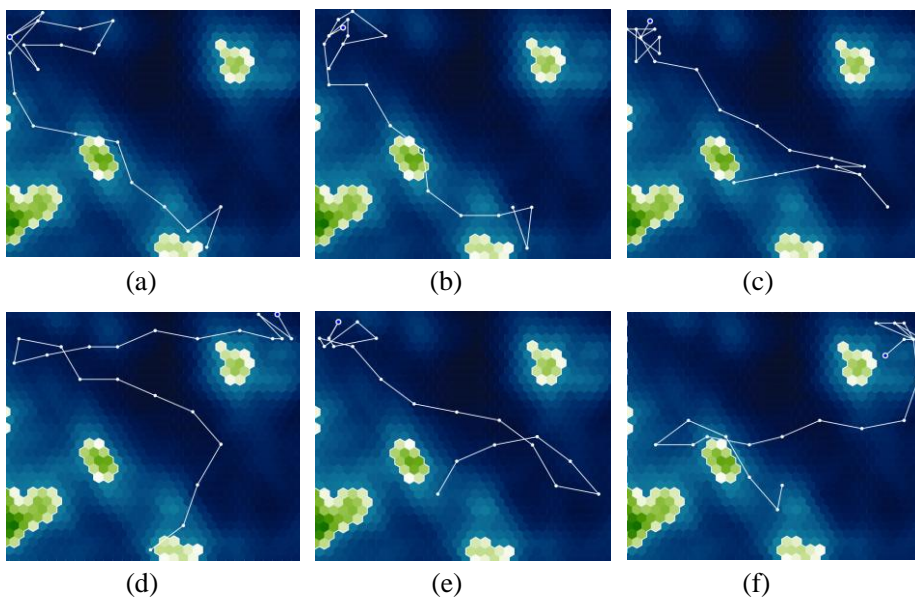


Figur 4. Randomisering av nästa vägpunkt. Sannolikhetsfördelningen som slumpade ut den första (a), den andra (b), den tredje (c), den fjärde (d), den sjätte (e) och den nionde (f) vägpunkten i ett visst sökmönster. Sannolikhetsfördelningen visualiseras med hjälp av viridis-färgkartan, skalad till minimi- och maximivärmen (ljusare grönt/gult indikerar högre sannolikhet). Låga sannolikheter har filtrerats bort.

Anpassning till vattenvolym

Beslutsstöden hos USA:s flotta försöker identifiera *ett sökmönster* som är "akustiskt optimalt", ett unikt sökmönster som är perfekt anpassat till vattenvolymen. DAMP å sin sida försöker i stället att identifiera en *randomisering* som är "akustiskt optimal", en randomisering som är perfekt anpassad till vattenvolymen. Följande exempel illustrerar hur akustiken påverkar oförutsägbarheten.

Exempel 2. Anta för exemplets skull att större efterklang i grundare vatten gör det svårare att detektera mål i scenariot ovan. Figur 5 visar några sökmönster som DAMP slumpar fram med denna akustik. Jämfört med tidigare (figur 3) pendlar jakten nu i mindre utsträckning mellan öster och väster, avslutas oftare vid det västra skyddsobjektet och fokuserar (som väntat) mer på grundare vatten.

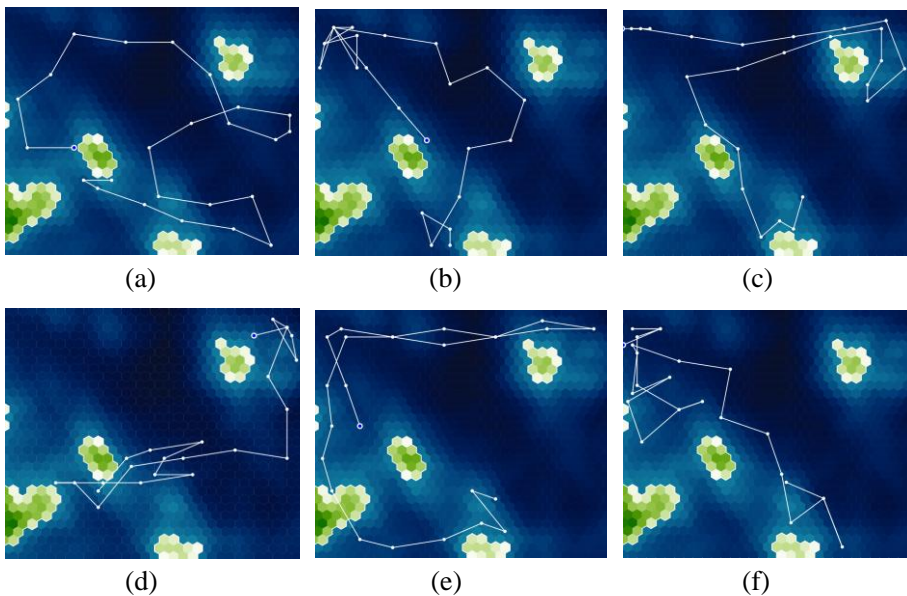


Figur 5. Randomiserade sökmönster när grundare vatten ger större efterklang.

