



Datavetenskap

Opponent:

Anders Broström, Niclas Kihlstradius

Respondenter:

Henrik Andersson, Peter Orelund

**City Mobility Model with Google Earth
Visualization**

1 Sammanfattat omdöme av examensarbetet

Ett fullgott examensarbete med en bra uppgift, och ett mycket intressant ämne som antagligen kommer att vara aktuellt för lång tid framöver. Synd att man inte nådde hela vägen. Det verkar som att arbetet förhindrades av den färdiga kod som användes som bas, programvara som var problematisk att installera och avrundningsfel som gäckade. Hade man inte haft dessa problem hade man förmodligen kommit mycket längre i sitt arbete, något som verkligen hade varit intressant att få ta del av. Icke desto mindre är det ett väl utfört examensarbete.

2 Synpunkter på uppsatsen knuten till examensarbetet

Uppsatsen beskriver hur en mängd verktyg har utvecklats för att möjliggöra visualisering av simulering av så kallade mobila ad hoc-nätverk i stadsektioner. Bakgrunden presenteras så att läsaren får förståelse för uppgiften och tillvägagångssättet vid utveckling. Författarna har skrivit kort och koncist, och koncentrerat sig på vad som är relevant. Intrycket är att de har haft god förståelse för sitt arbete. Vidare har texten luftats upp med bilder och exempel. Det gör att rapporten känns lättläst och fokuserad.

Uppsatsen inleder med en lättöverskådlig sammanfattning ("abstract") som återger arbetet ganska bra. Bakgrundskapitlet (kapitel 2) ger förklaringar som är mycket bra och lättförståeliga. Kapitlet upplevs dock till viss del som osammanhängande, och den röda tråden är inte så tydlig (man går exempelvis direkt från beskrivning av nätverk till förklaring av märkspråket XML). Vidare saknas en introduktion till nätverkssimulering. Även om simulering i sig inte har varit en del av arbetet hade det varit intressant om det nämnts något om motivationen bakom, och hur simulering fungerar. Resten av uppsatsen är en beskrivning av arbetet, där varje kapitel fått motsvara varsitt delmoment. Bakgrundsinformation, tillvägagångssätt, problem och resultat beskrivs därför om vartannat. Det hade kanske varit lämpligare att skapa särskilda avsnitt eller kapitel för ändamålen, så att exempelvis alla problem beskrivits på ett enda ställe. På samma sätt hade ett kapitel kunnat ägnas helt åt resultat och utvärdering.

Som icke-insatt läsare kan man ibland önska mer hjälp på traven. Följande kommentarer ska dock inte ses som krav på förändring. Det förekommer enstaka avsnitt i uppsatsen där man först efter att ha läst texten flera gånger om och gjort vissa antaganden kan ana vad det var som författarna vill få sagt (se beskrivningarna av varje kapitel). Det som saknas är också en lista med definitioner i början av rapporten, där enkla förklaringar kan ges av terminologin. En kort sammanfattning i slutet av varje kapitel hade också varit önskvärt.

2.1 Titel

Titeln kan eventuellt uppfattas som missvisande. ”City Mobility Model with Google Earth Visualization” får det att låta som att mobilitetsmodellen visualiseras (eller har stöd för visualisering) i Google Earth, vilket inte är fallet. Man skulle även kunna förtydliga kopplingen till simulering av mobila ad hoc-nätverk, kanske med en undertitel. Kanske går det att inflika att det rör sig om en utökning av en befintlig modell.

2.2 Uppsatsens disposition

Uppsatsen kan sägas ha en disposition som följer arbetsgången i arbetet som beskrivs, vilket gör att rapporten kan läsas ”flytande” från början till slut. Dock är det svårt att snabbt få en överblick av exempelvis resultat och utvärdering, eftersom dessa ligger utspridda över kapitel 3-6. Det samma gäller de olika problem som författarna beskriver. Det är förmodligen en smaksak, men rapporten hade kunnat delas upp i tydliga, specifika kapitel. Exempelvis implementering, resultat (för en tydlig överblick av vad som faktiskt åstadkommit) och utvärdering. Sammanfattningar i slutet av varje kapitel hade också varit önskvärt.

2.3 Begreppsapparat

I uppsatsen förekommer termer och förkortningar, som i de flesta fall förklaras inom parentes eller som fotnoter. Förklaringarna sker dock inte alltid enhetligt. Det vore önskvärt om en förkortning eller term alltid presenteras första gången den används. I början av kapitel 1 nämns exempelvis ”Google Earth” och ”XML” utan någon förklaring, och det är först i kapitel 2 som det framgår vad dessa innebär. Det finns även förkortningar som inte förklaras alls, t.ex. ”REAL” i avsnitt 2.7. För att åtgärda problemet hade en lista med definitioner kunnat läggas i början av uppsatsen.

