

Metodbeskrivning

MMS modell för Totalvideo





Innehåll

Inledning	3
Historik	4
1 Datakällor	5
1.1 Videopaneldata.....	5
1.2 AAM-paneldata.....	5
2 Hantering av dubbelmätning för program	6
3 Total räckviddsmodellering	8
3.1 Koncept för total räckvidd.....	8
3.2 Överlappstargets.....	10
3.3 Panelistkopplingar	10



Inledning

Syftet med MMS modell för totalvideo är att få en heltäckande bild av Sveriges befolknings samlade mediakonsumtion på TV och via onlineenheter. Detta uppnås genom att sammanfoga konsumtionen från MMS TV-mätning med den onlinekonsumtion som tas fram genom AAM-modellen.

För skapandet av totalmodellen används två datakällor. Den ena är en single source videopanel som mäter cirka 3000 svenska hushålls konsumtion av linjär-TV, samt även onlineinnehåll som spelats upp på storskärm (big screen) och digitala enheter (smartphone, surfplatta och PC). Den andra datakällan är en virtuellt skapad onlinepanel (hädanefter kallad AAM-panelen) med cirka 46 000 panelister som tas fram via den så kallade AAM-modellen (Audience Ascription Model). I jämförelse med TV är konsumtionen online mer fragmenterad med en lång svans av mindre titlar, vilket medför att AAM-panelens storlek ger en mer heltäckande bild av den svenska befolkningens onlinetittande jämfört med videopanelen. På grund av detta utgörs totalmodellen av TV-konsumtionen från videopanelen och onlinekonsumtionen från AAM-panelen.

En utmaning med datakällorna är en mätteknisk problematik som innebär att streamat onlinetittande under vissa omständigheter även fångas upp i videopanelens TV-mätning. Med andra ord finns en dubbelmätning där innehåll som registreras som ett onlinetittande även registreras som ett TV-tittande. Inom totalmodellen hanteras detta genom att identifiera och exkludera dubbelmätt konsumtion från TV-sidan.

Utöver problemet med dubbelmätning är beräkningarna av mått relaterade till tittartid relativt enkla. Det innebär i praktiken en rak summering av de individuella tittartiderna på TV och online. Beräkning av räckvidd är något mer komplex med tanke på att en individ kan konsumera en kombination av innehåll både på TV och online. Detta innebär att en rak summering av de individuella räckvidderna skulle ge en överskattning. Eftersom det inte finns någon direkt länk mellan panelisterna i videopanelen och AAM-panelen som tar hänsyn till detta dubbelkonsumtionsmönster, hanteras detta genom en total räckviddsmodellering.

Den totala räckviddsmodelleringen sker genom att det i ett första steg, med information från videopanelen, beräknas hur stor andel av onlineräckvidden som överlappar med TV-räckvidden för ett antal givna scenarion. Dessa kallas för överlappstargets och används som underlag för att skapa kopplingar mellan panelisterna i videopanelen och AAM-panelen.

Syftet med panelistkopplingarna är att skapa en ny totalpanel där varje totalpanelist utgörs av en kombination av en panelist från videopanelen och en panelist från AAM-panelen. Genom användningen av överlappstargets kombineras då TV-tittandet från videopanelen med onlinetittandet från AAM-panelen på ett sådant sätt att den totala beräknade räckviddsnivån speglar det sanna dubbelkonsumtionsmönstret i befolkningen.

I det första avsnittet beskrivs datakällorna närmare. Sedan följer en genomgång om hur dubbelmätningen hanteras. Det sista avsnittet behandlar total räckvidd genom att först förklara konceptet djupare, för att sedan avsluta med en beskrivning av hur överlappstargets beräknats och hur panelistkopplingarna mellan videopanelen och AAM-panelen skapas.

