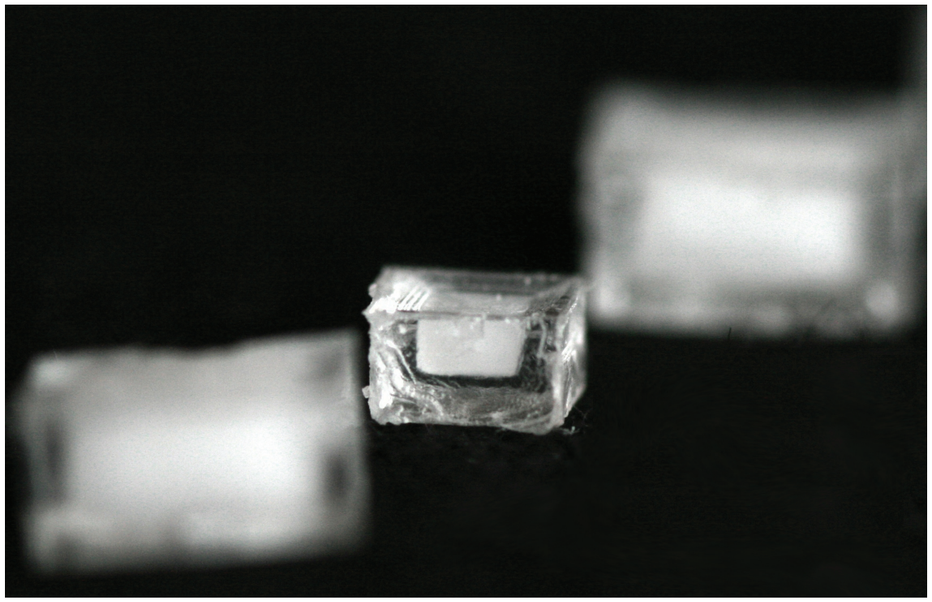
Framtida metoder för värdering av IT-säkerhet

JOHAN BENGTSSON, JONAS HALLBERG,
KRISTOFFER LUNDHOLM



FOI är en huvudsakligen uppdragsfinansierad myndighet under Försvarsdepartementet. Kärnverksamheten är forskning, metod- och teknikutveckling till nytta för försvar och säkerhet. Organisationen har cirka 1000 anställda varav ungefär 800 är forskare. Detta gör organisationen till Sveriges största forskningsinstitut. FOI ger kunderna tillgång till ledande expertis inom ett stort antal tillämpningsområden såsom säkerhetspolitiska studier och analyser inom försvar och säkerhet, bedömning av olika typer av hot, system för ledning och hantering av kriser, skydd mot och hantering av farliga ämnen, IT-säkerhet och nya sensorers möjligheter.

FOI
Totalförsvarets forskningsinstitut
Informationssystem
Box 1165
581 11 Linköping

Tel: 013-37 80 00
Fax: 013-37 81 00

www.foi.se

FOI-R--3094--SE
ISSN 1650-1942

Användarrapport
December 2010

Informationssystem

Johan Bengtsson, Jonas Hallberg,
Kristoffer Lundholm

Framtida metoder för värdering av IT-säkerhet

Titel	Framtida metoder för värdering av IT-säkerhet
Title	Future methods for IT security assessment
Rapportnr/Report no	FOI-R--3094--SE
Rapporttyp Report Type	Användarrapport / User report
Månad/Month	December
Utgivningsår/Year	2010
Antal sidor/Pages	56 p
ISSN	ISSN 1650-1942
Kund/Customer	Försvarsmakten
Projektnr/Project no	E53077
Godkänd av/Approved by	Magnus Jändel

FOI, Totalförsvarets Forskningsinstitut
Avdelningen för Informationssystem
Box 1165
581 11 Linköping

FOI, Swedish Defence Research Agency
Information Systems
Box 1165
SE-581 11 Linköping

Sammanfattning

Användandet av informationssystem inom Försvarsmakten ökar samtidigt som systemen blir allt mer omfattande, mer sammankopplade och hanterar mer verksamhetskritisk information. Detta leder till ökade behov av att säkerställa att systemen håller en tillräckligt hög nivå av IT-säkerhet. Olika ansatser till att avgöra IT-säkerhetsnivåer passar mer eller mindre bra under de olika stegen i Försvarsmaktens IT-livscykelmodell.

För att skapa en tydlig bild av vilka typer av framtida IT-säkerhetsvärderingsmetoder som kan vara relevanta för Försvarsmakten genomfördes en litteraturstudie. Litteraturstudien resulterade i ett antal artiklar vilka beskriver olika angreppssätt för att avgöra IT-säkerhetsnivån i ett system. Liknande angreppssätt grupperades ihop, vilket resulterade i fyra metodgrupper. Dessa fyra metodgrupper testades med hjälp av testproceduren TSAR. För att få en indikation på relevansen för Försvarsmakten testades varje metodgrupp gentemot behoven av IT-säkerhetsvärdering under varje steg i IT-livscykelmodellen. Behoven av IT-säkerhetsvärdering identifierades med hjälp av en enkätundersökning.

De genomförda testerna resulterade i ett mått på hur relevant varje metodgrupp anses vara för Försvarsmakten under varje steg i IT-livscykelmodellen. Utifrån testresultatet ges en rekommendation angående vilka typer av IT-säkerhetsvärderingsmetoder som kan vara av intresse för Försvarsmakten i framtiden.

Nyckelord: IT-säkerhet, värdering, IT-livscykelmodell, TSAR

Summary

The use of information systems within the Swedish Armed Forces is increasing while the systems are becoming larger, more interconnected, and manages more mission-critical information. This leads to increasing needs to ensure that the systems have a sufficient level of IT security. Different approaches to determining the IT security levels are more or less appropriate during the various stages of the IT lifecycle model used by the Swedish Armed Forces.

To create a clear picture of what types of future IT security assessment methods that may be relevant to the Swedish Armed Forces a literature study was conducted. The literature study resulted in a number of articles that describe different approaches to determine the IT security level of a system. Similar approaches were grouped together, resulting in four groups of methods. These four method groups were tested using the test procedure TSAR. To get an indication of the relevance to the Swedish Armed Forces, each method group was tested against the needs of IT security assessment at each stage of the IT life cycle model. The needs of IT security assessment were identified through a survey.

The tests resulted in a measure of how relevant each method group is considered to be for the Swedish Armed Forces during the steps of the IT life cycle model. Based on the test results, recommendations are provided regarding what types of IT security assessment methods that may be of interest to the Swedish Armed Forces in the future.

Keywords: IT security, assessment, lifecycle model, TSAR

Innehåll

1 Inledning

I takt med att informationssystemen blir allt mer omfattande, mer sammankopplade och hanterar mer verksamhetskritisk information, ökar behoven av ändamålsenlig IT-säkerhet. IT-säkerhet är en kvalitet hos informationssystem som beror på det aktuella systemets design, realisering, konfiguration, administration och användning. Det finns alltså många faktorer som påverkar IT-säkerheten hos informationssystem. Dessutom är säkerhet en subjektiv egenskap och de underliggande faktorerna prioriteras olika av olika intressenter. För att kunna fatta rationella beslut relaterade till IT-säkerheten hos informationssystem är det nödvändigt med en förmåga att avgöra hur dessa påverkar IT-säkerheten. Detta kräver bedömningar av nuvarande och framtida säkerhetsnivåer. För att minska subjektiviteten hos dessa bedömningar av IT-säkerhetsnivåer är det nödvändigt att nyttja strukturerade metoder för värdering av IT-säkerhet.

1.1 Syfte och målsättning

Det övergripande syftet med det arbete som presenteras i denna rapport är att beskriva de resultat som den pågående forskningen och utvecklingen avseende metoder för värdering av IT-säkerhet kan förväntas leda till. Mer konkret är målet att identifiera ansatser till framtida metoder för värdering av IT-säkerhet som kan förväntas bära frukt samt avgöra vilken relevans de kan få för de olika stegen i Försvarsmaktens IT-livscykelmodell. Dessutom ska det klargöras vilka övergripande egenskaper dessa metoder kommer att ha samt vilka typer av resultat som de kan förväntas kunna leverera.

1.2 Metod

För att uppfylla syftet och nå de uppsatta målen genomfördes en studie med följande delaktiviteter.

- Litteraturstudie inriktad mot nyligen publicerade arbeten för att identifiera vad det för närvarande forskas om inom området
- Gruppering av identifierade publikationer för att forma metodgrupper, vilka kan ses som inriktningar inom området
- Enkätstudie för att identifiera vilka behov relaterade till värdering av IT-säkerhet som kan kopplas till vart och ett av de sju stegen i Försvarsmaktens IT-livscykelmodell
- Testning av metodgrupperna för att avgöra deras relevans kopplad till de sju stegen i Försvarsmaktens IT-livscykelmodell

1.3 Bidrag

De resultat som beskrivs i denna rapport inkluderar:

- Formulering av fyra stycken metodgrupper som övergripande beskriver de ansatser till framtida metoder som har identifierats
- Sju uppsättningar med behov relaterade till IT-säkehetsvärdering kopplade till de sju stegen i Försvarsmaktens IT-livscykelmodell
- Relevansbedömningar avseende hur väl de olika metodgrupperna motsvarar Försvarsmaktens behov

1.4 Läsanvisningar

Då denna rapport fokuserar på utvecklingen inom området värdering av IT-säkerhet förutsätts grundläggande kunskaper inom områdena IT-säkerhet [1],[2] och värdering av IT-säkerhet [3]. Dessa begrepp förklaras kortfattat i avsnitt 2.1.

Kapitel 2 innehåller övrig bakgrund. I kapitel 3 beskrivs den genomförda litteraturstudien. I kapitel 4 presenteras de identifierade metodgrupperna. I kapitel 5 beskrivs hur de attribut som används för att testa metodgrupper valts ut. I kapitel 6 redovisas resultaten av testningen. I kapitel 7 diskuteras testresultaten. I kapitel 8 diskuteras framtida metoder för värdering av IT-säkerhet.

2 Bakgrund

Detta kapitel beskriver bakgrundskunskap som underlättar läsandet av denna rapport. Avsnitt 2.1 innehåller definitioner för ett antal olika begrepp som används. Avsnittet följs av 2.2 där testproceduren *Test of Security Assessment Relevance* beskrivs. Slutligen beskrivs Försvarsmaktens IT-livscykelmodell i 2.3.

2.1 Begreppsanvändning

Ackreditering utgör ”dels ett sådant godkännande av ett IT-system från säkerhetssynpunkt som avses i 12 § tredje stycket säkerhetsskyddsförordningen (1996:633), dels ett godkännande från säkerhetssynpunkt i övrigt av övriga IT-system” [4].