Användandet av versaler i vissa termer är inkonsekvent. Exempelvis ns-2/NS-2 och Nam/NAM. Det samma gäller användandet av siffror/bokstäver, såsom "record type 1" och "record type one". I avsnitt 3.2 skrivs "complete chain basic record" med samma formatering som texten i övrigt, medan en relaterad term skrivs "*Complete Chain Shape Coordinates*", vilket sticker ut tydligare. Samma skrivsätt borde användas för båda. (Att i vissa fall använda kursiv stil för termer och i andra fall inte förekommer också på andra håll.) Slutligen kan nämnas att "mobility model" ibland förkortas till "model", vilket är mindre tydligt.

2.4 Argumentering och slutsatsdragning

Argumenteringen är oftast god, och författarna förklarar varför de gjort sina val. Ett bra exempel är avsnitt 3.1 "Why the census TIGER database".

I andra fall kvarstår dock frågetecknen. I avsnitt 3.2 skrivs att definitionen av en komplett kedja ("complete chain") är intressant för författarna, men som läsare får man inte veta varför.

I beskrivningen av mobilitetsmodellen (avsnitt 4.1) så skriver författarna att de fokuserar på implementeringen snarare än att undersöka den avancerade matematiska bakgrund som ligger bakom. Dock ges ingen anledning varför.

Även motiveringen till att inte införa stöd för stationära noder i mobilitetsmodellen hade kunnat göras bättre. Anledningen, att modellen har en inledande initialiseringsfas som placerar ut noder slumpvis känns inte helt övertygande. Tydligare argument borde ha lagts fram varför inte modellen kunde modifieras – t.ex. för att det skulle påverka simuleringen negativt.

I avsnitt 4.2 skriver författarna att "we did not want negative values so 180 (degrees) was added to all coordinates", vilket man som läsare har svårt att förstå (vad menas?) och någon motivering till varför negativa koordinater inte uppskattas ges inte.

Det kunde ha argumenterats varför simulering av mobila ad hoc-nätverk är intressant, även om detta som sagt kanske ligger i periferin. Arbetet motiveras dock av författarna med att det är avsett att hjälpa forskare inom ämnet.

I det avslutande kapitlet drar författarna slutsatser kring det åstadkomna arbetet och problemen de påträffat. Det hade också kunnat nämnas slutsatser kring vad som kunnat göras annorlunda (för att undvika problem) för att underlätta i eventuella framtida, liknande projekt.

2.5 Sammanfattningen

Sammanfattningen är i stort bra och tar upp de viktigaste punkterna i rapporten. Det finns dock enstaka saker att anmärka på. Som sagts tidigare hade det varit intressant med en motivering av nätverkssimulering, även om det inte var en del av arbetet. I övrigt anmärks bara på några detaljer: det kunde ha inflikats varför en modifikation av mobilitetsmodellen var nödvändig, och meningen "this data will then be used as input to the modified mobility model and finally run through the ns-2 simulator" är kanske missvisande. Det borde förtydligas att det är utmatningen från mobilitetsmodellen som används i ns-2, och inte samma data som används som indata till modellen. I andra stycket skriver man att man skapat "utilities" (plural) när man syftar på TIGERconv (*ett* program). I slutet nämner författarna att "de flesta" målen uppnåddes, och vilka mål som inte hanns med, men det är ottydligt vilka mål som ursprungligen sattes upp.

2.6 Språkbehandling

Språket är i regel bra och lättförståeligt. Amerikansk engelska används, och det är en fråga om stil huruvida kortformer som exempelvis "don't" istället för "do not" bör användas eller inte, bara det görs konsekvent. Enstaka grammatiska fel, stavfel och platser för uteblivna ord har markerats i rapporten.

2.7 Referat och källförteckning

I källförteckningen anges bara vissa webb-källor med datum för besök. Källorna [7], [14] och [15] i källförteckningen hänvisas inte till någonstans. Referenslistan är felaktig. Numreringen stämmer inte alltid med vad som refereras. I avsnitt 2.4 refereras [3] och [4] för information om KML. Åtminstone [3] innehåller inget KML-relaterat material, och troligtvis inte [4] heller att döma av namnet. I avsnitt 2.6 hänvisas exempelvis till [6] (en KML-relaterad källa) när man skriver om TIGER, vilket knappast är korrekt. På liknande sätt används [8] felaktigt i avsnitt 2.7 (handlar om TIGER och inte ns-2). [11], [12] och [13] används också felaktigt. Författarna bör rätta till listan och därefter noggrant kontrollera att allt stämmer en gång till. Bortser man från numreringen verkar referenserna vara korrekta.