Anpassning till målsättning

Beslutsstöden hos USA:s flotta anpassar förstås sökmönster även till målsättningen för jakten, såsom var skyddsobjekten är och hur nära inpå som ett hot kan tillåtas. Följande exempel visar hur DAMP optimerar oförutsägbarheten (randomiseringen) med avseende på jaktens målsättning.

Exempel 3. *I ett gråzonsläge med en lägre konfliktnivå kan det vara av politiskt-strategiskt intresse att få bevis för intrångsförsök. Jakten kan då tänkas inte bara vilja hindra den misstänkta undervattensfarkosten att nå ett skyddsobjekt, utan även vilja hitta den. Figur 6 visar några sökmönster som DAMP slumpar fram om målsättningen i scenariot ovan även inkluderar att hitta den misstänkta undervattensfarkosten. Sökenheten har nu inte lika bråttom som tidigare (figur 3) att arbeta sig upp mot skyddsobjekten och kan vända tillbaka djupt söderut (bort från skyddsobjekten) även under senare skeden av jakten.*



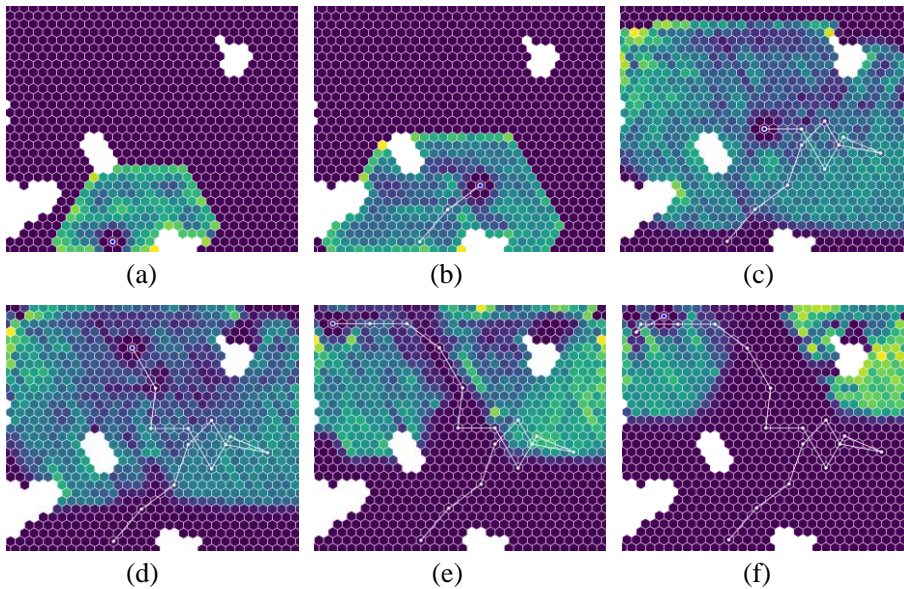
Figur 6. Randomiserade sökmönster när målsättningen för jakten även inkluderar att hitta undervattensfarkosten.

Positionsskattning

Beslutsstöden hos USA:s flotta upprätthåller en positionsskattning som estimerar var den misstänkta undervattensfarkosten sannolikt befinner sig givet hur området har sökts av hitintills under jakten. Positionsskattningen visualiseras som en värmekarta över sjökortet; ju varmare en position är, desto större sannolikhet är det att undervattensfarkosten befinner sig där.

I DAMP utgör positionsskattningen en *worst-case analysis*. Värmekartan estimerar positionen hos en undervattensfarkost som känner till vattenvolymen och manövrerar med en välavvägd oförutsägbarhet för att nå ett skyddsobjekt. Följande exempel illustrerar positionsskattningen i DAMP.

