Minimax Adaptive Control and Estimation

Kjellqvist, Olle

2024

Minimax Adaptiv Reglering och Estimering

Olle Kjellqvist
Institutionen för Reglerteknik

Populärvetenskaplig sammanfattning av doktorsavhandlingen *Minimax Adaptive Control and Estimation*, Maj 2024. Avhandlingen kan laddas ner från: https://www.control.lth.se/publications

---

I dagens samhälle förlitar vi oss på stora komplexa system för el, vatten, transport och kommunikation. Dessa system måste styras effektivt för att bevara våra resurser, men deras ökande komplexitet gör detta allt svårare. Min forskning fokuserar på att utveckla metoder för att optimera och säkra komponenter i dessa kritiska system.

Effektivitet är bra, men kommer ofta med en prislapp i form av ökad sårbarhet. Vi drar oss till minnes vintern 2022 när skånska hushåll varnades för periodvis avstängning och elpriserna slog alla rekord och COVID-19 restriktionernas katastrofala påverkan på distributionskedjan med t.ex. brist på microchips och en pågående hungerkris som följd. Dessa exempel visar hur viktigt det är att inte bara sträva efter effektivitet, utan också att säkerställa robusthet och pålitlighet i våra system.

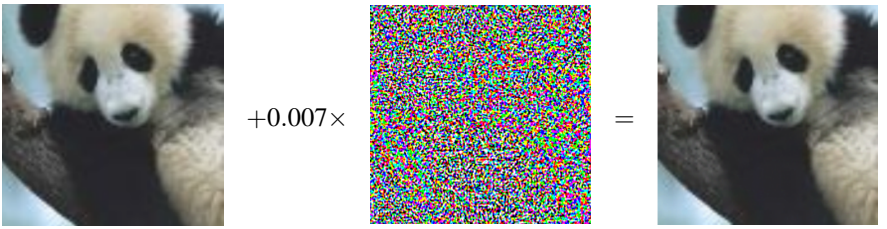
Reglerteori har växt fram som ett verktyg för att styra och övervaka tekniska system inom t.ex. telecom, processindustri, bilindustrin, flygindustrin m.fl. Teorin har sina rötter i matematiken, fysiken och datavetenskapen och studerar systematiska metoder för beslutsfattning och informationsinsamling. Att samla in information, alltså att mäta på ett system, och att fatta beslut, alltså att styra, baserat på dessa mätningar kallas för återkoppling. Reglerteorin innehåller metoder för att analysera, designa och implementera återkopplade system.

Den teori som finns idag är tyvärr inte direkt tillämpbar på de stora system som vi ser idag. En av de främsta anledningarna är att för att kunna få en god teoretisk förståelse av ett system behövs noggranna matematiska modeller. Arbetet med att skapa dessa modeller är tid- och resurskrävande och det går inte att genomföra på den skala som krävs.

För att reducera arbetsbördan och öka tillämpbarheten av reglerteori har jag i min avhandling tagit fram metoder för att uppskatta omätbara kvantiteter och designa återkopplade *komponenter* från högst osäkra modeller. Metoderna är baserad på en matematisk formulering av prestanda och sårbarhet i form av värsta-fallet analys och är systematiskt härledda för att optimera värsta fallets prestand. Metoderna är tillämpbara på en mängd olika system och har visat sig vara effektiva i simuleringar. De samlar in information om systemet, kompimerar informationen och använder den för att fatta beslut. De lär sig alltså av systemet och anpassar sig efter det. Det som utmärker mina metoder är att vi har en mycket god förståelse av

deras känslighet mot störningar och approximeringsfel.

Att känsligheten är transparant är viktigt för att vi ska kunna lita på systemet, och förstå hur det kommer att beté sig. Flera forskare har varnat för att dagens AI-lösningar är sårbara för så kallade *adversarial examples*, där små störningar i indata kan få systemet att fatta helt felaktiga beslut, som i klassificiering av pandan nedan.



Panda
58% säkerhet.

Brus

Gibbon.
99% säkerhet.

Exempel från *Explaining and Harnessing Adversarial Examples* av Goodfellow et al. på sårbarhet i maskininlärning. AI-systemet luras att felklassificera en pandabild genom att lägga till brus, så svagt att det inte syns med blotta ögat. Motsvarande brusbilder har lyckats lura självkörande bilar att en stoppskylt är en hastighetsskylt.

Att genom smarta AI-lösningar införa den typen sårbarheter är oacceptabelt för samhällskritiska system, som klassas som högrisksystem i EUs AI Act och är hårt reglerade. Därför är den transparanta känsligheten i mina metoder ett välkommet framsteg för att göra vår kritiska infrastruktur mer robust och effektiv.