

---

# HEAT MAPS I REALTID: Uppgiftsbeskrivning

DA225X: Examensarbete inom datavetenskap

Jesper NORBERG - [jenor@kth.se](mailto:jenor@kth.se)

KTH, 2014-12-19

---

# 1 Bakgrund och förutsättningar

*Varför och var ska den praktiska delen av examensarbetet genomföras?*

Den praktiska delen av examensarbetet ska utföras på Carmenta, ett företag med fokus på geografiska IT-lösningar. Jag kommer att under exjobbets gång att arbeta i deras Stockholmskontor. Jag valde Carmenta för att jag tyckte att deras exjobbsförslag verkade väldigt intressanta, samt att företaget gett mig ett starkt intryck av att vara kompetent. De utvecklar t.ex. sina egna bibliotek och produkter från grunden.

## 2 Vetenskaplighet

### 2.1 Frågan

*Ett examensarbete ska utreda en specifik vetenskaplig/teknisk fråga. Vilken är denna?*

Frågan som ska utredas är hur man effektivt kan skapa en heat map givet en stor mängd tvådimensionella datapunkter. Heat map:en är en grafisk representation av datan, som för ett område visar hur tätt datapunkterna är placerade. Den åstadkommer detta genom att mappa datatätheten till en färg. Varma färger indikerar hög datatäthet och kalla färger indikerar låg datatäthet, därav kommer namnet heat map. Då mängden datapunkter kan vara väldigt omfattande (>500000 punkter) kommer heat map:en att behöva implementeras effektivt. Sannolikt kommer parallellisering att krävas, för vilken användning av GPU och multi-core vore önskvärt. Då datamängden är omfattande och önskemål har uttryckts av Carmenta om skalbarhet, kan en heuristisk approach vara aktuell. Målet är att heat map:en ska kunna ritas upp i realtid, ie. på mindre än en sekund. Detta skulle möjliggöra för en användare att filtrera datan på olika parametrar och få direkt visuell feedback på hur datatätheten förändras.

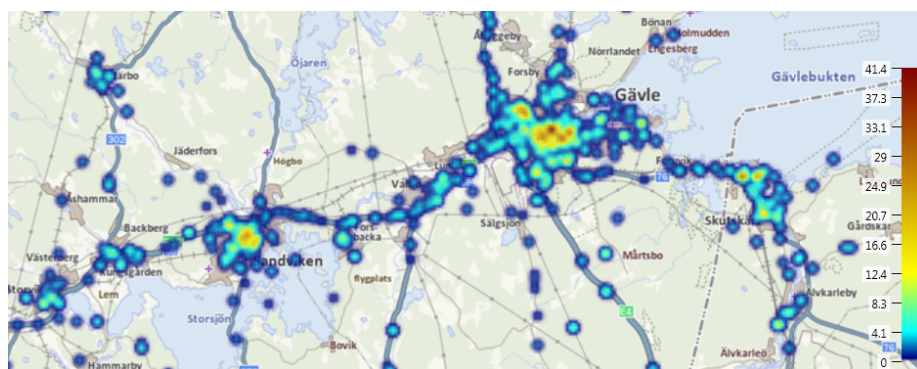


Figure 1: En heat map på en karta. Färgen indikerar datatätheten.

## 2.2 Koppling till forskning

*Beskriv på vilket sätt uppgiften anknyter till aktuell forskning eller utveckling. Förklara varför frågan är av intresse och för vem?*

Projektet är tänkt att användas främst på kartor och möjliggör visualisering av stora mängder geografisk data. Projektet är applicerbart inom områden där en sådan visualisering är relevant, i synnerhet då man vill studera förändring över tid. Den tänkta huvudanvändaren är statliga institutioner, i synnerhet SOS Alarm och liknande, som med hjälp av en visualisering av denna typ bättre kan planera sin resursallokering. Således kommer projektet vara applicerbart inom bl.a. forskning kring förbättring av effektiviteten hos SOS-alarm, vilket studeras på forskningscentret TUCAP.

## 2.3 Undersökningsmetod

*Hur ska den angivna frågan undersökas?*

Frågan kommer att undersökas genom att man först skapar en fungerande grundimplementation av heat maps och sedan försöker förbättra denna i nya implementationer. Den primära optimeringen är parallellisering via GPU och multi-core. Går inte problemet att lösa exakt på utsatt tid kommer en heuristisk approach att bli aktuell. De olika implementationerna jämförs sedan prestandamässigt med vetenskaplig metodik. Relevanta parametrar för utvärderingen inkluderar bl.a. snabbhet, exakthet och minnesåtgång.

## 2.4 Hypotes

*Vad är möjliga/sannolika utfall av undersökningen?*

En fungerande implementation av heat maps är extremt sannolik. Om problemet går att lösa exakt eller bara approximativt är oklart, men med tanke på datamängden och önskan om skalbarhet är det sannolikt att en heuristisk implementation kommer att krävas. Jag tror att projektet kommer att sluta lyckat, frågan är snarare hur bra den slutliga implementationen blir. Förhoppningsvis uppfyller den Carmentas behov med råge.

## 2.5 Utvärdering

*Hur kan man avgöra om målet med examensarbetet uppfyllts och om frågan besvarats tillräckligt?*

För att utvärdera projektet behövs en vetenskaplig studie kring hur effektiva dom olika implementationerna är, i termer av bl.a. snabbhet, exakthet och minnesåtgång. Givet att dom relevanta parametrarna ligger inom efterfrågade nivåer har programmet lyckats från Carmentas perspektiv. Den vetenskapliga delen av rapporten kommer även att innefatta en studie kring hur väl olika implementationer fungerar, samt en analys kring deras respektive fördelar och nackdelar.

### 3 Exjobbarens bakgrund

*Beskriv vilka kunskaper (kurser och/eller erfarenheter) du har som gör att detta är en lämplig uppgift för dig.*

Min datalogiska bakgrund är nästan uteslutande från KTH. Förutom dom obligatoriska kurserna har jag även läst kurser som fokuserar starkt på algoritmer, så som "Program System Construction using C++" och "Programming and Problem Solving under Pressure". Jag har presterat starkt i dessa kurser och känner att min specialitet inom datalogi ligger inom det algoritmiska. Projektet är ett algoritmiskt problem och är således väl lämpat för min bakgrund. Jag har ingen tidigare erfarenhet av att jobba med multi-core och GPU, men då denna programmering kan utföras i C++ är jag optimistisk.

### 4 Avgränsning och resurser

*Vad finns redan tillgängligt hos uppdragsgivaren i form av tidigare arbeten, programvara, kompetens etc?*

En tidigare implementation av heat maps finns på företaget, men den är betydligt långsammare än vad som krävs. Hela kartsystemet som heat map:en ska appliceras på finns färdigt som bibliotek utvecklat av Carmenta, så planen är att mitt program bara ska ta emot indata i form av datapunkter och skicka utdata i form av en heat map.

### 5 Företagskontakter

*Beskriv vilka som kommer att handleda dig på företaget och vad dessa har för roll och involvering i exjobbet.*

Min handledare på företaget kommer att vara Christoffer Rydberg. Han kommer främst att bidra med stöd kring Carmentas system, men även bistå som diskussionspartner genom arbetsprocessen. Carl-Johan Lundell är den ansvarige för exjobbet från företagets sida. Jag kommer även ha kontakt med det tekniska teamet som implementerat den gamla versionen av heat maps.