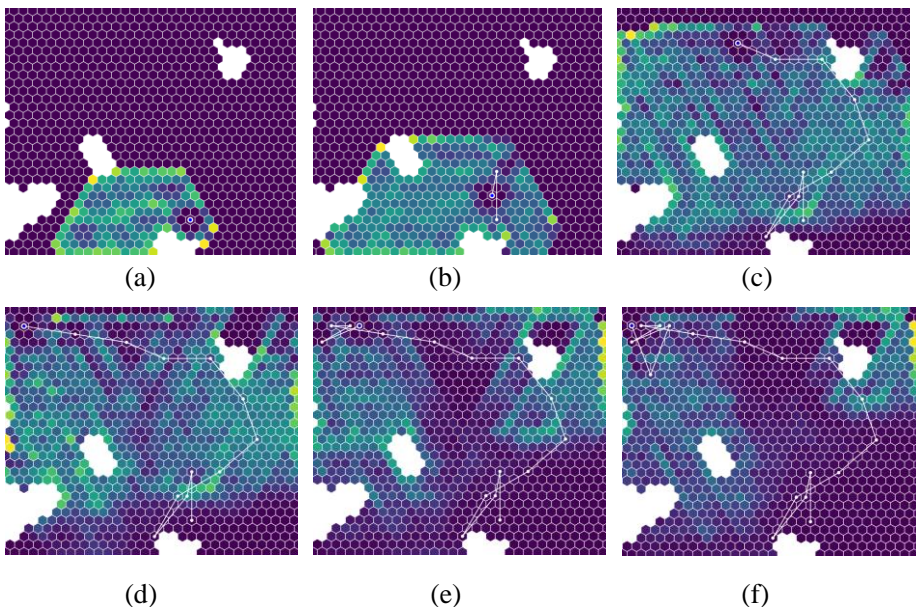
Exempel 1 (fortsättning). *Figur 7 visar hur positionsskattningen kan utvecklas över tid i scenariot i kapitel 2. Efter de första (a) och tredje (b) sonardoppen i figur 7 är fördelningen någorlunda jämn inom osäkerhetsområdet, med något mer sannolikhetsmassa mot perimetern och med försumbar sannolikhet kring själva doppstället. Efter de elfte (c) och trettonde (d) doppen är fördelningen nu mindre homogen, med en övervikt västerut. Efter de sextonde (e) och artonde doppen (f) är fördelningen än mindre homogen och koncentrerad kring de två skyddsobjekten, med en tydlig övervikt kring det östra.*



Figur 7. Positionsskattning i DAMP. Positionsskattning efter negativa sonarrapporter direkt efter första (a), tredje (b), elfte (c), trettonde (d), sextonde (e) och artonde (f) doppställena utmed ett visst sökmönster. Sannolikhetsfördelningen visualiseras med viridis-färgkartan och skalas till respektive minsta och högsta sannolikheter (ljusare grönt/gult indikerar högre sannolikhet). Låga sannolikheter har filterats bort.

Positionsskattningen, som utgör en *worst-case analysis* för spaningsuppgiften, antar att undervattensfarkosten agerar för en målsättning som är direkt motsatt jaktens. Följande exempel visar hur målsättningen kan påverka positionsskattningen.

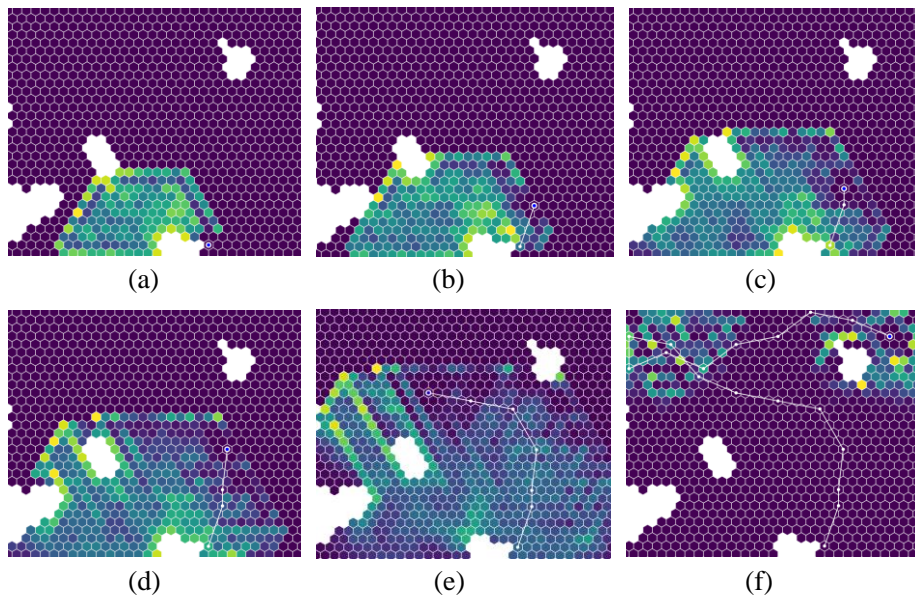
Exempel 3 (fortsättning). *Figur 8 visar hur positionsskattningen i scenariot i kapitel 2 kan utvecklas över tid om målsättningen för spaningsuppgiften även inkluderar att hitta den misstänkta undervattensfarkosten; positionsskattningen antar nu att undervattensfarkostens mål även innefattar att försöka undkomma upptäckt. Sannolikhetsfördelningen i början av jakten i figur 8 (a)–(c) är nu mer jämnt fördelad inom osäkerhetsområdet jämfört med positionsskattningen med den snävare målsättningen i figur 7 (a)–(c). Mot slutet av jakten i figur 8 (e)–(f) är sannolikhetsfördelningen nu mer utspridd över spaningsområdet jämfört med positionsskattningen med den snävare målsättningen i figur 7 (e)–(f).*



Figur 8. Positionsskattningar givet den utökade målsättningen, efter negativa sonarrapporter, direkt efter första (a), tredje (b), elfte (c), trettonde (d), sextonde (e) och artonde (f) doppställena utmed ett visst sökmönster. Låga sannolikheter har filtrerats bort.

