

# BENEFITS OF APPLYING COLLABORATIVE FILTERING ON STEAM PLATFORM – AN EXPLORATIVE STUDY

Examensarbete Systemarkitekturutbildningen

Martin Bergqvist  
Jim Glansk

HT 2016:KSAIXX



UNIVERSITY OF BORÅS  
SCHOOL OF BUSINESS AND IT

**Systemarkitekturutbildningen** är en kandidatutbildning med fokus på programutveckling. Utbildningen ger studenterna god bredd inom traditionell program- och systemutveckling, samt en spets mot modern utveckling för webben, mobila enheter och spel. Systemarkitekten blir en tekniskt skicklig och mycket bred programutvecklare. Typiska roller är därför programmerare och lösningsarkitekt. Styrkan hos utbildningen är främst bredden på de mjukvaruprojekt den färdige studenten är förberedd för. Efter examen skall systemarkitekter fungera dels som självständiga programutvecklare och dels som medarbetare i en större utvecklingsgrupp, vilket innebär förtrogenhet med olika arbetssätt inom programutveckling.

I utbildningen läggs stor vikt vid användning av de senaste teknikerna, miljöerna, verktygen och metoderna. Tillsammans med ovanstående teoretiska grund innebär detta att systemarkitekter skall vara anställningsbara som programutvecklare direkt efter examen. Det är lika naturligt för en nyutexaminerad systemarkitekt att arbeta som programutvecklare på ett stort företags IT-avdelning, som en konsultfirma. Systemarkitekten är också lämpad att arbeta inom teknik- och idédrivna verksamheter, vilka till exempel kan vara spelutveckling, webbapplikationer eller mobila tjänster.

Syftet med examensarbetet på systemarkitekturutbildningen är att studenten skall visa förmåga att delta i forsknings- eller utvecklingsarbete och därigenom bidra till kunskapsutvecklingen inom ämnet och avrapportera detta på ett vetenskapligt sätt. Således måste de projekt som utförs ha tillräcklig vetenskaplig och/eller innovativ höjd för att generera ny och generellt intressant kunskap.

Examensarbetet genomförs vanligen i samarbete med en extern uppdragsgivare eller forskningsgrupp. Det huvudsakliga resultatet utgörs av en skriftlig rapport på engelska eller svenska, samt eventuell produkt (t.ex. programvara eller rapport) levererad till extern uppdragsgivare. I examinationen ingår även presentation av arbetet, samt muntlig och skriftlig opposition på ett annat examensarbete vid ett examinationsseminarium. Examensarbetet bedöms och betygssätts baserat på delarna ovan, specifikt tas även hänsyn till kvaliteten på eventuell framtagna mjukvara. Examinator rådfrågar handledare och eventuell extern kontaktperson vid betygssättning.



HÖGSKOLAN I BORÅS  
INSTITUTIONEN HANDELS- OCH IT-HÖGSKOLAN

BESÖKSADRESS: JÄRNVÄGSGATAN 5 · POSTADRESS: ALLÉGATAN 1, 501 90 BORÅS  
TFN: 033-435 40 00 · E-POST: INST.HIT@HB.SE · WEBB: WWW.HB.SE/HIT

**Svensk titel:** Fördelar med att applicera Collaborative Filtering på Steam – En utforskande studie

**Engelsk titel:** Benefits of Applying Collaborative Filtering on Steam – An explorative study

**Utgivningsår:** 2016

**Författare:** Martin Bergqvist, Jim Glansk

**Handledare:** Henrik Linusson

**Abstract**  
(på engelska)

**Keywords:** Collaborative filtering, Recommender system, Steam.

## **Sammanfattning**

(på svenska)

**Nyckelord:** Collaborative filtering, Rekommendationssystem, Steam.

## Innehållsförteckning

1	Inledning.....	- 1 -
2	Problemdiskussion.....	- 2 -
3	Problemformulering & Syfte .....	- 2 -
4	Referenser.....	- 3 -

# 1 Inledning

Datorspel har blivit en djupt rotad företeelse i vår kultur vilket tydligt visas genom att spelindustrin ständigt växer sig större. Spelindustrin omsatte 91.8 miljarder USD 2015 och växer mellan 6-9 procent varje år (Newzoo 2016). En av anledningarna till framgången är digital distribution och ökad bandbreddskapacitet, vilket gör att spel på ett enkelt sätt kan köpas och laddas ner över nätet (Toivonen & Sotamaa 2010).