Auktorisation innebär ”bemyndigande för produktägare att utveckla Försvarsmaktens IT-verksamhet” [4].

Behov beskriver aktiviteter eller resurser som är nödvändiga för att kunna genomföra uppgifter och uppnå mål. Behov kan vara antingen medvetna eller omedvetna, verkliga eller upplevda samt tillfredsställda eller otillfredsställda. Påtalade behov är ofta relaterade till någon form av inneboende krav på åtgärd.

Behov av IT-säkerhetsvärdering (värderingsbehov) beskriver vilka brister på kunskap som behöver åtgärdas. Dessa utgör motiv för att genomföra IT-säkerhetsvärdering.

Informationssystem avser system som samlar in, bearbetar, lagrar och distribuerar information. Innebörden är allmän, men oftast avses datorstödda informationssystem [5]. För att undvika otydligheter avseende vad informationssystem faktiskt inkluderar kan termen IT-system nyttjas.

IT-system avser system med teknik för att hantera och utbyta information med omgivningen [4]. Detta innebär en tydlig koppling till tekniska system.

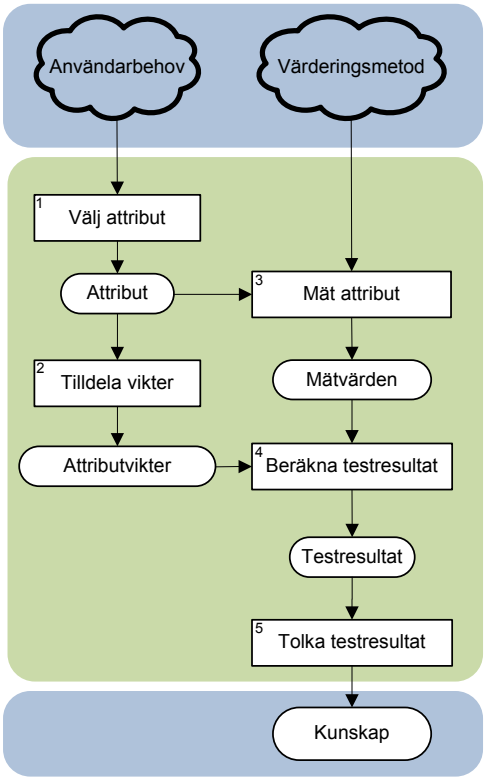
IT-säkerhet avser ett IT-systems förmåga att förhindra obehörig åtkomst och obehörig eller oavsiktlig förändring eller störning vid databehandling samt dator- och telekommunikation [6]. Säkerheten i IT-system beror dock inte endast på systemets tekniska lösningar. Därmed behöver beaktande av IT-säkerhet inkludera även andra aktiviteter och rutiner som påverkar den tekniska utrustningen, såsom hanteringen av lösenord, även om dessa aktiviteter och rutiner i sig inte ingår i IT-systemet.

IT-säkerhetsvärdering (säkerhetsvärdering, värdering) syftar till att öka kunskapen om kvaliteter avseende IT-säkerhet hos system, detta genom att fastställa nivåer för relevanta IT-säkerhetsegenskaper.

2.2 Test of Security Assessment Relevance

Det finns i dagsläget flera olika ansatser, och i vissa fall även färdiga metoder, för hur säkerhetsnivån i IT-system kan avgöras. De olika metoderna skiljer sig ofta avseende vilket underlag som krävs för att en värdering ska kunna genomföras. Detta medför att olika metoder fungerar mer eller mindre bra beroende på de givna förutsättningarna och vilka säkerhetsrelaterade egenskaper hos IT-systemet som ska värderas. Beskrivningen nedan baseras på metodbeskrivningen av TSAR-proceduren [7].

Test of Security Assessment Relevance (TSAR) är en testprocedur som utvärderar relevansen hos olika säkerhetsvärderingsmetoder. TSAR utgår från aktuella behov av säkerhetsvärdering för att avgöra hur väl dessa tillgodoses av de testade värderingsmetoderna. Ett test med TSAR består av de fem aktiviteterna *Välj attribut*, *Tilldela vikter till valda attribut*, *Mät attribut*, *Beräkna testresultat* samt *Tolka testresultat*. Aktiviteterna beskrivs i avsnitt 2.2.1 - 2.2.5. En övergripande bild av arbetsflödet illustreras i Figur 1.



Figur 1: Övergripande bild av arbetsflödet i TSAR.

2.2.1 Välj attribut

Den första aktiviteten i ett test är att ta fram vilka attribut som testet ska grunda sig på. De attribut som tas fram ska motsvara de identifierade användarbehoven. För att underlätta denna aktivitet finns en fördefinierad uppsättning egenskaper ur vilken attribut kan väljas (Bilaga A). Ett attribut är en vald egenskap som motsvarar ett eller flera identifierade behov. De fördefinierade attributen är kategoriserade i en trädstruktur.

Denna aktivitet utförs av testledaren i samarbete med de individer som identifierade behoven av säkerhetsvärdering. På så sätt minskar risken att de valda attributen inte skulle motsvara de identifierade behoven. Resultatet av denna aktivitet är en uppsättning attribut som motsvarar de identifierade behoven av säkerhetsvärdering.

2.2.2 Tilldela vikter

De attribut som valdes i 2.2.1 är oftast av olika betydelse. För att visa om attribut är mer eller mindre betydande tilldelas de vikter. Viktningsförfarandet kan utföras på många olika sätt. TSAR beskriver två olika viktningsförfaranden av olika noggrannhet. Det enklare viktningsförfarandet beskrivs här då detta använts senare i denna rapport.

Då attributen är placerade i en trädstruktur inleds viktningsförfarandet med att först vikta löven i varje kategori för att sedan vikta nivån ovanför. När de kategorier som löven tillhör har viktats, viktas de kategorier som dessa kategorier ingår i. Detta förfarande upprepas tills trädets toppkategori är nådd.

Tilldelning av vikter sker genom att attributen i varje kategori delas upp i de två grupperna *Mindre viktiga (α)* och *Mer viktiga (β)*. De två grupperna ska innehålla lika många attribut, alternativt ett attribut mer i α eller β om det är ett ojämnt antal attribut. Attribut av typ β ska ha en vikt som är dubbelt så stor som de av typ α och summan av alla vikter ska vara 1. Samma viktningsförfarande genomförs sedan för varje kategori ovanför den nyss viktade tills trädets toppkategori nås. Då har samtliga löv erhållit sina lokala vikter, som summerar till 1 för den aktuella kategorin. För att kunna jämföra vikten hos löv som tillhör olika kategorier beräknas globala vikter för alla löv. Den globala vikten för ett löv beräknas genom att multiplicera dess lokala vikt med produkten av vikterna för alla ovanliggande kategorier. Viktningsförfarandet kan genomföras i omvänd ordning, men det rekommenderas att börja på lövnivå då detta ger en tydligare förståelse för vad varje kategori representerar.

Exempel

Antag att kategori A innehåller de tre attributen A1, A2 och A3, medan kategori B innehåller attributen B1, B2, B3 och B4. Tilldelningen av vikter inleds med attributen i kategori A. Antag att A1 är mer viktig (β) än

A2 och A3 (α). Samma sak görs för kategori B där B2 och B4 antas vara mer viktiga (β) än B1 och B3 (α). Slutligen antas även att kategori A är mer viktig (β) än kategori B (α).

Attributen i kategori A som har 1 β och 2 α får då 2+2·1 = 4 viktandelar. För attributen i kategori A resulterar detta i α = 1/4 och β = 2/4. Attributen i kategori B som har 2 β och 2 α får då 2·2+2·1 = 6 viktandelar. För attributen i kategori B resulterar detta i α = 1/6 och β = 2/6.

Slutligen görs vikttilldelningen av kategorierna där A får en vikt på 2/3 medan B får en vikt på 1/3. Resultatet från tilldelningen av vikter kan ses i Tabell 1 där varje attributs lokala och globala vikt återges. Den lokala vikten anger ett attributs betydelse i jämförelse med övriga attribut i samma kategori, medan den globala vikten anger ett attributs betydelse i jämförelse med alla andra attribut. Den globala vikten räknas ut genom att multiplicera ett attributs vikt med den lokala vikten av dess kategori.

Tabell 1: Exempel på tilldelning av vikter.

	Lokal vikt	Global vikt
Kategori A	2/3	
Attribut A1	2/4	6/18
Attribut A2	1/4	3/18
Attribut A3	1/4	3/18
Kategori B	1/3	
Attribut B1	1/6	1/18
Attribut B2	2/6	2/18
Attribut B3	1/6	1/18
Attribut B4	2/6	2/18

2.2.3 Mät attribut

Mätning görs genom att uppfyllnad och täckning avgörs för varje attribut. Uppfyllnadsvärdet återspeglar om den testade säkerhetsvärderingsmetoden uppfyller ett attribut eller inte, medan täckningsvärdet återger om metodbeskrivningen innehåller tillräcklig information för att avgöra uppfyllnadsvärdet för ett attribut.

Om den testade säkerhetsvärderingsmetoden kan anses uppfylla ett attribut sätts uppfyllnadsvärdet till 1 och täckningsvärdet till 1. Om det däremot är helt tydligt att metoden inte uppfyller attributet sätts uppfyllnadsvärdet till 0 och täckningsvärdet till 1.

Om beskrivningen för en metod antyder att ett visst attribut skulle kunna uppfyllas, men inte ingående beskriver hur, sätts uppfyllnadsvärdet till 0 och täckningsvärdet till 0. Detta betyder att metoden inte uppfyller attributet sett till

den befintliga beskrivningen, men att detta skulle kunna ändras om metodbeskrivningen utökas. Täckningsvärdet kan därmed indikera en metods möjligheter att öka uppfyllnadsvärdet.

Exempel

I Tabell 2 återges de möjliga kombinationerna av uppfyllnadsvärde och täckningsvärde. Attributet C1 anses vara uppfyllt. C2 anses inte vara uppfyllt, men beskrivningen av den testade säkerhetsvärderingsmetoden är bristfällig vilket innebär att det med säkerhet inte kunnat avgöras huruvida attributet är uppfyllt. Attribut C3 anses inte vara uppfyllt. Den globala vikten som återges i tabellen påverkar inte uppfyllnad och täckning. Den är inkluderad för det exempel som återfinns i avsnitt 2.2.4.

Tabell 2: Exempel på uppfyllnadsvärden och täckningsvärden.

Attribut	Global vikt	Uppfyllnad	Täckning
C1	2/5	1	1
C2	2/5	0	0
C3	1/5	0	1

2.2.4 Beräkna testresultat

Testresultatet består av ett relevansvärde och en täckningsgrad. Relevansvärdet tas fram genom att summera de globala vikterna för alla uppfyllda attribut. Metodens relevans återspeglas då av ett värde i intervallet 0 till 1.