Undervattensfarkosten antas känna till akustiken i vattenvolymen och anpassa sitt oförutsägbara agerande till denna, liksom de jagande enheterna gör. Exemplet nedan illustrerar hur akustiken kan påverka positionsskattningen.

Exempel 2 (fortsättning). *Figur 9 visar hur positionsskattningen i scenariot i kapitel 2 kan utvecklas över tid om efterklang i grundare vatten gör det svårare att detektera mål. Sannolikhetsfördelningen är nu mindre jämnt fördelad inom osäkerhetsarean jämfört med positionsskattning med mer homogen akustik (figur 7).*



Figur 9. Positionsskattningar under ett antagande om större efterklang i grundare vatten, efter negativa sonarrapporter, direkt efter första (a), andra (b), tredje (c), fjärde (d), sjunde (e) och artonde (f) doppställena utmed ett visst sökmönster. Låga sannolikheter har filtrerats bort.

Som förklaras i nästa avsnitt blir sökmönstren som DAMP genererar inte mindre, utan mer, effektiva för spaningsuppgiften om undervattensfarkosten som jaktinsatsen möter råkar ha en annan målsättning än vad som antas i positionsskattningen eller om undervattensfarkosten inte agerar med det taktiska kunnande som antas i positionsskattningen.

Undervattensspaning som poker

DAMP genererar sökmönster och positionsskattningar med en variation av en form av förstärkningsinlärning som utvecklats för att spela poker en så optimal oförutsägbarhet som möjligt. I det här avsnittet återges översiktligt hur förstärkningsinlärningen har överförts från poker till undervattensspaning i DAMP. Läsaren hänvisas till Limér m.fl. (2025) för tekniska detaljer.

Poker-AI

Professionella pokerspelare spelar oförutsägbart. Men spelarna gör sig inte oförutsägbara med hjälp av psykologi och list, som i sällskapspoker, utan med en noggrant kalibrerad randomisering. Varje beslut i spelet baseras på en randomisering som anpassas till situationen – de egna privata korten, de publika korten, potten och spelronden. Spelarna försöker, så gott de kan, slumpa beslut med en sannolikhetsfördelning som är spelteoretiskt optimal för situationen i spelet, en så kallad Nashjämvikt (Nash, 1950).

Med Nashjämvikten viktas beslutsalternativ på ett sätt som gör att det inte uppstår några (statistiska) regelbundenheter i upprepade spel som motspelaren kan observera och exploatera till sin fördel.

Nashjämvikten kan ses som en sorts försiktighetsprincip: den är effektiv oaktat hur motspelaren faktiskt spelar. Ju mindre spelteoretiskt optimalt motspelaren spelar, desto större blir den förväntade utdelningen för spelaren. Motspelaren kan inte exploatera Nashjämvikten genom att spela "irrationellt" eller "helt oväntat".

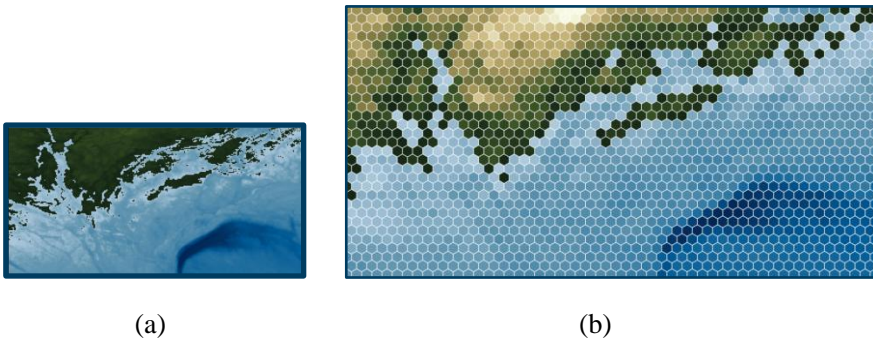
En pokerbot spelar enligt samma princip som professionella pokerspelare och försöker även den att slumpa drag enligt Nashjämvikten. Boten lär sig en approximation av Nashjämvikten genom att spela mot sig själv i stora mängder simulerade pokerpartier.