Till skillnad från andra e-Tjänster som NetFlix, Spotify, Amazon, YouTube etc. som distribuerar annan digital media, så ligger digital spel-distribution efter i hänseende att predicera bättre rekommendationer för sina kunder. Det finns flera tusen speltitlar att välja mellan vilket skapar svårigheter i att finna ett spel som passar specifika användare. En av de största aktörerna på marknaden idag är Steam, som har över 125 miljoner aktiva användarkonton (Kotaku 2015), och över 10000 speltitlar (Steam 2016). Steam hade 2011 50-70% av marknaden för digital spelförsäljning till PC (Forbes 2011).

Under 2015 hade spel för 3,5 miljarder USD sålts via Steam. Ca. 25% av spelen som såldes har aldrig spelats (Medium 2016), och ytterligare 19% av spelen har spelarna ägnat inte mer än en timmas speltid åt (Ars Technica, 2014).

Det finns mycket forskning gjord på hur man optimerar rekommendationer åt användare, men är då sällan riktad mot spel, utan andra medietyper, såsom film, böcker & musik.

Rekommendationssystem är till för att hitta relevanta biprodukter för den specifika användaren, eller, om man så vill, en maskininlärningsmetod för direktriktad reklam.

Det första rekommendationssystemet som använde termen *collaborative filtering* (CF) var Tapestry (Goldberg 1992). Idag används termen något annorlunda, och har inte ett givet samband med rekommendationssystem. Anledningen till att skilja termerna åt är att rekommendationssystem kan vara byggda utan CF, och då vanligtvis antingen *Content-based filtering* (CBF); Att filtrera efter liknande produkter, eller ha fler aspekter än vad som ryms i CF (Hybrid mellan CF och CBF) .

Rekommendationssystem tittar på flera sorters input. Där det finns, är explicit feedback det mest användbara; Om användaren värderar produkten ger det mycket pålitlig data. I de fall där explicit feedback saknas, kan istället implicit feedback nyttjas, som indirekt påvisar korrelation genom observation av användarens beteendemönster (Oard 1998).

CBF går enkelt uttryckt ut på att klassificera alla produkter så noggrant som möjligt och väga dessa taggar mot användarens angivna och/eller historiska preferenser, för att rekommendera de produkter som liknar de användaren gillat tidigare. Steam använder sig av detta system i grunden för sina rekommendationer, där användarens köpta spel, kombinerat med användarens vänners spel, utgör basen för rekommendationerna (Steam 2014).

Fördelarna med det här systemet är att resultaten är enkla för användaren att förstå, det är billigt att använda gällande beräkningsmängder, och det är möjligt med personliga resultat även utan en användarprofil. Nackdelarna å sin sida blir att resultaten blir ytliga och uppenbara (Shi 2008).

CF arbetar utifrån principen att liknande användare gillar liknande saker, och analyserar beteendemönster för att sedan hitta de användare som ligger närmast i beteende, och använda deras preferenser som rekommendationer. LinkedIn är ett bra exempel, där du rekommenderas nya företag att följa baserat på vad liknande profiler följer (Wu, Shah, Choi, Tiwari & Posse 2014). Fördelarna med CF är att det fungerar mot vilken domän som helst & “lyckoträffar”, när CF hittar till synes orelaterat material av intresse. Nackdelarna rör främst skalbarheten, då en liten databas kommer att ge svaga kopplingar; allmänt kallat *kallstart*, och en omfattande databas kommer ge starka kopplingar, men använda mer beräkningsresurser (Shi 2008). Det anses allmänt vara så, att CF är effektivare än CBF på befintliga datamängder, där man kan bortse från kallstarts-aspekten.

Hybrid Filter är ett samlingsnamn för de system som på något sätt kombinerar både CBF & CF. Detta går att göra på många olika sätt, vilket Netflix är ett utmärkt exempel på, då de 2007 använde en lösning baserad på 107 algoritmer för att rekommendera film (Bell, Koren & Volinsky 2007). Fördelen med att kombinera olika lösningar blir att man kombinerar styrkorna från varje delösning, och kompenserar i viss mån svagheter (Shi 2008).

Netflix annonserade 2006 en tävling, där man lovade 1 miljon USD i prispengar till den som lyckades finna ett sätt att optimera deras rekommendationssystem med 10 procent. När man avslutade tävlingen 2009 hade man nått en optimeringsgrad på 10.6 procent och engagerat över 41305 teams från 186 länder (Netflix Prize 2016).

## 2 Problemdiskussion

Stears plattform använder i nuläget content-based filtering genom explicit data och forskningen som är gjord pekar på att collaborative filtering är en mer mångsidig och exakt metod för att utvärdera mer implicit data mellan olika användare. Andra tjänster för digital distribution som t.ex. Netflix, Amazon, Youtube använder sig av mer komplexa rekommendationssystem för att skapa bättre rekommendationer för sina kunder.

