

EXAMENSARBETE Scalable processing of globally crowd-sourced geolocation data**STUDENT** Oskar Jermakowicz**HANDLEDARE** Marcus Klang (LTH), Rikard Windh (Combain Mobile AB)**EXAMINATOR** Krzysztof Kuchcinski (LTH)

Skalbar inomhuspositionering

POPULÄRVETENSKAPLIG SAMMANFATTNING **Oskar Jermakowicz**

Att ta reda på var en försvunnen apparat befinner sig i ett sjukhus kan vara viktigt. Är denna apparat uppkopplad till the Internet of Things så är det möjligt att ta sensordata och beräkna vilket rum på sjukhuset den befinner sig i på några sekunder.

Enheter inom the Internet of Things ökar exponentiellt. Att hålla reda på enheters positionering inomhus blir betydligt mer relevant inom dagens samhälle. Detta kräver bearbetning av stora mängder data för att uppnå höga precisioner.

Ett nuvarande system, Combain Positioning Solutions (CPS), jämfördes med en prototyp som jag utvecklade i en big data miljö.

	Prototyp	CPS
Precision	95 meter	15 meter
Datapunkter / s	16 tusen	15 st
Datapunkter / USD	69 miljoner	120 tusen
Skalbart	Linjärt	Nej

Jämförelse mellan prototypen och CPS.

Jämförelsen visar att prototypens precision är mindre bra, men alla andra aspekter är otroligt lovande, i synnerhet skalbarheten där prototypens prestanda ökar i samband med ökning av datorkraft.

CPS är ett system utvecklat av Combain Mobile AB. I dagsläget får Combain Mobile AB in cirka 50 miljoner datapunkter varje dag och deras databas har över 48 miljarder positioner, över 1,4 miljarder Wi-Fi nätverk och över 88 miljoner mobilmaster. Denna mängd data behandlas av CPS dagligen. Inomhuspositioneringsdelen av denna

lösning har en genomsnittlig precision på 15 meter när den ska lokalisera enheter. För att ständigt förbättra precisionen, så måste programmet systematiskt lära sig genom att behandla otroligt stora mängder data.

Att behandla dagens inkommande data är inga problem för CPS, men när antalet användare ökar som förväntat med 100 till 10'000 gånger så läggs en stor tyngd på att CPS ska kunna vara skalbart. Detta är en datamängd på upp till 14 miljarder datapunkter, eller upp till 4 terrabytes data att behandla varje dag. Därför har det föreslagits att undersöka hur CPS kan anpassas till en så kallad *big data* lösning med fokus på skalbarhet.

Under examensarbetet så har jag utvecklat en prototyp i ett big data ramverk och den består av de grundläggande funktionerna i CPS. Prototypen implementerades på molnet, där flertal tester och mätningar som skalbarhet, prestanda, kostnadseffektivitet och precision genomfördes. Med hjälp av dessa mätningar så kunde jag göra en översiktlig jämförelse mot CPS.

Då antalet användare och inkommande datapunkter ökar så kommer det enligt resultatet inte räcka att bara skruva upp mängden datorkraft. Därmed kommer någon slags mer skalbar lösning att behövas. Prototypen visar flertal lovande aspekter och jag drar slutsatsen att prototypen är en potentiell grund och startpunkt för en framtida big data lösning.