Pokerbottar är i dag övermänskligt oförutsägbara (Brown & Sandholm, 2019). En modern pokerbot vinner ofelbart i en turnering med mänskliga motspelare, och detta även om de mänskliga motspelarna innan turneringen får studera hur pokerboten har agerat i tidigare turneringar.

Anpassning till aktiv spaning

DAMP genererar beslutsstöd för en spaningsuppgift genom att översätta uppgiften till ett pokerliknande spel som analyseras med förstärkningsinlärning hämtad från det senaste inom pokerbottar (Schmid, et al., 2023) (Sokota, et al., 2023). För att minska tiden det tar att beräkna Nashjämvikten har förstärkningsinlärningen anpassats till dynamiken i undervattensspaning och till parallell beräkning på accelererad beräkningshårdvara.

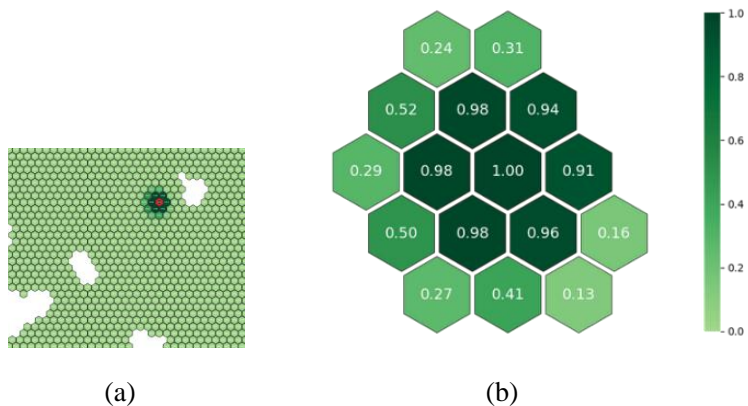
Jaktspel. Spaningsuppgiften har överförs till ett dubbelsidigt jaktspel (eng. *pursuit-evasion game*) (Isaacs, 1965) där sökplattformar och undervattensfarkost ("spelpjäserna") rör sig i en diskret modell av vattenvolymen med en vald upplösning (figur 10).



Figur 10. Diskret vattenvolym. Data för en utvald del av Östersjön (a) översätts till homogena hexagoner (b) i ett två- eller tredimensionellt rum med vald upplösning (antal hexagoner). I de pedagogiskt förenklade exemplen i denna rapport har den diskreta vattenvolymen bara två dimensioner och hexagonerna har bara ett attribut, havsdjup.

Den akustiska krigsdimman i vattenvolymen modelleras av detektionssannolikheter (figur 11) som kan bero på avståndet till målet, målets fart, anslagsvinkeln (med vilken ljudet som sonaren avger träffar målet), vilket djupskikt som målet befinner sig i, bottenpografien kring sonaren med mera.⁶ Detektionssannolikheterna ges av stokastiska sonarekvationer (Wagner, Mylander, & Sanders, 1999) eller av en uppslagstabell som beräknats med annat verktyg.

Förstärkningsinlärning. Jaktspelen i den diskreta vattenvolymen analyseras med förstärkningsinlärning som är hämtad från poker-AI. Spelteoretiskt sunda abstraktioner och heuristiker, specifikt anpassade för det underliggande spelsystemet för undervattensspaning, stabiliserar och snabbar upp inlärningen utan att ändra Nashlösningen som beräknas. Förstärkningsinlärningen är även anpassad för accelererad beräkningshårdvara; tillförs mer beräkningsresurser minskar beräkningstiden.⁷



Figur 11. Diskret akustik. (a) Ett visst doppställe i exempel 2. (b) Sannolikheter för att upptäcka en undervattensfarkost i omkringliggande positioner till just detta doppställe. I exemplet beror detektionssannolikheter på avstånd och havsdjup enligt en enkel schablon.

⁶ Den momentana detektionssannolikheten är en valfri funktion av (i) undervattensfarkostens position i den tredimensionella vattenvolymen, (ii) undervattensfarkostens fart, (iii) undervattensfarkostens riktning, (iv) sökenheternas positioner och (v) sökenheternas fart (för ytfarkoster med släpsonar).

⁷ Spelsystemet (simulatorn) är implementerat i JAX (Bradbury, 2018) och anpassat för parallell simulering på accelererad beräkningshårdvara.

Beslutsstöd. Sökmönster och positionsskattning kan genereras i realtid med en enklare dator när spaningsuppgiften genomförs. Sökmönster genereras genom att slumpa nästa vägpunkt utifrån den förberäknade Nashjämvikten. Positionsskattningen uppdateras under jakten genom att kombinera Nashjämvikten (för undervattensfarkosten) med inkommande sonarrapporter.