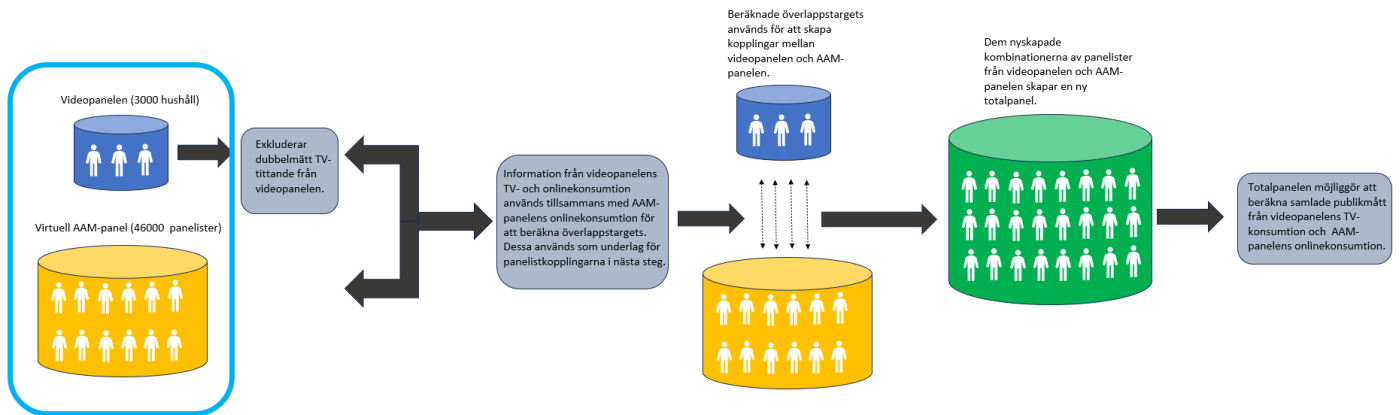


Historik

MMS lanserade Totalmodellen v2 2023-12-13 med konsumtion från dagen 2023-12-11.

Datum	Händelse
2023-12-13	Lansering av Totalmodellen v2

1 Datakällor



1.1 Videopaneldata

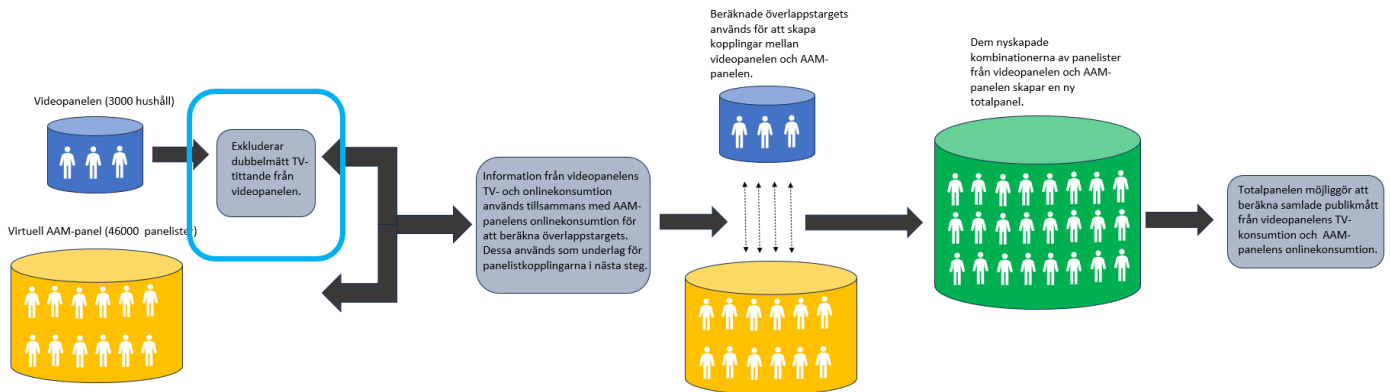
MMS mäter dagligen den svenska befolkningens TV-konsumtion genom en panel bestående av ca 3000 hushåll. Till en början mättes endast panelisternas TV-tittande, men sedan 2022-09-26 används i stället en single source videopanel för både TV och online. Detta innebär att mätningen av de hushåll som ingår i TV-mätningen, nu även inkluderar panelisternas onlinekonsumtion på storskärm (big screen) och på digitala enheter (smartphone, surfplatta och PC). Videopanelen inkluderar tillräckligt många hushåll för att täcka det tittande som konsumeras på TV, men på onlinesidan är tittandet för fragmenterat för att få en heltäckande bild av den långa svansen bestående av många mindre titlar. Men genom att videopanelen innehåller både TV- och onlinekonsumtion för samma individer, kan relevanta dubbelkonsumtionsmönster identifieras och extrapoleras till den större AAM-panelen.