Täckningsgraden beräknas genom att summera alla täckningsvärden och dividera med antalet attribut. Täckningsgraden visar således hur stor andel av attributen som det varit möjligt att avgöra huruvida de är uppfyllda.

Exempel

Baserat på de värden som återges i Tabell 2 beräknas relevansvärdet och täckningsgraden på följande vis:

Relevansvärde = $(2/5)\cdot 1+(2/5)\cdot 0+(1/5)\cdot 0 = 2/5 = 0{,}4$

Täckningsgrad = $(1+0+1)/3 = 2/3 \approx 67\%$

2.2.5 Tolka testresultat

Att tolka testresultatet kan anses som ganska rättframt då metoden med högst relevansvärde är den metod som är mest relevant och således bäst motsvarar de identifierade behoven. Det finns dock några egenskaper att beakta.

Till att börja med finns det inget gränsvärde som definierar när en metod kan anses vara relevant. En metod som har ett högre relevansvärde anses vara mer

relevant än en metod som har ett lägre relevansvärde, men det finns inget gränsvärde definierat som säger att en metod är relevant om den har ett relevansvärde över exempelvis 0,7.

Vidare bör även täckningsgraden tas i beaktning när resultaten tolkas. En låg täckningsgrad tyder på att metodbeskrivningen varit bristfällig och att metoden potentiellt kan nå ett högre relevansvärde. Detta måste beaktas när metoder väljs för fortsatt granskning eljest riskeras att relevanta metoder förbises.

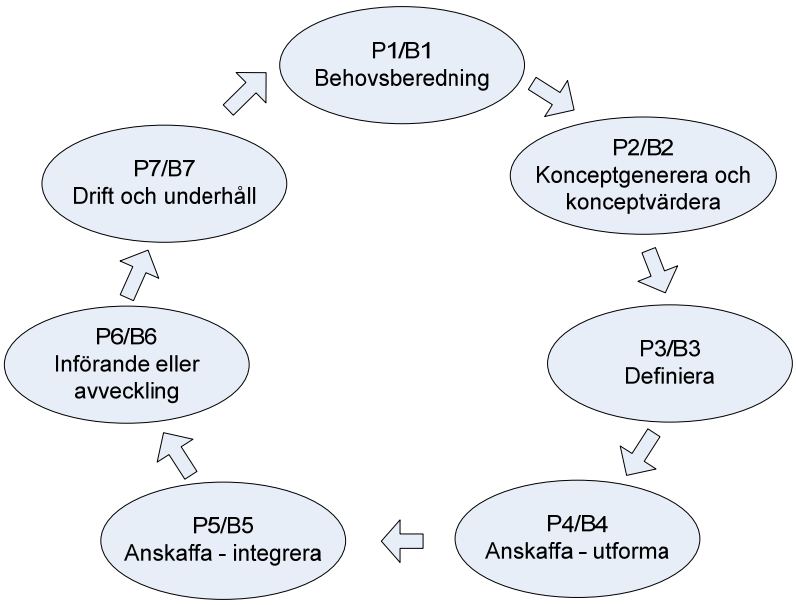
Efter det att testresultaten tolkats och diskuterats ska det beslutas om den mest relevanta metoden ska användas, eller om de mest relevanta metoderna bör studeras djupare innan metodval görs.

2.3 Försvarsmaktens IT-livscykelmodell

Försvarsmaktens IT-livscykelmodell består av två parallella spår. Det ena spåret är auktorisationsprocessen som är ett bemyndigande att utveckla Försvarsmaktens IT-verksamhet. Det andra spåret är ackrediteringsprocessen som utmynnar i ett godkännande av ett IT-system ur säkerhetssynpunkt [4].

IT-livscykelmodellen, återgiven i Figur 2, består av sju steg (P1-P7) och sju beslutspunkter (B1-B7). Auktorisationsprocessen sträcker sig fram till och med beslutspunkt B4, medan ackrediteringsprocessen löper genom hela livscykeln. De olika stegen och beslutspunkterna beskrivs i Direktiv för Försvarsmaktens informationsteknikverksamhet [8].

I denna rapport nyttjas Försvarsmaktens IT-livscykelmodell som grund för att identifiera behov avseende säkerhetsvärdering. Genom att testa säkerhetsvärderingsmetoder mot de behov som finns under ett IT-systems olika faser i IT-livscykelmodellen, fås en tydligare bild över när i utvecklingen olika säkerhetsvärderingsmetoder ger relevanta resultat.



Figur 2: De sju stegen och beslutspunkterna i Försvarsmaktens IT-livscykelmodell.

3 Litteraturstudie

Detta kapitel beskriver den litteraturstudie som genomfördes för att kartlägga aktuella trender inom forskningsområdet *värdering av IT-säkerhet*. Kapitlet inleds med avsnitt 3.1 som beskriver den metod som användes vid genomförandet av litteraturstudien. Avsnittet följs av 3.2 som beskriver genomförandet och resultatet av litteraturstudien.

3.1 Metod

Litteraturstudien inleds med att identifiera en uppsättning söktermer. Dessa söktermer används sedan för att söka efter publikationer i relevanta databaser. De publikationer som fås från dessa sökningar gallras baserat på titel. För de publikationer som anses relevanta baserat på titel görs ytterligare en gallring baserat på publikationens sammanfattning följt av en sista gallring baserat på hela publikationen. De ingående stegen beskrivs mer utförligt i återstoden av detta avsnitt.

3.1.1 Val av söktermer

De termer som senare används vid databassökningar tas fram genom att forskargruppen genomför en brainstorming. Från de termer som framkommer från denna övning väljer projektledaren i samråd med övriga medlemmar ut de termer som ska användas. Termerna ska vara relaterade till området säkerhet i datorbaserade informationssystem.

3.1.2 Databassökning

Efter det att söktermerna fastställts används dessa för att genomföra sökningar i utvalda databaser. För att få med så många träffar som möjligt ska de resultat som tas med endast begränsas om antalet träffar som annars skulle erhållas anses vara oöverskådligt. Hur denna begränsning görs beror på vilka möjligheter som finns under rådande förutsättningar. Då syftet med litteraturstudien är att hitta trender för framtida metoder beaktas endast publikationer från 2008 och framåt. De databaser som avses användas till litteraturstudien är IEEE Xplore, ACM Portal och Springer Link. Resultatet från sökningarna är en lista med titlar på publikationer för respektive databas.

3.1.3 Identifiering av relevanta publikationer

När de tre listorna med titlar har sammanställts, inleds gallringen av de identifierade publikationerna. Den första gallringen sker baserat på

publikationernas titel och genomförs först enskilt av de deltagande forskarna för att sedan ensas vid en gemensam diskussion. Vid denna, första gallring är det titlarna som studeras och de som anses ha potential att innehålla en ansats till en metod för värdering av IT-säkerhet väljs ut. Resultatet från denna diskussion är en ny lista för varje databas med de utvalda titlarna.

Inför nästa gallring kompletteras titlarna som valts ut i föregående steg med tillhörande sammanfattning. Gallring sker på liknande sätt som vid den första gallringen. De deltagande forskarna läser alla sammanfattningar var för sig och väljer ut de som anses kunna innehålla relevanta ansatser till metoder för värdering av IT-säkerhet. Dessa enskilda val ensas sedan vid en gemensam diskussion. Resultatet från detta steg är en lista med utvalda titlar och tillhörande sammanfattning.

I det sista steget av gallringen läses hela de utvalda publikationerna och resultatet är den uppsättning med publikationer som används som grund för det fortsatta arbetet.

3.2 Identifiering av artiklar

En publikation syftar på någon typ av resultat som publiceras som ljud, text, bild eller film. I denna litteraturstudie begränsas studien till publikationer i form av vetenskapliga artiklar. Följande avsnitt beskriver framtagandet av den uppsättning med artiklar som i kapitel 4 används för att skapa grupper med framtida metoder. En sammanfattning av det antal artiklar som identifierades samt hur många artiklar som ansågs relevanta för vidare studier i varje steg finns i Tabell 3.

Tabell 3: Antal artiklar som utgjorde resultatet från de olika stegen i litteraturstudien fördelat på ursprunglig källa

Steg i litteraturstudien	IEEE	ACM	Springer
Databassökning	782	36	58
Gallring baserat på titel	200	8	15
Gallring baserat på sammanfattning	81	1	7
Gallring baserat på hela artiklar	33	0	5

3.2.1 Val av sökuttryck

Framtagandet av de termer som användes för att göra databassökningar inleddes med en brainstormingsession. Under denna session producerades en uppsättning med termer relaterade till området *säkerhet i datorbaserade informationssystem*. Genom att kombinera dessa termer erhölls en lista bestående av 306 sökuttryck. Rimligheten i dessa sökuttryck testades genom att se hur många träffar de gav i

Google Scholar. Resultatet blev totalt 160341 sökträffar i Google Scholar. De enskilda sökuttrycken gav 0 till 9990 sökträffar. Medianvärdet för antalet sökträffar var 113. Detta test visade att sökuttrycken nödvändigtvis måste vara mer avgränsande än generella begrepp såsom ”security”. Från listan med sökuttryck gjordes ett urval baserat på gruppens erfarenheter inom området samt antalet sökträffar för de olika sökuttrycken.

För att förankra valet av termer genomfördes även en enklare validering med syftet att förankra de utvalda termernas relevans. Valideringen genomfördes med hjälp av följande steg.

- Titlarna på alla artiklar ur IEEE Xplore från 2009 lades in i en databas.
- Baserat på databasen togs frekvensen för alla förekommande ordpar fram.
- Relationen mellan de vanligaste ordparen, som innehöll ”security”, i databasen och de vid brainstormingen framtagna termerna studerades för att validera (1) de framtagna söktermernas relevans samt (2) att ingen säkerhetsrelaterad term missats.

Under valideringen framkom att termen ”network security” saknades i den ursprungliga uppsättningen av termer. En ny kombinering av termer gjordes vilket resulterade i 408 sökuttryck. Ett nytt urval av sökuttryck gjordes från den nya listan baserat på gruppens erfarenheter inom området samt antalet sökträffar för respektive sökuttryck i Google Scholar. Sökuttrycken som valdes ut finns listade i Tabell 4. Sökuttrycken är uppdelade i tre termer. Den första termen utgörs av en grundterm innehållande ordet ”security”. Den andra och tredje termen är till för att begränsa träffarna till områden som rimligen ger artiklar inom området värdering av IT-säkerhet. Anledningen till att grundtermen står inom citationstecken är att det var på det sättet sökningen genomfördes för att endast få träffar på den sammansatta grundtermen, inte dess delar.