4 Experimentell verifiering

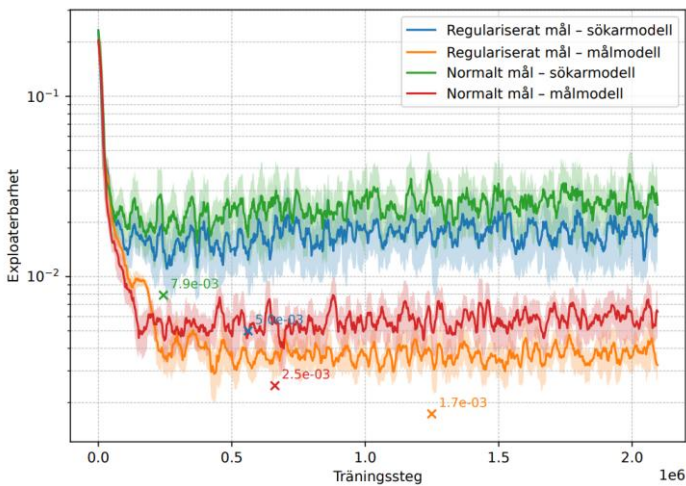
Detta kapitel beskriver hur DAMP:s oförutsägbarhet har utvärderats genom ett formellt spelteoretiskt mått och genom upprepade spel mot mänskliga motspelare. Simulatoren, det vill säga modellen av operationsmiljön, har inte verifierats.

Spelteoretisk exploaterbarhet

Inom spelteori bedöms en algoritms oförutsägbarhet i termer av så kallad *exploaterbarhet* (eng. *exploitability*), som är ett formellt mått på hur robust algoritmen är mot adaptiva motspelare.⁸ Exploaterbarhet i ett mer komplext bluff- och vilseledningsspel är dock praktiskt omöjligt att beräkna. Exploaterbarhet hos AI för poker verifieras därför inte i spelet poker självt, utan verifieras indirekt i extremt förenklade pokervarianter med några få spelkort och några få spelronder, såsom Kuhnpoker (Kuhn, 1950).

DAMP:s exploaterbarhet har mätts i ett artificiellt miniatyrscenari (12 hexagoner). Sökmönstergenereringens exploaterbarhet visade sig vara försumbar (figur 12). Se Limér m.fl. (2025) för tekniska detaljer.

Experimentet styrker att DAMP:s anpassning av de senaste poker-AI-algoritmerna till aktiv spaning är spelteoretiskt sund.



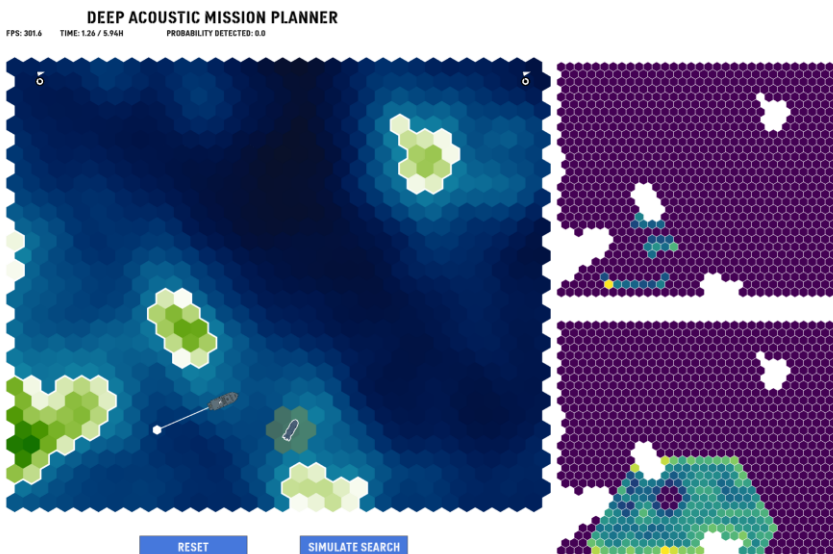
Figur 12. Exploaterbarhet i miniatyrscenari. *x-axel*: träningssteg, *y-axel*: exploaterbarhet för DAMP med olika inslag av heuristik. En adaptiv motspelare kan inte öka sin vinstchans (sannolikhet att nå till ett skyddsobjekt) med mer än 0,5–1 procent genom att anpassa sig till hur DAMP spelar.

⁸ Lite mer formellt är exploaterbarhet ett mått på hur mycket som motspelaren kan öka sitt förväntade resultat genom att anpassa sig till algoritmen, dvs. genom att optimera sitt spelande utifrån en kännedom om hur algoritmen randomiserar drag.

