

Plottuppgift i TSRT04 (Datatrafik): Prognos för ökning av mobilanvändning

22 augusti 2016

Detta är en av plottuppgifterna i kursen TSRT04. För att bli godkänd på kursen måste ni lösa en av plottuppgifterna och redovisa den på något av examinationstillfällena. Redovisningen sker på engelska så skriv även er kod och kommentarer på engelska (det är en bra övning för framtiden!). Uppgiften ska lösas i grupper om två, eller individuellt. Eftersom det rör sig om ett examinationsmoment är det inte tillåtet att dela eller visa MATLAB-kod/anteckningar för andra studenter. Det är däremot okej att diskutera uppgiften muntligen, exempelvis för att dela med sig av goda råd!

1 Bakgrund

Mobilnäten byggdes ursprungligen för att vi skulle kunna prata i mobiltelefon med varandra. Idag är däremot mobilsamtal och sms bara en liten del av trafiken i mobilnäten. Det är trådlös datatrafik från olika populära Internet-tjänster som driver utvecklingen; exempelvis sociala nätverk, webbsidor och strömmande video. Den ökande datatrafiken kräver en ständig utbyggnad av mobilnäten med fler basstationer (på master och hustak) och nya teknologier (t.ex. 4G/LTE). Det verkar därför rättvist och transparent att det ska vara mängden dataanvändning som avgör månadskostnaden på mobilabonnemang nu och i framtiden – istället för hur många mobilsamtal som vi gör.

2 Information om dataserien

I denna uppgift ska vi studera en prognos som företaget Cisco har gjort för hur datatrafiken förändras i Sverige:

www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html

Studien beskriver hur många petabyte (10^{15} byte) av data som skickades i mobilnäten under 2013 och ger en prognos för hur det kommer utvecklas under de närmsta fem åren. Trafiken är uppdelad i tre kategorier: "video", "filöverföring" och "webb och annat". Dataserien finns i filen `datattraffic.mat` som kan laddas ner från kurshemsidan.

Ladda in dataserien i MATLAB med kommandot

```
» load datattraffic
```

Årtalen finns i vektorn `years`, medan den årliga mängden datatrafik finns i matrisen `traffic`. Den första kolumnen innehåller mängden datatrafik (per år) från video, den andra kolumnen innehåller mängden data från filöverföringar (per år) och den tredje kolumnen visar samma sak för "webb och annat". Alla siffror är i petabyte (se ovan).

3 Undersökning och visualisering av dataserien

Uppgiften är att undersöka dataserien genom att plotta datan på olika sätt.

1. Bestäm ifall ni vill skriva din kod som ett skript, i olika funktioner eller som en kombination av båda. Ifall ni gör ett skript så kan det vara bra att inleda med `clear` (för att tömma Workspace) och `close all` (för att stänga alla öppna figurer/grafar).
2. Plotta mängden datatrafik för olika typer av data som en funktion av tiden. Visa allt i samma figur med hjälp av `plot`. Använd `xlabel`, `ylabel`, `title` och `legend` för att beskriva vad som syns i figuren. Fundera på vilken typ av datatrafik som dominerar och driver utvecklingen (både när det gäller datamängd och lutningen på kurvorna).
3. Denna typ av information brukar ofta redovisas i stapeldiagram. Plotta samma sak men hjälpa av `bar`. Lägg till beskrivande text som ovan.
4. Det finns olika typer av stapeldiagram. Kolla dokumentationen för att se hur man väljer mellan "grouped" där varje datatrafiktyp får en egen stapel och "stacked" där datatypernas staplar ställs ovanpå varandra. Använd kommandot `subplot` för att plotta dessa två stapeldiagram i två olika delbilder, förslagsvis under varandra.
5. Återgå till linjeplotten som ni gjorde med `plot`. Vad betyder siffrorna för en medelperson i Sverige? Utgå ifrån att det finns 9,7 miljoner svenskar och att antalet är samma över alla åren i dataserien. Skala om `traffic` så att den visar datatrafiken för en "medelsvensk". Det kan även vara bra att ändra skalan från petabyte (10^{15} byte) till gigabyte (10^9 byte). Fundera på om siffrorna verkar rimliga och ifall de stämmer in på ert mobilanvändande.
6. Fundera på hur linjeplotten skulle se ut i utskrift. Det finns två viktiga saker att tänka på: 1) Använd "kraftfulla" färger, såsom svart, rött och blått. Grönt och särskilt gult brukar bli svåräst i tryck. 2) Säkerställ att kurvorna går att särskilja även ifall de är utskrivna i gråskala. Detta kräver olika linjestilar (t.ex. solid, dotted, dashdotted och dashed). Ni kan använda `plot` för att ställa in färg och linjestil.
7. Lägg till en linje för den totala datatrafiken i linjeplotten. Det underlättar läsandet av en graf om den förklarande rutan (skapad med `legend`) listar kurvorna i samma ordning som kurvorna ligger i bilden. Se till så att det är fallet.
8. Spara era figurer, dels i MATLAB-formatet `.fig` och dels i följande fyra vanliga bildformat: `.jpg`, `.png`, `.pdf`, och `.eps`. Öppna de sparade filerna för att jämföra bildkvaliteten. Zooma in på bilderna och se vilka format som alltid är skarpa (s.k. vektorgrafik) och vilka som blir oskarpa vid inzoomning. Pröva sedan att importera bilderna till ett ordbehandlingsprogram (t.ex. Word eller Open Office). Vilka bildformat är bäst att använda i rapporter och liknande sammanhang? Gå tillbaka till MATLAB och dra i det nedre högra hörnet på en figur för att ändra dess storlek och dimensioner på skärmen. Spara filen en gång till och se om dina förändringar följde med till den sparade filen eller inte.