Videopanelen har därmed två syften i totalen. Dels bidrar den med konsumtionen från TV-sidan, dels bidrar den med information om relevanta dubbelkonsumtionsmönster mellan TV och online.

1.2 AAM-paneldata

I maj 2021 lanserade MMS sin onlinemätning baserat på en modellering kallad AAM (Audience Ascription Model). AAM-modelleringen kommer till rätta med videopaneldatats bristfälliga täckning av onlinetittande genom att komplettera videopanelen med trafikvolym som mäts på censusedata. Detta censusedata mäts direkt från dem spelare som hårdvarutillverkare implementerat för att spela upp online innehåll och ger därmed en mer komplett bild av onlinekonsumtionen. Med hjälp av videopanelen, som följer faktiska individer, målgruppsprofileras dessa uppspelningar och kopplas sedan till nya artificiellt skapade panelister. Slutresultatet är en virtuell onlinepanel med ca 46 000 panelister. Viktigt att notera är att även om videopanelens onlinetittande används i AAM-modelleringen, finns det inte någon direkt koppling mellan videopanelens TV-konsumtion och AAM-panelisternas onlinekonsumtion. Detta är anledningen till att dubbelkonsumtionsmönstret måste modelleras explicit för att kunna beräkna korrekta totalräckvidder.

2 Hantering av dubbelmätning för program



I ett första steg måste dubbelmätningen för program hanteras i det ingående datamaterialet. Dubbelmätningen innebär att innehåll som i videopanelen skett genom streamat online tittande på TV-skärmen, även kommer fångas upp som TV-tittande. Denna dubbelmätning sker i de fall innehållet som spelats upp online, sänts på TV någon gång under de senaste 28 dagarna. Dubbelmätningen bidrar precis som dubbelkonsumtionen till en överskattad räckvidd, men medan dubbelkonsumtionen är en naturlig del av hur innehåll konsumeras, är dubbelmätningen en avvikelse som innebär att även tittartid överskattas. Med andra ord är skillnaden att dubbelkonsumtionen måste hanteras inom modellen som en naturlig del av faktiskt tittarbeteende, medan dubbelmätningen måste exkluderas från det ingående datamaterialet.

Då reklamen som visas inom och mellan program, oftast skiljer sig mellan den linjära sändningen och motsvarande innehåll online, utgör dubbelmätningen endast ett problem på programsidan och hanteras därför endast för program. Dubbelmätningen kan rent teoretiskt exkluderas från antingen TV-sidan eller onlinesidan, men det finns huvudsakligen två argument till att det mest lämpliga är att exkludera konsumtionen på TV-sidan¹. Det ena är att källan till innehållet faktiskt är en online plattform och det andra är att den större fragmentiseringen av online innehåll innebär att mer information går förlorad om exkluderingen sker på onlinesidan.

Dubbelmätningen hanteras genom att använda en kombination av tillgänglig information från videopanelens konsumtion i form av streamingmätare på TV-sidan och mätning av diverse onlineenheter på onlinesidan. Vi har två kriterier som kan användas för att identifiera dubbelmätt TV-tittande.

Kriterium 1 är att ett tittarpass på TV matchar med ett tittarpass online genom att

- Tittarpassen överlappar tidsmässigt.
- Tittarpassen sker på samma TV-enhet.
- Tittarpassen har samma registrerade titelkod.