Empirisk exploaterbarhet

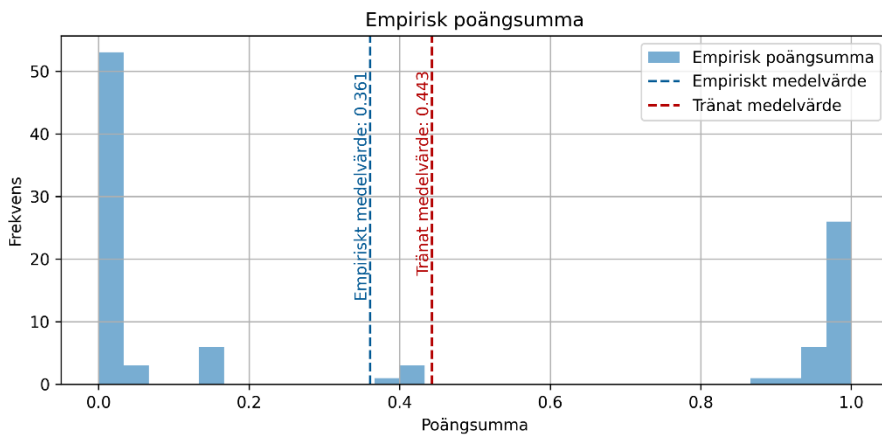
När exploaterbarhet är praktiskt omöjligt att beräkna är det praxis inom spelteoretisk forskning att pröva algoritmens oförutsägbarhet genom upprepade spel mot en mänsklig motspelare, varvid motspelaren får veta att den spelar mot en och samma algoritm i varje parti. Poker-AI, exempelvis, verifieras genom upprepade spel mot professionella pokerspelare (Brown & Sandholm, 2019).

Personal på FOI fick spela som undervattensfarkost mot DAMP i sammanlagt etthundra spel i scenariot i kapitel 2. Deltagarna uppmanades att försöka hitta sätt för undervattensfarkosten att exploatera sökenheten (dvs. DAMP:s AI). För att underlätta för deltagarna att hitta eventuell exploatering, visade användargränssnittet hur sökenheten ”tänker” (figur 13). En värmekarta (den nedre i figur 13) visade var sökenheten bedömer att undervattensfarkosten är just nu och en annan värmekarta (den övre i figur 13) visade hur sökenheten slumpar nästa vägpunkt.



Figur 13. Användargränssnitt för spel mot DAMP. Gränssnittet visar sökenhetens positionsskattning (nedre värmekarta) och sökenhetens randomisering av nästa vägpunkt (övre värmekarta). Låga sannolikheter filtreras bort i värmekartorna.

Figur 14 visar utfallet i försöket. Spelarna når ett skyddsobjekt med 36 procent sannolikhet. Om DAMP spelar mot sig själv i samma scenario når undervattensfarkosten fram med större sannolikhet, 44 procent, vilket indikerar att deltagarna inte hittat något sätt att exploatera sökenheten. För att styrka detta granskades de spelidéer (taktiker) som deltagarna haft störst framgång med under spelen; ingen av spelidéerna nådde upp till 44 procent sannolikhet vid upprepade spel mot DAMP.



Figur 14. Utfall i spel mot DAMP. Grafen visar spelens fördelning över utfall. *x-axel*: poäng för undervattensfarkosten. *y-axel*: andel spel. *Streckade vertikala linjer*: genomsnittlig poäng för undervattensfarkost när den spelas av försökspersoner (blå linje) respektive av DAMP (röd linje). Undervattensfarkostens poäng i spelet ges av sannolikheten att den når fram till ett skyddsobjekt givet hur sökenheten rörde sig under spelet.

Scenariot som spelades var tillrättalagt för att vara lätt att förstå och resonera kring för människor. Det är inte orimligt att förvänta sig att mänskliga spelare skulle prestera än sämre jämfört med DAMP i ett scenario med mer komplex akustik och fler sökenheter.

Resultatet ska tolkas med försiktighet eftersom att försöket inte var noggrant strukturerat. Vissa deltagare spelade med eftertanke och upprepade gånger, medan andra spelade snabbt och bara någon enstaka gång.

5 Slutsatser

AI för pokerliknande spel representerar en milstolpe inom modern AI.⁹ Sedan ett par år tillbaka tillämpas den militärt i en serie uppmärksammade teknikutvecklingsprojekt i USA, men under företagssekretess där metoder, algoritmer och beslutsproblem endast är vagt antydda (StrategyRobot, 2025). Prototypen som introducerats i denna rapport, *Deep Acoustic Mission Planner*, ser ut att vara den första publikt rapporterade tillämpningen till ett militärt beslutsproblem.

Prototypen beräknar sökmönster och positionsskattning för jakt med aktiva sonarer, liksom görs av *Acoustic Mission Planner* som finns ombord på ubåtsjakthelikoptrar i USA:s flotta.

Sökmönster och positionsskattning beräknas på ett sätt som – om prototypen fungerar som avsett – inte kan exploateras av en adaptiv motståndare. Prototypens exploaterbarhet har utvärderats genom ett formellt spelteoretiskt mått och genom upprepade spel mot mänskliga motspelare enligt praxis inom spelteoretisk forskning. Givet att den underliggande modellen av vattenvolymer är rimlig, indikerar resultaten att prototypen inte kan exploateras av en motståndare som studerar spaningsinsatser för att lära sig hur prototypen ”tänker”.