9. Förhandsgranska hur figurerna ser ut i gråskala. Detta kan ni göra genom att gå till menyn "File" i en figur och välja "Print Preview". Under fliken "Color" kan ni skifta mellan färg och gråskala. Ifall det inte går att urskilja de olika kurvorna i gråskala så gör ni nödvändiga förändringar och förhandsgranskar på nytt.

För att bli godkänd på denna uppgift ska ni kunna uppvisa och köra det skript (eller funktion) som ni skapat för att lösa de olika deluppgifterna. Ni ska ha lagt in kommentarer i koden och sett till så att plottarna är självförklarande.

4 Kort om ökningen i datatrafik

Prognosen från Cisco (som dataserien härstammar från) har fått stort genomslag, inte minst bland teleoperatörer och forskare som planerar för framtida mobilnät. Ökningen i trådlös datatrafik ligger stabilt på cirka 35% per år i prognosen. För att förstå vad detta betyder så låter vi $t = 0$ vara vårt startår (t.ex. 2013 i prognosen). Om trafikmängden betecknas d (enhet: byte/år) vid vårt startår, då blir trafikmängden efter t år ungefär lika med

$$d \cdot 1.35^t. \quad (1)$$

Notera att detta rör sig om en exponentiell ökning med åren. Varje gång man ser en exponentiell ökningstakt så är det rimligt att fråga hur länge den kan hålla i sig – om det finns fysiska begränsningar som förr eller senare kommer att hindra en fortsatt ökning.

Det ligger nära till hands att tro att det är vårt nya it-samhälle (med smarta telefoner och surfplattor) som skapat den exponentiella ökningen i trådlös datatrafik. Men faktum är att det sett ut så här sedan begynnelsen av trådlös kommunikation. Martin Cooper, som var en pionjär inom mobiltelefoni på Motorola, myntade på 90-talet något som kallas *Cooper's law*. Den säger att mängden trådlös kommunikation fördubblats varje 2,5 år. Denna tumregel har stämt bra under 1900-talet och således har trafik ökat med en faktor 1 miljon under de senaste 45 åren. Bryter man ner detta till en årlig ökningstakt, $2^{1/2.5} = 1.32$, så ser vi att tumregeln föreskriver en årlig ökning med 32 %. Detta ligger förvånansvärt nära Ciscos prognos för Sverige. Med andra ord, är den exponentiella ökningstakten inget nytt, men det ska bli intressant att se vart den leder oss i framtiden.

Om ni vill kan ni läsa mer om Cooper's law här:

<http://www.arraycomm.com/technology/coopers-law/>

5 Referens

Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018. URL:

http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html