Tabell 4: De söktermer som användes vid sökning i databaser.

Grundterm	**Avgränsningar**	
”information security”	evaluation	
”information security”	assessment	
”information security”	metric	
”information security”	analysis	method
”information security”	analysis	framework
”information security”	analysis	methodology
”cyber security”	evaluation	
”cyber security”	assessment	
”cyber security”	metric	
”cyber security”	analysis	method
”cyber security”	analysis	framework
”cyber security”	analysis	methodology
“it security”	evaluation	

Grundterm	Avgränsningar	
“it security”	assessment	
“it security”	metric	
“it security”	analysis	method
“it security”	analysis	framework
“it security”	analysis	methodology
“network security”	evaluation	
“network security”	assessment	
“network security”	metric	
“network security”	analysis	method
“network security”	analysis	framework
“network security”	analysis	methodology

3.2.2 Databassökning

Vid sökning i samtliga tre databaser *IEEE Xplore*, *ACM Portal* och *Springer Link* användes sökuttrycken i Tabell 4 för att hitta alla matchande artiklar. Sökningen genomfördes för artiklar publicerade från 2008 fram till det datum sökningen genomfördes (2010-05-10).

Vid sökning i IEEE Xplore användes ett sammansatt uttryck med vilka alla de identifierade sökuttrycken kombinerades till ett logiskt uttryck. Sökningen gav ett antal träffar som ansågs vara för omfattande, varpå en begränsning gjordes genom att bara ta med träffar från ämnesområden som kunde anses som relevanta. De områden som valdes ut var: ”Computing & Processing (Hardware/Software)”, ”Communication, Networking & Brodcasting”, ”Signal Processing & Analysis” samt ”General Topics for Engineering (Math, Science & Engineering)”. Det är värt att notera att i IEEE Xplore kan en artikel tillhöra flera kategorier, vilket innebär att alla för litteraturstudien relevanta artiklar bör tillhöra någon av de valda kategorierna. Sökningen i IEEE Xplore gav 782 träffar.

Även vid sökningen i ACM Portal kombinerades alla sökuttryck till ett logiskt uttryck. Då det är nödvändigt att specificera vilka fält som ska ingå i en sökning i ACM gjordes sökningen i fälten *Title* och *Abstract*. Antalet träffar i denna databas var betydligt färre, jämfört med IEEE Xplore, och det ansågs därför inte nödvändigt att göra några ytterligare begränsningar. Sökningen i ACM Portal gav 36 träffar.

Då Springer Link inte tillåter logiska uttryck längre än 10 termer fick sökningen delas upp i multipla sökningar. De fält som användes var *Title* och *Summary*. Liksom sökningen i ACM var den mängd träffar som erhölls hanterbar, vilket betydde att inga ytterligare begränsningar var nödvändiga. Sökningen i Springer Link gav 58 träffar.

3.2.3 Identifiering av relevanta artiklar

Nästa steg i litteraturstudien var att göra en gallring bland de 876 träffarna och avgöra vilka artiklar som skulle ligga till grund för testet av framtida metoder.

Första steget i denna gallring genomfördes baserat på artiklarnas titel. Aktiviteten genomfördes av tre personer och inleddes med att varje person enskilt markerade vilka titlar som denne ansåg intressanta vilket följdes av en diskussion där dessa markeringar ensades. Gallringen baserat på titel resulterade i att 223 artiklar valdes ut för vidare studier.

Nästa steg i gallringen skedde på samma sätt som det föregående, med undantag för att gallringen i detta steg skedde baserat på artiklarnas sammanfattning istället för på titel. Resultatet var att 89 artiklar valdes ut.

Sista steget i gallringen skedde baserat på en genomläsning av hela artiklar. Då det inte ansågs möjligt att genomföra genomläsningen av alla 89 artiklar på samma sätt som i de tidigare stegen delades uppgiften att läsa varje artikel upp mellan de tre personerna. Dock rådfrågade personerna varandra om artiklar låg på gränsen till vad som ansågs vara relevant för att se en trend för framtida metoder. Från genomläsningen av de kvarvarande artiklarna valdes 38 artiklar ut. Dessa artiklar presenteras i Tabell 5.

Tabell 5: De artiklar som valdes ut under litteraturstudien

ID	Titel	Referens
A1	An Information System Security Evaluation Model Based on AHP and GRAP	[9]
A2	Research on System Integrative Evaluation Method Based on Survivability	[10]
A3	A framework for security quantification of networked machines	[11]
A4	A Network Security Evaluation Model based on Common Criteria	[12]
A5	Cyber Security Risks Assessment with Bayesian Defense Graphs and Architectural Models	[13]
A6	Research on Network Security Risk Assessment Based on Group Consistency	[14]
A7	GARNET: A Graphical Attack Graph and Reachability Network Evaluation Tool	[15]
A8	A Simple, Smart and Extensible Framework for Network Security Measurement	[16]
A9	Assessment of Information Security Levels in Power Communication Systems Using Evidential Reasoning	[17]
A10	Security Situation Assessment Based on the DS Theory	[18]
A11	Information Security Risk Assessment Based on AHP/DST	[19]
A12	The analysis of uncertainty of network security risk assessment using Dempster-Shafer theory	[20]
A13	Study on assessment method for computer network security based on	[21]

ID	Titel	Referens
	rough set	
A14	Method of risk evaluation of information security based on neural networks	[22]
A15	A novel approach to network vulnerabilities quantitative evaluation based on Mamdani-Style fuzzy logic	[23]
A16	A Method of Security Evaluation based on Fuzzy Mathematics	[24]
A17	A Method of Information Security Risk Assessment Using Fuzzy Number Operations	[25]
A18	Information Security Risk Assessment Based on Analytic Hierarchy Process and Fuzzy Comprehensive	[26]
A19	Network security situation quantitative evaluation based on the classification of attacks in attack-defense confrontation environment	[27]
A20	Network Security Situation Generation and Evaluation Based on Heterogeneous Sensor Fusion	[28]
A21	Network Security Situation Assessment Based on Data Fusion	[29]
A22	Network threat assessment based on attribute recognition	[30]
A23	Research on Network Risk Situation Assessment Based on Threat Analysis	[31]
A24	A Logic-based Approach to Network Security Risk Assessment	[32]
A25	A Game Theoretical Attack-Defense Model Oriented to Network Security Risk Assessment	[33]
A26	iMeasure Security (iMS): A Novel Framework for Security Quantification	[34]
A27	VEA-bility Security Metric: A Network Security Analysis Tool	[35]
A28	Applying Attack Graphs to Network Security Metric	[36]
A29	A Mobile Ambients-Based Approach for Network Attack Modelling and Simulation	[37]
A30	A Flexible Approach to Measuring Network Security Using Attack Graphs	[38]
A31	A Method for Global Attack Graph Generation	[39]
A32	An Attack Graph-Based Probabilistic Security Metric	[40]
A33	An Approach for Security Assessment of Network Configurations Using Attack Graph	[41]
A34	Estimating a System's Mean Time-to-Compromise	[42]
A35	Access Graph Based Risk Assessment Model for Network Information System Using Access Graph to Analyze Network Vulnerabilities	[43],[44]
A36	An OVAL-based active vulnerability assessment system for enterprise computer networks Study of generating attack graph based on privilege escalation for computer networks	[45],[46]
A37	Information system security compliance to FISMA standard: A quantitative measure Information system security compliance to FISMA standard: A quantitative measure	[47],[48]
A38	A Novel Quantitative Approach For Measuring Network Security	[49]

4 Identifierade metodgrupper

Resultatet från litteraturstudien utgörs av 38 vetenskapliga artiklar. För att erhålla en grund för resonemang om framtida utvecklingen inom området värdering av IT-säkerhet grupperades dessa artiklar. I detta kapitel presenteras arbetet med och resultatet av denna gruppering som utgörs av fyra metodgrupper.

4.1 Gruppering av artiklar

För att gruppera artiklarna nyttjades en metod enligt följande steg.

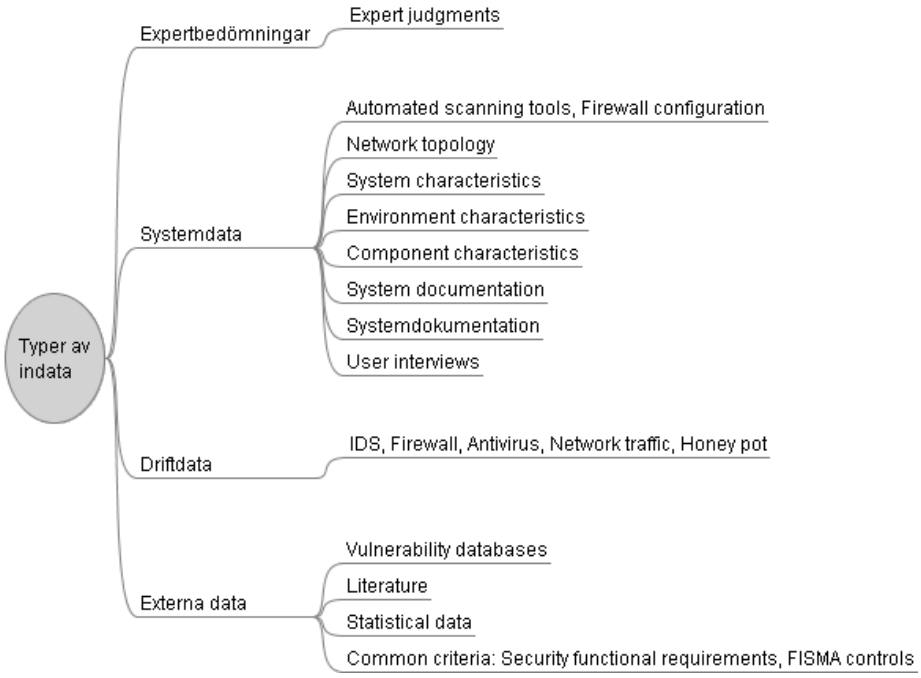
- Varje artikel klassificeras genom att specificera vilka typer av indata som behövs samt vilka karaktäriserande metoder som nyttjas för att genomföra värderingar i enlighet med den aktuella artikeln.
- En tabell med olika typer av indata i horisontell ledd och olika karaktäriserande metoder i vertikal ledd skapas. I tabellen förs de klassificerade artiklarna in.
- De olika typer av indata och karaktäriserande metoder som har identifierats grupperas till mer generella kategorier.
- En kategoritabell skapas med kategorier avseende indata och karaktäriserande metoder i horisontell respektive vertikal ledd. Även i denna tabell förs de olika artiklarna in.
- Utifrån kategoritabellen identifieras metodgrupper.