¹ Ett undantag från detta är linjära livesändningar (s.k. simulcast) som konsumerats via onlineplattformar. Denna typ av tittande filteras bort från onlinekonsumtionen och ska därför behållas på TV-sidan.



Kriterium 2 är att ett tittarpass på TV har en domän registrerad via streaming-mätare som matchar respektive kanalhus egen site-domän för onlineinnehåll.

Dessa två kriterier ska egentligen alltid stämma överens, men på grund av mättekniska skäl kan det ibland vara så att egentliga matchningar inte hittas mellan tittarpassen online och tittarpassen på TV, samt att streamingmätaren inte alltid registrerat en site-domän trots att tittandet faktiskt skett online.

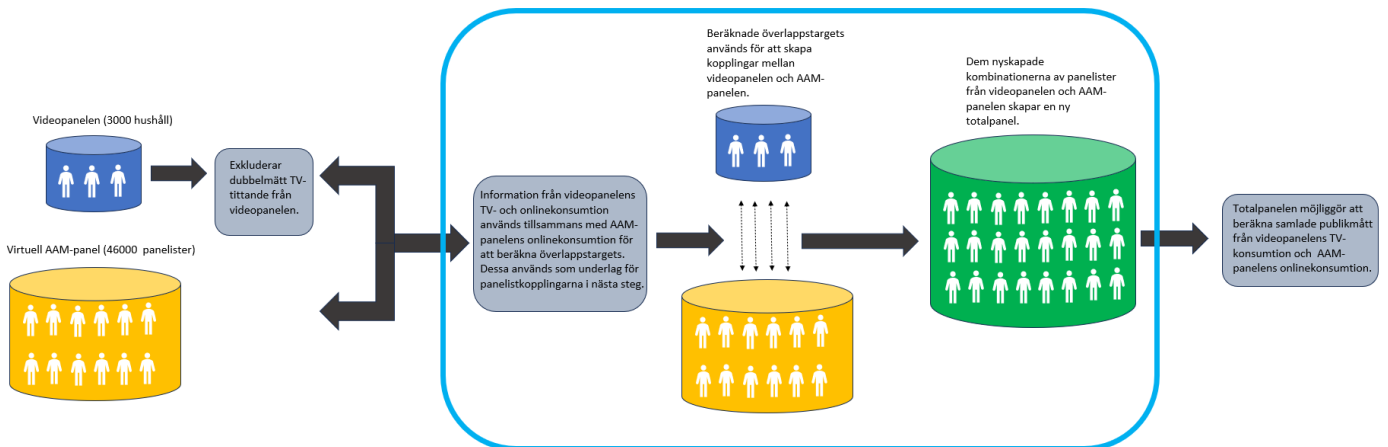
Hur dessa kriterier kan användas är beroende av om TV-tittandet skett under sändningsdag eller inte. Tittande som sker live eller timeshiftat inom sändningsdagen kan endast identifieras med kriterium 1 då visst TV-tittande skett via linjär feed och därför ska behållas på TV-sidan (se tidigare fotnot).

Tittarpass som skett på annan dag än sändningsdag kan däremot identifieras med båda kriterierna, det vill säga att det räcker med att ett av kriterierna är uppfylla för att tittandet ska märkas ut som att vara dubbelmätt.

Ovanstående summeras i tabellen nedan.

	Kriterium 1: Tittarpass på TV matchar med ett tittarpass online genom att <ul style="list-style-type: none"> • Tittarpassen överlappar tidsmässigt. • Tittarpassen sker på samma TV-enhet. • Tittarpassen har samma registrerade titelkod. 	Kriterium 2: Tittarpass på TV har en registrerad domän via streaminmätare som matchar respektive kanalhus egen site-domän för onlineinnehåll.
Live och timeshift under sändningsdag	X	
Timeshift dag -1 till -28	X	X