I avsnitt 2.8 skrivs det om Nam utan att någon källa anges (inte nödvändigtvis ett fel). När man i avsnitt 4.1 skriver att "the models have perfect sampling, which means that they start in a steady state, avoiding any initial deviations" låter det som man använt en källa, men någon

sådan anges inte – åtminstone inte vid ”referatet” (återigen, inte nödvändigtvis ett fel). I avsnitt 3.2 görs ett citat i slutet av fjärde stycket som inte har någon tillhörande referens.

Det kan påpekas att det är otydligt varifrån man hämtat vilken information när referenser skrivs på rad (exempelvis ”[8, 9, 10]” i avsnitt 2.7) och även om det är behändig kortform för att göra texten mindre hackig så kan man eventuellt skriva varje referens vid rätt plats, dock är det inget krav.

2.8 Övriga kommentarer

Beskrivningen av hur verktyget TGRConv används (kommandoradsgränssnitt) som förekommer i avsnitt 3.4 hade kanske gjort sig bättre i en bilaga. Bilagan hade kunnat beskriva steg för steg ett användningsfall av alla verktyg, från konvertering av vägdata till visualisering av simulering. Eventuellt hade en bilaga kunnat ägnas åt att förklara själva simuleringen – hur den går till.

Simuleringsdelen känns en aning förenklad, och det är svårt att bilda sig en uppfattning om kopplingen mellan mobila ad hoc-nätverk och ”punkter” som rör sig i Nam till en början. När man har läst igenom rapporten 2-3 gånger börjar frågetecknen klaras ut.

3 Genomgång av uppsatsen kapitelvis

3.1 Kapitel 1

Kapitlet ger en kort introduktion till arbetet och förklarar uppsatsens innehåll. Kapitlet kunde ha haft mer tydliga mål (vad exakt ska göras) och syften (varför är nätverkssimulering intressant). Det beskrivs i andra stycket att examensarbetet fokuserar på simulering av mobila ad hoc-nätverk, vilket är missvisande - rapportens vikt har lagts på utveckling av verktyg för visualisering av simulering. Då är istället efterkommande mening mer passande som beskrivning av arbetet. Det skrivs i samma stycke att ”we will convert the Google Earth data to a mobility model”, vilket borde skrivas om så att det tydligare framgår att konvertering sker så att data kan läsas av mobilitetsmodellen. Det skrivs även om noder, och det är oklart vilken typ av noder som avses – nätverksnoder? En figur som beskriver arbetet översiktligt, exempelvis alla ingående verktygs relationer, skulle underlätta förståelsen.

3.2 Kapitel 2

Kapitlet presenterar bakgrundsmaterial som är nödvändigt för att förstå resten av arbetet. Beskrivningarna av varje enskilt område, filformat eller verktyg är bra – i synnerhet med figurer och exempel. Dock känns dessa stycken lite avhuggna i från varandra. Efter presentationen av trådlösa nätverk i allmänhet och ad hoc-nätverk i synnerhet kastas läsaren abrupt in i en beskrivning av märkspråket XML. Även om man i efterhand förstår kopplingen så hade det varit lämpligt med ett eller flera mellanliggande avsnitt som gör övergången mer naturlig. Här skulle simulering av trådlösa nätverk kunnat presenteras i generella termer.

Figur 2 innehåller ogiltig XML-kod, vilket bör åtgärdas (salary-taggen). I avsnitt 2.5 står det att "it is also possible to store some information in KML-files", vilket skulle kunna preciseras en aning. Vilken information? Figur 4 visar en bild av programmet Nam som visualiserar en simulering. Figuren (eller programmet Nam) borde kunnas förklaras ytterligare. Att endast påpeka att en samling noder rör sig åt olika håll gör det svårt att se någon koppling till simulering av trådlösa nätverk.

3.3 Kapitel 3

Kapitlet fokuserar på konvertering av data från TIGER-databasen till KML-formatet, och utvecklingen av verktyget för att göra detta. Kapitlet har en tydlig röd tråd och man får som läsare bra förståelse för innehållet. Avsnitt 3.2 innehåller dock en del frågetecken. Definitionen av en komplett kedja (complete chain) som refereras är svår att förstå. Förbryllande då den i sin tur använder sig av ny, för läsaren kanske obekant terminologi (vänster- och högerpolygoner). Det vore bra om författarna hade skrivit en egen, enklare, förklaring efteråt – kanske illustrerad med en bild. Senare i samma avsnitt nämns "the TIGER/Line ID", som hade kunnat förklaras ytterligare redan här.