Resultatet av den initiala klassificeringen av artiklarna återfinns i Tabell 6. För att undvika svårtolkade översättningar har de engelska namnen på de olika typerna av indata och de karaktäriserande metoderna behållits i Tabell 6. I tabellen nyttjas intervall för att identifiera flera artiklar med angränsande numrering, t ex avser A7-8 de två artiklarna A7 och A8.

Tabell 6: Klassificering av artiklar baserat på typ av indata, som anges i kolumn 1, och karaktäriserande metod för att ta fram säkerhetsvärden, som anges i rad 1.

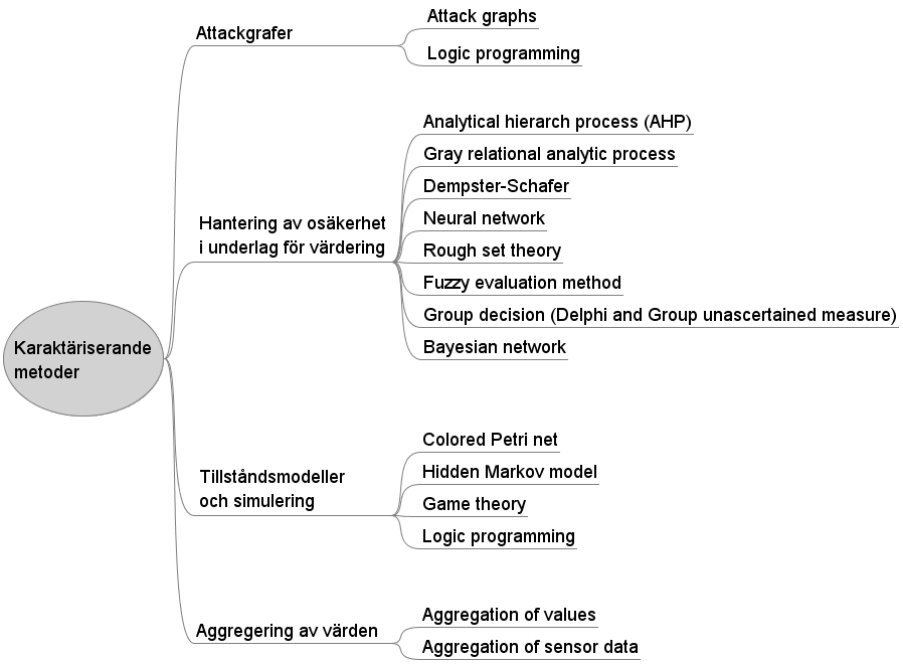
	Analytical hierarchy process	Gray process	Dempster-Shafer	Neural network	Rough set theory	Fuzzy evaluation method	Group decision	Bayesian network	Aggregation of values	Aggregation of sensor data	Pathfinder network	Colored Petri net	Hidden Markov model	Game theory	Logic programming	Attack graph
Expert judgment	A1 A11-14 A16-18	A1	A9 A11 A12	A14	A13	A14-18	A12 A17	A5	A9, A32-33 A38	A19 A22	A37		A3	A25	A24	A26 A28 A30 A32-35
Autom. scan. tool, Firewall configuration									A27	A19 A23					A36	A7-8 A27 A29 A31 A36
Vulnerability database									A27 A32-33 A38						A36	A7 A26-27 A29 A31-33 A36
Network topology									A27 A38							A26-27 A29
IDS, Firewall, Antivirus, Network traffic, Honeypot			A10 A21	A20					A10	A19-20 A22 A23		A21	A3	A25		
Literature								A5								
Statistical data								A5								
User interview			A9						A9							
Common criteria FISMA control	A1 A16	A1				A16					A37					
Sys. charac.									A2							
Environ. charac.									A2							
Component characteristic									A2 A4				A3			
Sys. Document			A9						A9							

Grupperingen av indatatyper resulterade i de fyra kategorierna *expertbedömning*, *systemdata*, *driftdata* och *externa data* (eng. *expert judgment*, *system data*, *operational data* and *external data*). Vilka typer av indata som ingår i respektive kategori framgår av Figur 3. Liksom tidigare i avsnittet har de engelska benämningarna på de olika typerna av indata behållits i Figur 3.



Figur 3: Olika typer av indata, med de enskilda typerna till höger och kategorierna de tillhör i mitten

Grupperingen av karaktäriserande metoder resulterade i de fyra kategorierna *attackgrafer*, *hantering av osäkerhet i underlag för värdering*, *tillståndsmodeller och simulering* samt *aggregering av värden*. Ett specialfall utgörs av den karaktäriserande metoden *logikprogrammering* som tillhör två olika kategorier beroende på vilken typ av indata som avses. Om indata är av typen *expertbedömingar* infaller den i kategorin *tillståndsmodeller och simulering*. I övriga fall klassificeras den som *attackgrafer*. Vilka karaktäriserande metoder som ingår i respektive kategori framgår av Figur 4. På samma sätt som tidigare har de engelska benämningarna på de karaktäriserande metoderna behållits i Figur 4.



Figur 4: Olika karaktäriserande metoder, med de enskilda metoder till höger och kategorierna de tillhör i mitten

Den resulterande kategoritabellen utgörs av Tabell 7. Baserat på denna tabell kan metodgrupper väljas som en kombination av kategorier för typer av indata respektive karaktäriserande metoder. Då flertalet artiklar beskriver nyttjande av flera olika typer av indata valdes metodgrupperna att utgöras av de karaktäriserande metoderna, oberoende av typ av indata. De resulterande metodgrupperna blev således *attackgrafer*, *hantering av osäkerhet i underlag för värdering*, *tillståndsmodeller och simulering* samt *aggregering av värden*. Dessa metodgrupper beskrivs i följande avsnitt.

Tabell 7: Tabell med artiklar klassificerade enligt kategorier för typ av indata och karaktäriserande metoder.

	Attackgrafer	Hantering av osäkerhet i underlag för värdering	Tillståndsmodeller och simulering	Aggregering av värden
Expert-bedömningar	A26, A28, A30, A32, A33, A34, A35	A1, A5, A6, A9, A11, A12, A13, A14, A15, A16, A17, A18	A3, A25, A21	A9, A19, A22, A32, A33, A37, A38
Systemdata	A7, A8, A26, A27, A29, A31, A36	A9	A3	A2, A4, A9, A19, A23, A27, A38
Driftdata		A10, A20, A21	A3, A21, A25	A10, A19, A20, A22, A23
Externa data	A7, A26, A27, A29, A31, A32, A33, A36	A1, A5, A16		A27, A32, A33, A37, A38

4.2 Beskrivning av metodgrupper

I detta avsnitt beskrivs de fyra metodgrupper som identifierades i föregående avsnitt. Dessa beskrivningar utgör underlag för testet av dessa metodgruppers relevans för Försvarsmakten. Resultatet beskrivs i kapitel 6.

4.2.1 Attackgrafer

Metodgruppen attackgrafer utgår från systemstruktur och kända sårbarheter för att skapa en graf över hur en motståndare kan ta sig mellan de olika noderna i systemet. Denna graf används sedan som underlag för att genomföra olika analyser av IT-säkerheten.

Metoder baserade på attackgrafer inkluderar följande steg

- Samla data om systemet
- Koppla data avseende systemet mot data från sårbarhetsdatabaser
- Skapa attackgraf utifrån identifierade sårbarheter och nätverkstopologi
- Analysera graf
- Tolka analysresultat

Att samla in data om systemet är en stor svårighet och den del som finns minst beskrivet i de studerade artiklarna. Beskrivna tillvägagångssätt inkluderar användande av automatiska verktyg för identifiering av noder, nätverkstopologi samt de sårbarheter verktyget kan känna igen (A7). Vid insamling av data om systemet nyttjas även experter för att validera och komplettera automatiskt insamlad data.

Få av de studerade artiklarna fokuserar på skapandet av attackgrafer (A7), dvs. att gå från data om systemet till en graf som beskriver alla möjliga attacker baserade på kända sårbarheter. Oftast tas grafen fram av experter som utgår från insamlad data. Funktionaliteten hos verktyg för automatisk generering av attackgrafer är i dagsläget ganska begränsad.

Angående analys av den framtagna grafen finns ett antal olika angreppssätt. Beskrivna angreppssätt innefattar exempelvis:

- Beräkna antalet vägar in till skyddsvärda objekt (A29).
- Beräkna sannolikheten att en attack mot skyddsvärda objekt lyckas (A26, A28, A30, A31, A32, A33, A35).
- Beräkna genomsnittlig tidsåtgång för genomförande av attack (A34).
- Beräkna mått på hur välkonfigurerat ett nätverk är (A27).

Det finns i dagsläget få verktyg (A7, A8) för att producera utdata från attackgrafer, det vill säga att generera underlag för värdering av IT-säkerhet baserat på framtagna attackgrafer.

Noterbart är att det finns ett verktyg (A7) som ger stöd vid både import av insamlade data, skapande av attackgraf samt genomförande av analys. För övrigt innehåller de studerade artiklarna mestadels förslag på hur verktyg eller algoritmer skulle kunna realiseras (A30, A31, A32, A35, A36) snarare än färdiga implementationer.

Tolkning av vad den framtagna attackgrafen säger om systemets IT-säkerhet utgår från olika egenskaper hos grafen, men baseras helt på expertutlåtanden.

4.2.2 Hantering av osäkerhet i underlag för värdering

Metoderna i denna grupp är inriktade på att hantera osäkerheter i indata. Huvudsakligen fokuserar metoderna på hur icke fullständiga indata ska kunna användas. En annan inriktning är hur slutsatser ska kunna dras från subjektiva bedömningar.

Vanligaste källan för indata är experter. Exempel på övriga indata är intervjuer med användare, statistiska data om styrkan hos säkerhetsfunktionerna samt olika typer av systemloggar. Specificeringen av hur data ska samlas in är generellt sett vag. Övervägande delen av indata handlar om expertutlåtanden. På grund av

områdets dynamik kommer inte verktyg för automatisk insamling av expertkunskap, t ex via databas, att finnas i närtid, om ens någonsin. Även för insamling av driftdata krävs i nuläget mänsklig inblandning.

En konsekvens av fokuseringen på hantering av osäkerheter i underlag för värdering, snarare än själva värderingen av säkerhet i system, är att artiklarna inte är tydliga avseende vilka värderingsresultat som kan produceras. Möjliga resultat avseende värdering av IT-säkerhet inkluderar dock sannolikheter för lyckade attacker (A5), sannolikhetsvärden kopplade till olika sårbarheter (A15), sannolikhetsvärden för potentiella säkerhetsproblem av generell karaktär, såsom skadlig kod (A14), samt säkerhetsvärden som ej relaterar till sannolikheter (A1, A11, A12, A13, A16, A17, A18).