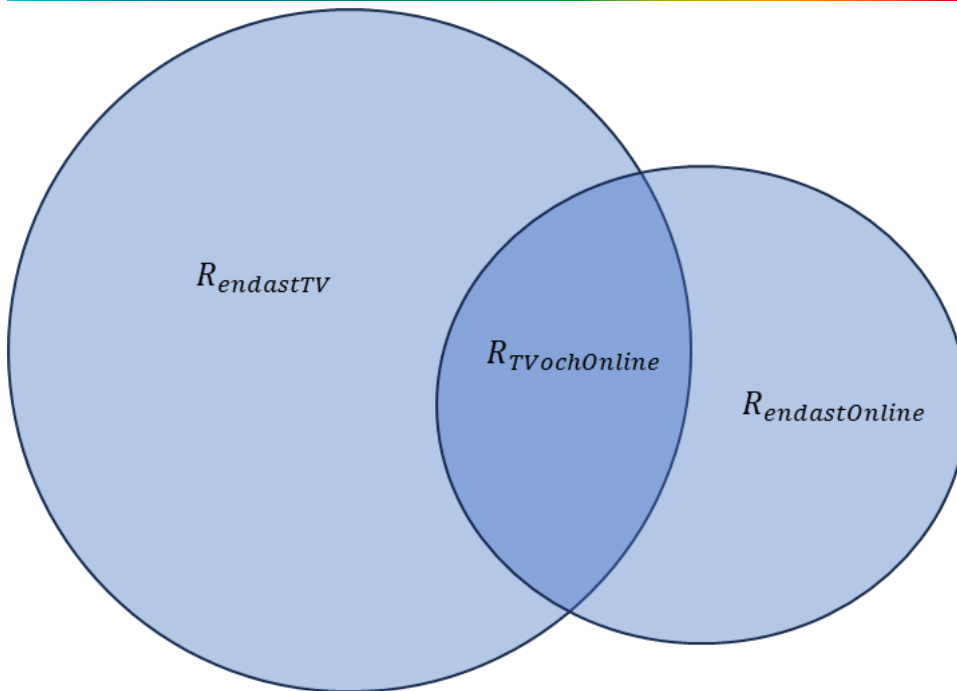
3 Total räckviddsmodellering



3.1 Koncept för total räckvidd

När dubbelmätningen identifierats och exkluderats, kan den totala tittartiden för ett innehåll adderas ett till ett. Syftet med den totala räckviddsmodelleringen är därför att ta hänsyn till dubbelkonsumtionen genom att matcha TV- och onlinekonsumtion på ett sätt som ger representativa totala räckviddsskattningar givet olika aggregeringsscenarion. Ett aggregeringsscenario definieras som en aggregering av ett innehåll i form av en eller flera titlar över en given tidsperiod. Dessa scenarion avser alltså olika kombinationer av innehåll som exempelvis i fallet med program tillhör samma serie, i andra fall tillhör olika serier men samma kanalhus, och i något ytterligare fall kanske är helt orelaterade.

Räckvidden kan då beskrivas som unionen av dem individer som via TV eller online enhet, konsumerat någon del av det innehåll som avses. Detta illustreras med figuren nedan.



Figur1. Antalet individer som endast konsumerat innehåll på TV illustreras av area $R_{endastTV}$, individer som endast konsumerat innehåll online illustreras av area $R_{endastOnline}$ och individer som konsumerat innehåll på både TV och online illustreras av area $R_{TVochOnline}$.

Figur 1 illustrerar räckviddssambandet mellan tv-konsumtion och onlinekonsumtion. Den totala räckvidden R_{Total} kan då beräknas som

$$R_{Total} = R_{endastTV} + R_{endastOnline} + R_{TVochOnline}$$

Det är viktigt att poängtera att $R_{TVochOnline}$ vid aggregering av innehåll inte nödvändigtvis endast består av individer som konsumerat exakt samma titel eller spot, utan även kan bestå av personer som konsumerat en titel på TV och en annan titel online. Detta innebär att, desto mer titlar eller spottar som ingår i den valda aggregeringen, desto större sannolikhet är det att en individ kommer tillhöra $R_{TVochOnline}$.