Slutsatsen från arbetet är att beslutsstöd för aktiv spaning kan nyttja AI för pokerliknande spel till att generera oförutsägbara sökmönster.

Framöver planeras prototypen prövas i en experimentmiljö med obemannade plattformar.

⁹ Marvin Minsky Medal for Outstanding Achievements in AI, https://www.ijcai.org/awards/minsky_medal.

6 Referenser

- An, B., Shieh, E., Tambe, M., Yang, R., Baldwin, C., & Meyer, G. (2012). PROTECT: A deployed game-theoretic system for strategic security allocation for the United States Coast Guard. *AI Magazine*.
- AN/UYQ-100 Undersea Warfare Decision Support System. (2021). Hämtat från <https://www.navy.mil/Resources/Fact-Files/Display-FactFiles/Article/2166791/anuyq-100-undersea-warfare-decision-support-system-usw-dss/>
- Bradbury, J. (2018). *JAX: composable transformations of Python+NumPy programs*. Hämtat från <http://github.com/jax-ml/jax>
- Brown, N., & Sandholm, T. (2019). Superhuman ai for multiplayer poker. *Science*.
- Brown, S. (1980). Optimal search for a moving target in discrete time and space. *Operations research*.
- Fang, F., Nguyen, T., Pickles, R., Lam, W., Clements, G., An, B., . . . Schwedock, B. (2017). PAWS: A deployed game-theoretic application to combat poaching. *AI Magazine*.
- Fang, F., Stone, P., & Tambe, M. (2015). When security games go green: designing defender strategies to prevent poaching and illegal fishing. *International Joint Conference on Artificial Intelligence (IJCAI)*.
- Försvarsmakten. (2021). *Reglemente Taktik för marina operationer – Del 2*.
- Hew, P., & Yiap, N. (2018). Optimally randomized patrolling of chokepoints for theatre antisubmarine warfare. *Military Operations Research*.
- Hughes, W. P. (2025). *Naval warfare*. Hämtat från <https://www.britannica.com/topic/naval-warfare>
- Isaacs, R. (1965). *Differential Games: A mathematical theory with applications to warfare and pursuit, control and optimization*. John Wiley & Sons.
- Kierstead, D. P., & DelBalzo, D. R. (2003). A genetic algorithm applied to planning search paths in complicated environments. *Military Operations Research*.
- Kierstead, D., & DelBalzo, D. (2003). A genetic algorithm applied to planning search paths in complicated environments. *Military Operations Research*.
- Kuhn, H. (1950). Simplified Two-Person Poker. i A. Tucker, *Contributions to the Theory of Games*. Princeton University Press.

- Limér, C., Brynielsson, J., Cohen, M., & Rydell, F. (2025). Anti-submarine warfare planning using public belief states and self-play. *International Conference of Machine Learning and Applications (ICMLA)*.
- Matej Moravčík, M. S. (2017). DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*.
- Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the national academy of sciences*.
- Sadowski, K. (2025). *Protecting maritime infrastructure from attack: new technologies and tactics*. Navy Lookout.
- Schmid, M., Moravcik, M., Burch, N., Kadlec, R., Davidson, J., Waugh, K., . . . Bowling, M. (2023). Student of Games: A unified learning algorithm for both perfect and imperfect information games. *Science*.
- Scuderi, J. (2021). *Daniel H. Wagner Associates: Over 40 years of success*. Department of the Navy, SBIR/STTR Transition Program.
- Sinha, A., Fang, F., An, B., Kiekintveld, C., & Tambe, M. (2018). Stackelberg security games: Looking beyond a decade of success. *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. International Joint Conferences on Artificial Intelligence.
- Sokota, S., D’Orazio, R., Ling, C., Wu, D., Kolter, J., & Brown, N. (2023). Abstracting imperfect information away from two-player zero-sum games. *International Conference on Machine Learning*.
- StrategyRobot*. (2025). Hämtat från <https://www.strategyrobot.ai/>
- Wagner, H., Mylander, W., & Sanders, T. (1999). *Naval Operations Analysis*. Naval Institute Press.
- Walsh, M., Menthe, L., Geist, E., Hastings, E., Kerrigan, L., Léveillé, J., & Donnelly, B. (2021). *Exploring the feasibility and utility of machine-assisted command and control: Volume 1, findings and recommendations*. RAND.
- Wang, Y., Shi, Z., Yu, L., Wu, Y., Singh, R., Joppa, L., & Fang, F. (2019). Deep reinforcement learning for green security games with real-time information. *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. AAAI Press.



ISSN 1650-1942

www.foi.se