I denna metodgrupp ingår metoder baserade på en rad olika typer av ansatser såsom Analytical Hierarchy Process (A1, A11, A12, A13, A14, A16, A17, A18), Gray Relational Analytic Process (A1), Dempster-Shafer theory (A9, A10, A11, A12, A21), Rough set theory (A13), Fuzzy evaluation method (A14, A15, A16, A17, A18), Bayesian networks (A5), Neural networks (A14, A20) och Delphi process (A12, A17).

4.2.3 Tillståndsmodeller och simulering

Metoderna i denna grupp använder formella modelleringstekniker för att beskriva olika tillstånd som de system som ska värderas kan befinna sig i. Använda modelleringstekniker är Hidden Markov Models (A3), Colored Petri Nets (A21) och Game Theory (A25). Tillstånden nyttjas för att beskriva säkerhetsläget i ett system (A21).

Indata utgörs av egenskaper för noder (A3), data som samlas in under drift (A3, A25) samt expertkunskap (A3, A21, A25). Expertkunskap krävs för att modellera nodernas påverkan på systemets säkerhetsnivå (A3), sannolikheter för övergångar mellan tillstånd (A21) samt förhållande mellan systemstatus och möjliga attacker (A25).

Resultaten utgörs av riskfaktorer för noder och sammanfattande värden för hela system (A3), sannolikheter för olika säkerhetstillstånd i systemet (A21) samt sannolikheter för lyckade attacker (A25).

4.2.4 Aggregering av värden

Metoderna i denna grupp fokuserar på problemet med hur olika säkerhetsrelaterade värden kan slås samman till en uppsättning sammansatta säkerhetsvärden. De metoder som används för att genomföra själva aggregeringen av värden inkluderar aritmetiska operationer (A2, A4, A9, A10, A19, A20, A22, A23, A37, A38) samt tilldelning och sammanslagning av sannolikheter (A10)

Då metoderna fokuserar på att aggregera många olika typer av data utgörs indata till metoderna exempelvis av IDS- och nätverksdata (A10, A19, A20, A22, A23), expertutlåtanden (A9, A37, A38), egenskaper hos entiteter (A2, A4, A9, A38) och säkerhetsstandarder (A37). Generellt sett består denna metodgrupp av två huvudspår som utgår antingen från sensordata eller från expertbedömningar alternativt dokumentation av systemens egenskaper.

Resultatet från metoderna är aggregerade värden som beskriver säkerhetsegenskaper hos system. Dessa säkerhetsvärden inkluderar säkerhetssituation i nätverk (eng. network security situation) (A4, A10, A19, A20, A23), tillförlitlighetsvärden (A9), avstånd till tillstånd utan sårbarheter (A37), överlevnadsförmåga hos systemet (A2) samt hotvärdering (A22, A38).

5 Val av attribut och vikter

TSAR-proceduren, som beskrivs övergripande i avsnitt 2.2, är en testprocedur som används för att avgöra hur väl en säkerhetsvärderingsmetod motsvarar befintliga säkerhetsvärderingsbehov. En förutsättning för att kunna genomföra ett test är således att befintliga behov har identifierats. Detta kapitel beskriver hur attribut, som motsvarar Försvarsmaktens behov av säkerhetsvärdering, valdes ut och viktades.

5.1 Val av attribut

För att identifiera vilka behov som finns hos Försvarsmakten, avseende säkerhetsvärdering av IT-system, genomfördes en enkätundersökning. Frågor formulerades, utifrån den uppsättning attribut som återfinns i metodbeskrivningen av TSAR-proceduren [7], för att identifiera vilka behov som finns i de olika stegen i Försvarsmaktens IT-livscykelmodell. Enkäten innehöll påståenden som motsvarar de olika attributen i den fördefinierade uppsättningen. För varje påstående kunde den svarande välja för vilket eller vilka steg i IT-livscykelmodellen påståendet ansågs vara relevant. Enkätundersökningen genomfördes webbaserat med hjälp av enkätsystemet Lime Survey[1]. Genomförandet av enkätundersökningen motsvarar den första aktiviteten i TSAR-proceduren, *Välj attribut*, som beskrivs i avsnitt 2.2.1.

Enkäten skickades ut till ett antal personer hos Försvarsmakten, Försvarets Materielverk och Totalförsvarets forskningsinstitut som jobbar med frågor relaterade till IT-livscykelmodellen. Enkätundersökningen resulterade i 24 enkätsvar, varav 12 var kompletta. Endast de 12 kompletta enkätsvaren beaktades. Varje enkätsvar resulterade i en svarsvektor.

För att identifiera anomalier bland de 12 kompletta enkätsvaren mättes varje svarsvektors Hammingavstånd [50]. Ett Hammingavstånd visar antalet positioner som har olika symboler vid jämförelse av två strängar. I detta fall mättes Hammingavståndet för varje svarsvektor till övriga svarsvektorer. De två svarsvektorer som hade störst Hammingavstånd till övriga svarsvektorer sorterades bort. En granskning av de två bortsorterade svarsvektorerna visade att dessa svarsvektorer i störst utsträckning uttryckt attribut som relevanta när de övriga ej ansett attributen som relevanta. De resterande tio svarsvektorerna utgjorde slutligen det underlag som användes för att identifiera de relevanta attributen för varje steg i Försvarsmaktens IT-livscykelmodell.

De attribut som återfanns i två eller flera svarsvektorer ansåg som relevanta och valdes därmed ut. Resultatet av enkätundersökningen, det vill säga de attribut

[1] http://www.limesurvey.org

som ansågs relevanta för de olika stegen i Försvarsmaktens IT-livscykelmodell, återges i Tabell 8. De fördefinierade attributen återges i Bilaga A.

Tabell 8: Utvalda attribut (1.1 till 5.3.2) för de olika stegen (B1 till B7) i Försvarsmaktens IT-livscykelmodell.

	B1	B2	B3	B4	B5	B6	B7
1.1	1	1	1	1	1	1	1
1.2	1	1	1	1	1	1	1
1.3	1	1	1	1	1	1	1
1.4		1		1	1	1	1
1.5	1	1	1	1	1	1	1
2.1.1		1	1		1	1	1
2.1.2							
2.1.3.1		1	1	1	1	1	1
2.1.3.2			1	1	1	1	1
2.1.3.3				1	1	1	1
2.1.3.4	1	1	1	1	1	1	1
2.1.3.5			1	1	1	1	1
2.2.1					1	1	1
2.2.2				1	1	1	1
2.2.3	1	1	1	1			
3.1		1	1	1	1	1	1
3.2	1	1	1	1	1	1	1
4.1		1	1	1	1	1	1
4.2	1	1	1	1	1	1	1
4.3	1	1	1	1	1	1	1
4.4		1	1	1	1		
4.5	1	1	1	1	1	1	1
5.1.1	1	1	1	1	1	1	1
5.1.2	1	1	1	1	1	1	1
5.1.3	1			1			
5.2.1	1	1	1	1	1	1	1
5.2.2	1	1	1	1	1	1	1
5.2.3					1		
5.3.1	1	1	1	1	1	1	1
5.3.2	1				1		

5.2 Val av attributvikter

Den andra aktiviteten i TSAR-proceduren, *Tilldela vikter*, beskriven i avsnitt 2.2.2, innebär att de attribut som valdes i 5.1 ska tilldelas vikter. Detta då de olika attributen oftast inte är av samma betydelse, vilket även bör återspeglas i testprocedurens resulterande relevansvärde.

Vikterna för de utvalda attributen togs fram av två personer genom att de var för sig tilldelade attributen vikter utifrån den förenklade viktningsmetoden. Dessa två förslag på viktfördelning jämfördes sedan och för de attribut där åsikterna skiljde sig åt fördes en diskussion tills konsensus nåddes. Detta förfarande genomfördes för varje steg i IT-livscykelmodellen, vilket resulterade i en viktvektor för varje steg i IT-livscykelmodellen. Det sammanställda resultatet återges i Tabell 9.

Tabell 9: Framtagna attributvikter för de olika stegen i Försvarsmaktens IT-livscykelmodell.

	B1	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**
1.1	0,083	0,063	0,095	0,041	0,063	0,036	0,036
1.2	0,083	0,063	0,048	0,020	0,031	0,071	0,036
1.3	0,042	0,031	0,095	0,041	0,063	0,071	0,036
1.4		0,031		0,020	0,063	0,071	0,018
1.5	0,042	0,063	0,048	0,020	0,031	0,036	0,018
2.1.1		0,014	0,016		0,056	0,063	0,063
2.1.2							
2.1.3.1		0,009	0,005	0,048	0,028	0,009	0,009
2.1.3.2			0,005	0,048	0,028	0,009	0,009
2.1.3.3				0,048	0,014	0,005	0,005
2.1.3.4	0,042	0,019	0,011	0,024	0,014	0,005	0,005
2.1.3.5			0,011	0,024	0,028	0,005	0,005
2.2.1					0,028	0,032	0,032
2.2.2				0,063	0,056	0,016	0,016
2.2.3	0,083	0,083	0,095	0,032			
3.1		0,083	0,048	0,095	0,083	0,095	0,095
3.2	0,125	0,042	0,095	0,048	0,042	0,048	0,048
4.1		0,036	0,041	0,071	0,036	0,024	0,048
4.2	0,063	0,036	0,041	0,036	0,036	0,024	0,048
4.3	0,125	0,071	0,082	0,071	0,071	0,048	0,095
4.4		0,071	0,082	0,071	0,071	0,048	
4.5	0,063	0,036	0,041	0,036	0,036		0,095
5.1.1	0,025	0,033	0,012	0,029	0,042	0,095	0,048
5.1.2	0,050	0,067	0,024	0,014	0,021	0,048	0,024
5.1.3	0,025			0,014			
5.2.1	0,017	0,017	0,012	0,038	0,016	0,048	0,048
5.2.2	0,033	0,033	0,024	0,019	0,008	0,024	0,024
5.2.3					0,008		
5.3.1	0,067	0,100	0,071	0,029	0,021	0,071	0,143
5.3.2	0,033				0,010		

6 Testning av metodgrupper

För att avgöra de fyra identifierade metodgruppernas relevans för Försvarsmakten, undersöktes uppfyllnad av de i föregående kapitel utvalda attributen. Urvalet av attribut är kopplat till de sju stegen i Försvarsmaktens IT-livscykelmodell, vilket betyder att varje metodgrupp får sju uppsättningar med värden som motsvarar stegen i livscykelmodellen. Testningen genomfördes med hjälp av den av FOI framtagna demonstratorn som realiserar TSAR [51].