Inför modelleringen finns för ett givet aggregeringsscenario information tillgänglig om storleken på $R_{TV} = R_{endastTV} + R_{TVochOnline}$ som utgörs av tv-räckvidden från videopanelen, samt $R_{Online} = R_{endastOnline} + R_{TVochOnline}$, som utgörs av onlineräckvidden från AAM-panelen. Totalräckvidden kan då skrivas om som

$$R_{Total} = R_{TV} + R_{Online} - R_{TVochOnline}$$

Vidare kan detta skrivas om som

$$R_{Total} = R_{TV} + R_{Online} - O_{TV} * R_{Online}$$

där $O_{TV} = \frac{R_{TVochOnline}}{R_{Online}}$ och alltså kan definieras som andelen av onlineräckvidden som överlappar med TV-räckvidden. Från respektive datakälla har vi för ett givet aggregeringsscenario skattningar av R_{TV} och R_{Online} , vilket innebär att problemet främst reduceras till att hitta representativa skattningar av överlappet O_{TV} .

Detta sker i två steg. Först används information om dubbelkonsumtionsmönstret från videopanelen för att skatta överlappstargets O_{TV}^* för ett antal givna aggregeringsscenarion. Dessa överlappstargets används sedan som underlag



för att skapa kopplingar mellan videopanelister och AAM- panelister. Denna process förklaras utförligare i nästa avsnitt.

3.2 Överlappstargets

För att bevara det dubbelkonsumtionsmönster som kan tydas inom videopanelen, beräknas ett antal överlappstargets O_{TV}^* som sedan används som underlag när videopanelen och AAM-panelen kopplas samman.

Det går inte att explicit ta hänsyn till varje aggregeringsscenario som kan tänkas uppstå, därför har ett mer begränsat antal scenarion definierats på förhand. Förväntningen är då att andra dubbelkonsumtionsmönster som inte explicit modellerats för, hanteras implicit genom att de korrelerar med dem aggregeringsscenarion som inkluderats i modellen. I nedanstående tabell visas vilka aggregeringsscenarion det beräknas explicita överlappstargets för.

	1 dagar	7 dagar	14 dagar	28 dagar
Kanalhus (program och reklam)	X	X	X	
Kanalhus (program)	X	X		X
Kanalhus (reklam)	X		X	X
Enskilda program (titelkoder)	X			
Programserier		X	X	X
Enskilda reklamer (filmkoder)	X			
Reklamkampanjer	X	X	X	

Varje aggregeringsscenario motsvarar alltså ett aggregerat innehåll i , i kombination med tidsperiodens längd bakåt i tid t och beräknas som

$$O_{TV_{i,t}}^* = \frac{R_{TVochOnline_{i,t}}^*}{R_{Online_{i,t}}^*}$$

Exempelvis beräknas det då från videopanelen hur stor del av onlineräckvidden på kanalhusnivå som inom varje målgrupp och tidsperiod även konsumerat respektive kanalhus innehåll på TV. Dessa överlappstargets används sedan som mål värden för liknande aggregeringsscenarion där onlinedelen i stället motsvaras av onlinekonsumtionen från AAM-panelen.