Med hänsyn till detta, i samband med det faktum att en stor andel av spelen som säljs via Steam faktiskt inte spelas, eller spelas mycket lite, konstateras att det finns ett intresse i att undersöka potentialen i ett system som följer principer närmare State-of-the-art.

## 3 Problemformulering & Syfte

Syftet med denna rapport är att undersöka vilken nivå av pricksäkerhet implementering av CF kan ge på Steams rekommendationer.

## 4 Referenser

Newzoo (2016). *The Global Games Market Reaches \$99.6 Billion in 2016*.

<https://newzoo.com/insights/articles/global-games-market-reaches-99-6-billion-2016-mobile-generating-37> [2016-11-25]

Goldberg, D., Nichols, D., Oki, B. & Terry, D. (1992). Using collaborative filtering to weave an information tapestry, *ACM*, NEW YORK.

Oard, D.W., Kim, J. (1998). Implicit Feedback for Recommender Systems, *AAAI Technical Report WS-98-08*.

Toivonen, S. & Sotamaa, O. (2010). Digital distribution of games: the players' perspective. *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, ACM, New York, ss. 199-206.

Steam (2014). *Recommendation Feed*.

<http://store.steampowered.com/about/newstore?l=swedish> [2016-11-24]

Shi, C. (2008). Exploring Movie Recommendation System Using Cultural Metadata, *IEEE*, ss.431.

Forbes (2011). *The Master of Online Mayhem*.

<http://www.forbes.com/forbes/2011/0228/technology-gabe-newell-videogames-valve-online-mayhem.html> [2016-11-24]

Wu, L., Shah, S., Choi, S., Tiwari, M. & Posse, C. (2014). *The browsemaps: Collaborative filtering at LinkedIn*.

Bell, R. M., Koren, Y., & Volinsky, C. (2007). *The BellKor solution to the Netflix prize*.

Netflix (2009). *Netflix Prize*. <http://www.netflixprize.com/leaderboard.html> [2016-11-25]



**Högskolan i Borås** är en modern högskola mitt i city. Vi bedriver utbildningar inom ekonomi och informatik, biblioteks- och informationsvetenskap, mode och textil, beteendevetenskap och lärarutbildning, teknik samt vårdvetenskap.

På **institutionen Handels- och IT-högskolan (HIT)** har vi tagit fasta på studenternas framtida behov. Därför har vi skapat utbildningar där anställningsbarhet är ett nyckelord. Ämnesintegration, helhet och sammanhang är andra viktiga begrepp. På institutionen råder en närhet, såväl mellan studenter och lärare som mellan företag och utbildning.

Våra **ekonomiutbildningar** ger studenterna möjlighet att lära sig mer om olika företag och förvaltningar och hur styrning och organisering av dessa verksamheter sker. De får även lära sig om samhällsutveckling och om organisationers anpassning till omvärlden. De får möjlighet att förbättra sin förmåga att analysera, utveckla och styra verksamheter, oavsett om de vill ägna sig åt revision, administration eller marknadsföring. Bland våra **IT-utbildningar** finns alltid något för dem som vill designa framtidens IT-baserade kommunikationslösningar, som vill analysera behov av och krav på organisationers information för att designa deras innehållsstrukturer, bedriva integrerad IT- och affärsutveckling, utveckla sin förmåga att analysera och designa verksamheter eller inrikta sig mot programmering och utveckling för god IT-användning i företag och organisationer.

**Forskningsverksamheten** vid institutionen är såväl professions- som design- och utvecklingsinriktad. Den övergripande forskningsprofilen för institutionen är handels- och tjänsteutveckling i vilken kunskaper och kompetenser inom såväl informatik som företagsekonomi utgör viktiga grundstenar. Forskningen är välrenommerad och fokuserar på inriktningarna affärsdesign och Co-design. Forskningen är också professionsorienterad, vilket bland annat tar sig uttryck i att forskningen i många fall bedrivs på aktionsforskningsbaserade grunder med företag och offentliga organisationer på lokal, nationell och internationell arena. Forskningens design och professionsinriktning manifesteras också i InnovationLab, som är institutionens och Högskolans enhet för forskningsstödjande systemutveckling.



**HÖGSKOLAN I BORÅS**  
VETENSKAP FÖR PROFESSION

BESÖKSADRESS: JÄRNVÄGSGATAN 5 • POSTADRESS: ALLÉGATAN 1, 501 90 BORÅS  
TFN: 033-435 40 00 • E-POST: [INST.HIT@HB.SE](mailto:INST.HIT@HB.SE) • WEBB: [WWW.HB.SE/HIT](http://WWW.HB.SE/HIT)