För var och en av metodgrupperna återfinns testresultaten i en tabell. Varje tabell innehåller sju rader, där dessa syftar på framtagandet av det underlag som behövs inför de sju beslutspunkterna B1 till B7. Varje rad innehåller värden för relevans, täckningsgrad samt R-max. Relevansen återges som ett värde från 0 till 1, där 1 indikerar hög relevans medan 0 indikerar ingen relevans. Täckningsgraden utgörs av andelen attribut som kunde mätas. R-max syftar på det maximala relevansvärde som skulle kunna uppnås om täckningsgraden vore 100%. Notera att R-max är ett viktat värde medan täckningsgraden är oviktad. Det går därför inte att räkna fram R-max med hjälp av de data som återfinns i tabellerna i avsnitt 6.1 till 6.4.

6.1 Attackgrafer

I Tabell 10 återges testresultaten för metodgruppen attackgrafer. Metodgruppen förutsätter till stor del att det finns ett system, eller åtminstone en väl utvecklad design av ett system, att samla data ifrån. Detta återspeglas i testresultaten genom att relevansvärdena för stegen B1 till B3 är lägre än för de senare stegen. Det är därmed till viss del osäkert hur metodgruppen kan utvecklas för att bättre hantera värderingar under de tidigare stegen (B1 till B3) i IT-livscykelmodellen. Denna osäkerhet återspeglas av att täckningsgraden generellt sett ökar ju närmare drift systemet befinner sig.

Tabell 10: Testresultat för metodgruppen attackgrafer.

	Relevans	Täckningsgrad	R-Max
B1	0,429	70,6%	0,783
B2	0,529	71,4%	0,810
B3	0,462	72,7%	0,881
B4	0,654	76,0%	0,905
B5	0,624	81,5%	0,849
B6	0,650	79,2%	0,798
B7	0,769	83,3%	0,905

6.2 Hantering av osäkerhet i underlag för värdering

I Tabell 11 återges testresultaten för metodgruppen hantering av osäkerhet i underlag för värdering. Metodgruppen baserar större delen av värderingarna på expertutlåtanden, vilket är applicerbart i alla IT-livscykelsteg. Dock fokuserar de flesta artiklar på de senare stegen i IT-livscykeln, vilket ger avtryck i R-max för B6 och B7. De generellt höga värdena på R-max beror på att metodgruppen innehåller många olika, men relaterade, ansatser.

Tabell 11: Testresultat för metodgruppen hantering av osäkerhet i underlag för värdering.

	Relevans	Täckningsgrad	R-Max
B1	0,521	64,7%	0,942
B2	0,559	66,7%	0,929
B3	0,474	68,2%	0,913
B4	0,602	72,0%	0,867
B5	0,624	74,1%	0,883
B6	0,646	70,8%	0,991
B7	0,616	70,8%	0,991

6.3 Tillståndsmodeller och simulering

I Tabell 12 återges testresultaten för metodgruppen tillståndsmodeller och simulering. Gruppen har generellt en ganska låg relevans i de olika stegen. Detta kan bero på att de artiklar som ingår är beskrivna på en övergripande nivå. Fokus ligger snarare på modelleringsmetoden än genomförandet av IT-säkerhetsvärdering. Denna fokusering avspeglas i de relativt låga relevansvärdena samt de höga värdena på R-max. Metodgruppen förutsätter att det finns ett system, eller åtminstone en väl utvecklad design av ett system, att samla data ifrån. Detta återspeglas tydligast i testresultaten avseende R-max som ökar ju närmare drift systemet befinner sig.

Tabell 12: Testresultat för metodgruppen tillståndsmodeller och simulering.

	Relevans	Täckningsgrad	R-Max
B1	0,521	76,5%	0,775
B2	0,402	61,9%	0,836
B3	0,406	63,6%	0,847
B4	0,546	72,0%	0,910
B5	0,540	70,4%	0,937
B6	0,539	66,7%	0,960
B7	0,509	70,8%	0,978

6.4 Aggregering av värden

I Tabell 13 återges testresultaten för metodgruppen aggregering av värden. Inom denna metodgrupp ligger fokus på värdering av IT-säkerhet snarare än vilka modellerings- och utvärderingsmetoder som används. Detta fokus gör det svårt att avgöra huruvida de olika attributen är uppfyllda eller ej, vilket avspeglas i relativt låg relevans och täckningsgrad. Samtidigt resulterar detta fokus i höga värden på R-max.

Tabell 13: Testresultat för metodgruppen aggregering av värden.

	Relevans	Täckningsgrad	R-Max
B1	0,396	52,9%	0,942
B2	0,374	47,6%	1,000
B3	0,321	45,5%	1,000
B4	0,451	52,0%	0,986
B5	0,526	59,3%	0,982
B6	0,546	54,2%	1,000
B7	0,516	58,3%	1,000

7 Diskussion

I denna rapport presenteras en litteraturstudie för att identifiera ansatser till framtida metoder för värdering av IT-säkerhet. Baserat på litteraturstudien har mer generella metodgrupper identifierats. Dessa metodgrupper definieras av vetenskapliga artiklar, vilka presenterar ansatser till framtida metoder för värdering av IT-säkerhet. För att avgöra hur relevanta dessa metodgrupper är för Försvarsmakten har de testats med hjälp av testproceduren TSAR.

Syftet med litteraturstudien var att hitta indicier för vad som kan utgöra grunden för framtida metoder, inte att utvärdera redan föreslagna metoder. Därmed kan underlaget te sig fragmentariskt och mer genomarbetade förslag till metoder saknas. Det är nämligen huvudsakligen konferensbidrag som återfinns och dessa är begränsade i sin omfattning. De flesta artiklar är inriktade på att lösa delproblem snarare än att ta fram en helhetslösning. Detta var anledningen till att det var metodgrupper bestående av ett antal artiklar, snarare än enskilda artiklar, som testades.

7.1 Rekommendationer

Utifrån de tester som genomförts med TSAR-proceduren kan rekommendationer göras angående vilka metodgrupper som anses ha störst relevans i framtiden för Försvarsmakten under de olika stegen i IT-livscykelmodellen. Genom att sammanställa testresultaten för de fyra metodgrupperna kan det identifieras vilka metoder som erhållit högst relevans för varje steg. Resultatet av denna sammanställning återges i Tabell 14.

Det kan vara värt att notera att metodgruppen innehållande artiklar som berör hantering av osäkerhet i underlag för värdering anses ha högst relevans i framförallt de tidiga stegen av IT-livscykelmodellen, där något system fortfarande inte har realiserats. Det är under dessa steg som metodgruppens förmågor kommer bäst till pass då en IT-säkerhetsvärdering kan genomföras även om underlaget inte är komplett eller om underlaget innehåller osäkerheter.

Under IT-livscykelmodellens senare steg har metodgruppen attackgrafer uppnått högst relevans. Artiklarna i denna metodgrupp fokuserar på realiserade system. Underlag samlas in från olika typer av loggar samt genom att skanna av systemet för att identifiera nätverkstopologi och de ingående nodernas egenskaper. Underlaget analyseras sedan för att med hjälp av olika sårbarhetsdatabaser ta fram en bild över vilka kända sårbarheter som existerar i systemets noder, samt hur olika sårbarheter kan kombineras för att en angripare ska kunna nå skyddsvärda noder. På så sätt ger attackgrafer en mer aktuell och detaljerad bild över vilka specifika sårbarheter som finns i de system som värderas. Det krävs dock fortfarande omfattande arbete för att skapa attackgrafer utifrån insamlade

indata. Därmed finns det utrymmer för såväl forskning som utveckling innan metodgruppen kan anses vara helt mogen. Trots det bedöms metodgruppen attackgrafer vara en intressant kandidat för värdering av IT-säkerhet under de senare stegen i IT-livscykelmodellen.

Tabell 14: Rekommenderade metodgrupper för de olika stegen i Försvarsmaktens IT-livscykelmodell.

Steg	Metodgrupp
B1	Hantering av osäkerhet i underlag för värdering / Tillståndsmodeller och simulering
B2	Hantering av osäkerhet i underlag för värdering
B3	Hantering av osäkerhet i underlag för värdering
B4	Attackgrafer
B5	Attackgrafer / Hantering av osäkerhet i underlag för värdering
B6	Attackgrafer
B7	Attackgrafer

7.2 Kommentarer avseende metodtest

Utfallet från testningen visar på att de olika metodgrupperna är mer eller mindre relevanta för de olika stegen i IT-livscykelmodellen. En jämförelse av de vikter som attributen har för de olika stegen ger vid handen att det är flera attribut som inte direkt varierar beroende på vilket steg som avses. Det är inte rättframt att se detta direkt i tabellen eftersom vikternas storlek också beror på hur många attribut som ska dela på vikten inom en kategori.

En slutsats är dock att en del attribut inte är tillräckligt specifika för att ge utslag vid testning av metodgrupperna. Exempelvis är tekniska faktorer av betydelse under alla stegen, även om de behandlas på olika abstraktionsnivå. En uppdelning i mer specifika attribut skulle dock kunna ge mer variation hos resultaten.

På samma sätt skulle en uppdelning i mer specifika behov leda till ökad variation avseende vilka attribut som väljs ut för en testomgång. Därmed skulle utfallet för de olika metodgrupperna bli annorlunda när det är mer specifika användningsfall som avses.

Tydligare, mer välspecificerade metoder som beskriver hela värderingsförfarandet skulle sannolikt ge högre relevansvärden för flera av de identifierande metodgrupperna. Detta gäller speciellt för metodgruppen *aggregering av värden*.

En noggrann undersökning av huruvida en värderingsmetod är relevant i ett specifikt användningsfall ligger utanför syftet med TSAR. Därmed innehåller TSAR inte några kortslutande attribut, dvs. attribut vilka om de inte uppfylls

automatiskt skulle ge relevansvärdet noll. Detta gör att irrelevanta metoder fortfarande kan få relativt höga relevansvärden.

7.3 Slutsatser

Som formuleringen av metodgrupperna visar finns det ett antal framtida angreppssätt avseende mer noggrann värdering av IT-säkerhet. Det kommer att vara av stor vikt att klargöra vilka behov som finns av värdering inom Försvarsmakten så att de mest lämpliga ansatserna kan väljas. Att avgöra vilka ansatser till värdering av IT-säkerhet som är mest lämpliga kräver dock att de omkringliggande processerna finns på plats.

Det kommer därför inte att uppstå någon universell metod som kan nyttjas i alla situationer. Olika situationer kommer inte bara att kräva olika metoder utan även vitt skilda ansatser till värdering.

En stor andel av de studerade ansatserna, inom alla metodgrupper, nyttjar expertutlåtanden som indata. Detta gör att de är flexibla och, helt eller delvis, applicerbara i alla steg i IT-livscykelmodellen. Nyttjandet av expertutlåtanden medför dock att metoderna blir beroende av subjektiva indata.