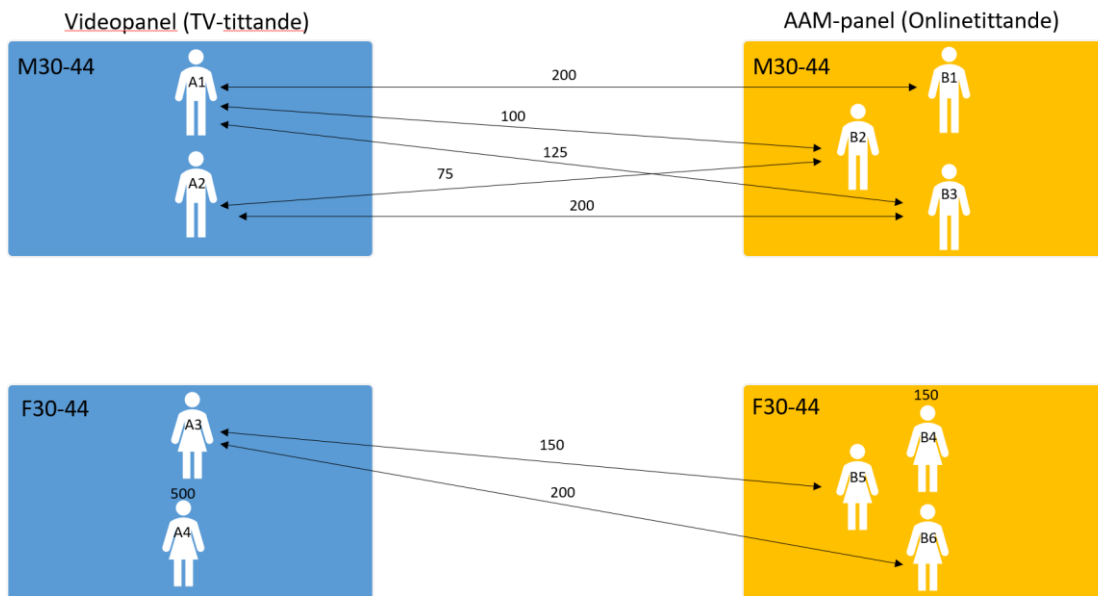
3.3 Panelistkopplingar

Tillvägagångssättet för beräkning av totalräckvidder grundar sig i att skapa kopplingar mellan panelister från videopanelen och panelister från AAM-panelen. Det är videopanelens vikter som utgör basen för den resulterande totalpanelens vikter, vilket innebär att TV-räckvidden förblir helt oförändrad. Man kan därför se det som att andelar av AAM-panelisterna fördelas till videopanelisterna. Detta innebär att onlineräckvidderna kan få mycket små förändringar jämfört med den initiala onlineräckvidd som beräknats från AAM².

² Detta är dock inte något som påverkar de siffror som ses i Onlineanalys då det är onlinesiffrorna från Totalmodellen som slutligen visas även där.

Syftet är att panelistkopplingarna ska ske på ett sådant sätt att dem tillsammans ska generera ett konsumtionsmönster som representerar det sanna konsumtionsmönstret i Sveriges befolkning. Detta uppnås genom att kopplingarna, med hjälp av en optimeringsfunktion, skapas med hänsyn tagen till de överlappstargets O_{TV}^* som beräknats enligt föregående avsnitt.

För att ta hänsyn till ett så flexibelt konsumtionsmönster som möjligt kan varje videopanelist kopplas till flera AAM-panelister och varje AAM-panelist kan i sin tur även kopplas till flera videopanelister. Kopplingar mellan panelister kan endast ske inom respektive målgrupp. Detta illustreras nedan i figur 2.



Figur 2. Visar hur panelister från videopanelen kopplas ihop med panelister från den AAM-panelen.

Varje koppling får här en egen vikt så exempelvis har panelist A1 i videopanelen en TV-konsumtion som motsvaras av en vikt av 425 individer i populationen. Av denna vikt på 425 representerar 200 denna panelists TV-konsumtion i kombination med onlinekonsumtionen från panelist B1. På liknande sätt har onlinekonsumtionen från panelist B2 en samlad vikt på 175, varav en vikt på 100 är i kombination med TV-konsumtionen från A1 och 75 är i kombination med TV-konsumtionen från A2. Det finns såklart även panelister som endast konsumerar TV respektive online, vilket representeras av panelist A4 och B4. Då kopplingarna baseras på videopanelens vikter, innebär detta att en summering av alla kopplingar relaterade till en videopanelist, resulterar i den panelists ursprungliga tv-vikt.

Sammantaget innebär detta att de olika kombinationerna, av TV-panelister från videopanelen och onlinepanelister från AAM-panelen, kan ses som nya "totalpanelister". Dessa totalpanelister kan nu användas för att beräkna tittarmått på totalnivå.