När KML-exemplet i figur 5 beskrivs så motiverar man sättandet av *extrude* till 1 med att "the line should be extended to the ground", vilket inte förklaras. En sista kommentar gäller avsnittet 3.3 där det står att vissa fält i filen som läses in inte är separerade, eller att fälten är uteblivna. Det får det att låta som att det är omöjligt att utläsa informationen, vilket förmodligen inte avses. Man blir då som läsare intresserad av hur detta problem kan lösas.

3.4 Kapitel 4

Kapitlet fokuserar på modifikation av mobilitetsmodellen. Bortsett från några oklarheter i början är kapitlet ganska välskrivet. I avsnitt 4.1, första stycket, så beskrivs mobilitetsmodellen, som består av tre delar, ”random waypoint and random walk with wrapping and reflection”, ”restricted random waypoint on a city section” och till sist ”random waypoint on a generalized domain” – dessa hade kunnat förklaras mer (vad de gör och om de påverkar varandra, exempelvis). I samma stycke skrivs också att ”a random waypoint model picks a path in a set of paths, according to some algorithm, and upon reaching its end chooses a new one”. Det vore bra om algoritmen kunde preciseras bättre. Och hur väljs en ny väg? Är det en angränsande väg? Är det noder som traverserar dessa vägar?

I andra stycket skrivs ”the models we are using”, vilket är förvirrande då det bara är *en* modell som används (se slutet av samma stycke). Dessutom skriver författarna att ”the models have perfect sampling, which means that they start in a steady state, avoiding any initial deviations”, vilket är svårt att förstå utan vidare förklaring.

I stycke tre nämns fordon (vehicles) i simuleringen, vilket inte tagits upp tidigare. I avsnitt 4.3 skrivs att ”throughout the model we handled data as doubles ... and not floats as was used in the mobility model”. Det är oklart vilken ”modell” som åsyftas först i meningen.

I avsnitt 4.3 nämns ett avrundningsfel, där -86.766319 och -86.766314 avrundades till samma tal, nämligen 899. Det talet känns väldigt främmande då det är en rejäl avrundningsmiss som har skett ifall -86.766319 kan bli 899. Stämmer siffrorna här? Om man syftar på att talet blir 899 efter anpassningen till koordinatsystemet bör detta påpekas för att undvika missförstånd.

3.5 Kapitel 5

Kapitlet beskriver simulering i ns-2 och problem som uppstod. I första meningen i första stycket bör ”input” ändras till ”output”, för tydlighets skull. I första stycket i avsnitt 5.1 står det att ”When the simulation is complete you have among other things an output file that is a NAM file...”, här skulle det kunna tänkas vara intressant att få veta vilka andra utmatningar som gavs, och varför dessa inte tillförde någonting.

Det borde förklaras varför noderna i figur 12 har 2d-koordinater, medan rörelserna i figur 13 görs i tre dimensioner (om figuren har tolkats rätt).

I det näst sista stycket i 5.2 så nämnas att Dr. Andreas Kassler genomförde en simulering med den fil som man författarna hade misslyckats med att köra i simulatorn. Här skulle man som läsare vilja veta mer om vad han gjorde och varför inte studenterna kunde göra detta.

3.6 Kapitel 6

Kapitlet avrundar rapporten och beskriver resultatet, problem och framtida vidareutvecklingar. Här nämns ”målen” igen, även denna gång utan att beskriva målen. Det enda som nämns är vad man inte hann med, inte vad det var tänkt att man skulle hinna med. I andra avsnittet skrivs ”vehicles, or nodes”. Indikerar detta att noderna alltså är fordon? Man talar även om ”road placement”, vilket förmodligen är en felskrivning (”node placement”).

3.7 Övriga kommentarer

I bilaga A presenteras en del av koden av den modifierade mobilitetsmodellen. Koden borde kunna formateras om en aning för att underlätta för läsaren av rapporten. Koden tycks även innehålla en del onödiga kommentarer, kanske av misstag kvarstående sedan testning.

4 Slutliga kommentarer

Efter att ha läst igenom uppsatsen ett par gånger börjar det mesta falla på plats. Det hade varit önskvärt med lite mer fokus på själva nätverkssimuleringen, även om det inte ingick någon sådan i själva arbetet. ”Noderna” som visas i utskriften från Nam är nämligen svåra att knyta an till just simulering av mobila ad hoc-nätverk. Och det är lite oklart vad dessa noder egentligen representerar. Fordon som åker runt i ett stadsnät? Som läsare kan man dock anta att noderna förmodligen kan användas för att simulera trådlösa nätverk genom att göra inställningar i ns-2. Kanske innehåller varje simulerat fordon en enhet som kan koppla upp sig mot mobila ad hoc-nätverk? Detta skulle kunna förklaras tydligare i kapitel 2 och 5.

I övrigt är rapporten bra, lättläst och med många exempel, och den ger en ganska tydlig bild av författarnas arbete.