Nyttjandet av formella metoder, vilket är speciellt vanligt förekommande i metodgruppen *hantering av osäkerhet i underlag för värdering*, borde leda till bra spårbarhet från resultat till bakomliggande orsaker. Subjektiva indata gör det dock svårt att spåra anledningen till presenterade resultat, det vill säga avgöra vilka de objektiva bakomliggande orsakerna till resultatet kan vara.

Att hantera subjektivitet utgör en av de stora utmaningarna för forskningen kring metoder för värdering av IT-säkerhet. Objektiva indata kräver subjektiva tolkningar och indata baserade på expertutlåtanden medför att hela värderingen baseras på subjektiva värderingar.

Effektiv och ändamålsenlig värdering av IT-säkerhet är en förutsättning för att kunna dra fördel av de möjligheter IT-system ger, samt kunna hantera de risker som uppstår i och med nyttjandet. Därmed är det av stor vikt att tillgodogöra sig de metoder för värdering av IT-säkerhet som framkommer. Exempelvis kommer ett effektivt användande av strukturerade metoder för hantering av informationssäkerhetsrisker att vara avgörande för förmågan att följa IT-livscykelmodellen och därmed få ett effektivt nyttjande av informationssystem.

Referenser

[1] R.J. Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems*, Wiley, 2008.

[2] D. Gollmann, *Computer security*, Chichester: Wiley, 2006.

[3] J. Hallberg, A. Hunstad, and N. Hallberg, *Handbok för IT-säkerhetsvärdering*, Linköping, Sverige: FOI, 2007.

[4] Försvarsmakten, *Handbok för Försvarsmaktens Säkerhetstjänst, Informationsteknik (H SÄK IT)*, Försvarsmakten, 2006.

[5] Nationalencyklopedin, "Informationssystem," *Nationalencyklopedin*, 2008.

[6] SIS, *SIS HB 550: Terminologi för informationssäkerhet, utgåva 3*, SIS Förlag, 2007.

[7] J. Bengtsson, K. Lundholm, J. Hallberg, A. Hunstad, and J. Löfvenberg, *The TSAR procedure rev. 1: Test of Security Assessment Relevance*, Swedish Defence Research Agency, FOI, 2010.

[8] Försvarsmakten, *Direktiv för Försvarsmaktens informationsteknikverksamhet*, 2004.

[9] X. Cuihua and L. Jiajun, "An Information System Security Evaluation Model Based on AHP and GRAP," *Web Information Systems and Mining, 2009. WISM 2009. International Conference on*, 2009, pp. 493 -496.

[10] W. hua Jiang, J. Du, and F. Liu, "Research on System Integrative Evaluation Method Based on Survivability," *Advanced Information Networking and Applications - Workshops, 2008. AINAW 2008. 22nd International Conference on*, 2008, pp. 573 -577.

[11] H. Wang, S. Roy, A. Das, and S. Paul, "A framework for security quantification of networked machines," *Communication Systems and Networks (COMSNETS), 2010 Second International Conference on*, 2010, pp. 1 -8.

[12] X. Wu, J. Li, and W. Yao, "A Network Security Evaluation Model based on Common Criteria," *Apperceiving Computing and Intelligence Analysis, 2008. ICACIA 2008. International Conference on*, 2008, pp. 416 -420.

[13] T. Sommestad, M. Ekstedt, and P. Johnson, "Cyber Security Risks Assessment with Bayesian Defense Graphs and Architectural Models," *System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on*, 2009, pp. 1 -10.

[14] L.X. qin and C. Xia, "Research on Network Security Risk Assessment Based on Group Consistency," *Software Engineering, 2009. WCSE '09. WRI World Congress on*, 2009, pp. 425 -428.

[15] L. Williams, R. Lippmann, and K. Ingols, "GARNET: A Graphical Attack Graph and Reachability Network Evaluation Tool," 2008.

[16] F. Cheng, C. Wolter, and C. Meinel, “A Simple, Smart and Extensible Framework for Network Security Measurement,” 2008.

[17] L. Nordstrom, “Assessment of Information Security Levels in Power Communication Systems Using Evidential Reasoning,” *Power Delivery, IEEE Transactions on*, vol. 23, 2008, pp. 1384 -1391.

[18] F. Xuewei, W. Dongxia, M. Guoqing, and L. Jin, “Security Situation Assessment Based on the DS Theory,” *Education Technology and Computer Science (ETCS), 2010 Second International Workshop on*, 2010, pp. 352 -356.

[19] Y. Qing, Z. Changhong, W. Xiaoping, and Z. Dingjun, “Information Security Risk Assessment Based on AHP/DST,” *Management and Service Science, 2009. MASS '09. International Conference on*, 2009, pp. 1 -4.

[20] H. Gao, J. Zhu, and C. Li, “The analysis of uncertainty of network security risk assessment using Dempster-Shafer theory,” *Computer Supported Cooperative Work in Design, 2008. CSCWD 2008. 12th International Conference on*, 2008, pp. 754 -759.

[21] L. Kong, X. Ren, and Y. Fan, “Study on assessment method for computer network security based on rough set,” *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*, 2009, pp. 617 -621.

[22] D. Zhao, J. Liu, and Z. Zhang, “Method of risk evaluation of information security based on neural networks,” *Machine Learning and Cybernetics, 2009 International Conference on*, 2009, pp. 1127 -1132.

[23] Q. Ma, X. Tan, D. Kong, and H. Xi, “A novel approach to network vulnerabilities quantitative evaluation based on Mamdani-Style fuzzy logic,” *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, 2008, pp. 3327 -3330.

[24] G. Lu, Z. Chen, X. He, and J. Li, “A Method of Security Evaluation based on Fuzzy Mathematics,” *Apperceiving Computing and Intelligence Analysis, 2008. ICACIA 2008. International Conference on*, 2008, pp. 106 -109.

[25] Y. Fu, Y. Qin, and X. Wu, “A Method of Information Security Risk Assessment Using Fuzzy Number Operations,” *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on*, 2008, pp. 1 -4.

[26] X. Long, Q. Yong, and L. Qianmu, “Information Security Risk Assessment Based on Analytic Hierarchy Process and Fuzzy Comprehensive,” *Risk Management Engineering Management, 2008. ICRMEM '08. International Conference on*, 2008, pp. 404 -409.

[27] Y. Shuping and G. Yingyan, “Network security situation quantitative evaluation based on the classification of attacks in attack-defense confrontation environment,” *Control and Decision Conference, 2009. CCDC '09. Chinese*,

2009, pp. 6014 -6019.

[28] L. Xiaowu, Y. Jiguo, and W. MaoLi, “Network Security Situation Generation and Evaluation Based on Heterogeneous Sensor Fusion,” *Wireless Communications, Networking and Mobile Computing, 2009. WiCom '09. 5th International Conference on*, 2009, pp. 1 -4.

[29] L. Mixia, Z. Qiuyu, Z. Hong, and Y. Dongmei, “Network Security Situation Assessment Based on Data Fusion,” *Knowledge Discovery and Data Mining, 2008. WKDD 2008. First International Workshop on*, 2008, pp. 542 -545.

[30] J. Si, Q. Zhang, D. Man, K. Wang, and W. Yang, “Network threat assessment based on attribute recognition,” *Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on*, 2009, pp. 1482 -1486.

[31] L. Huiying and C. Yuanda, “Research on Network Risk Situation Assessment Based on Threat Analysis,” *Information Science and Engieering, 2008. ISISE '08. International Symposium on*, 2008, pp. 252 -257.

[32] Y. Ji, D. Wen, H. Wang, and C. Xia, “A logic-based approach to network security risk assessment,” *Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on*, 2009, pp. 9 -14.

[33] W. He, C. Xia, H. Wang, C. Zhang, and Y. Ji, “A Game Theoretical Attack-Defense Model Oriented to Network Security Risk Assessment,” *Computer Science and Software Engineering, 2008 International Conference on*, 2008, pp. 498 -504.

[34] V. Vijayaraghavan and S. Paul, “iMeasure Security (iMS): A Novel Framework for Security Quantification,” *Networks and Communications, 2009. NETCOM '09. First International Conference on*, 2009, pp. 414 -421.

[35] M. Tupper and A. Zincir-Heywood, “VEA-bility Security Metric: A Network Security Analysis Tool,” *Availability, Reliability and Security, 2008. ARES 08. Third International Conference on*, 2008, pp. 950 -957.

[36] A. Xie, W. Wen, L. Zhang, J. Hu, and Z. Chen, “Applying Attack Graphs to Network Security Metric,” *Multimedia Information Networking and Security, 2009. MINES '09. International Conference on*, 2009, pp. 427 -431.

[37] V. Franqueira, P. van Eck, R. Wieringa, and R. Lopes, “A Mobile Ambients-Based Approach for Network Attack Modelling and Simulation,” *Availability, Reliability and Security, 2009. ARES '09. International Conference on*, 2009, pp. 546 -553.

[38] C. Feng and S. Jin-Shu, “A Flexible Approach to Measuring Network Security Using Attack Graphs,” *Electronic Commerce and Security, 2008 International Symposium on*, 2008, pp. 426 -431.

[39] D. Man, B. Zhang, W. Yang, W. Jin, and Y. Yang, “A Method for Global Attack Graph Generation,” *Networking, Sensing and Control, 2008. ICNSC*

2008. IEEE International Conference on, 2008, pp. 236 -241.

[40] L. Wang, T. Islam, T. Long, A. Singhal, and S. Jajodia, “An Attack Graph-Based Probabilistic Security Metric,” 2008.

[41] N. Ghosh and S. Ghosh, “An Approach for Security Assessment of Network Configurations Using Attack Graph,” *Networks and Communications, 2009. NETCOM '09. First International Conference on*, 2009, pp. 283 -288.

[42] D. Leversage and E. James, “Estimating a System's Mean Time-to-Compromise,” *Security Privacy, IEEE*, vol. 6, 2008, pp. 52 -60.

[43] X. Xiao, H. Wang, and G. Zhang, “Access Graph Based Risk Assessment Model for Network Information System,” *Frontier of Computer Science and Technology, 2008. FCST '08. Japan-China Joint Workshop on*, 2008, pp. 42 -48.

[44] X. Xiao, T. Zhang, and G. Zhang, “Using Access Graph to Analyze Network Vulnerabilities,” *Computational Intelligence and Industrial Application, 2008. PACIIA '08. Pacific-Asia Workshop on*, 2008, pp. 781 -786.

[45] X. Chen, Q. Zheng, and X. Guan, “An OVAL-based active vulnerability assessment system for enterprise computer networks,” *Information Systems Frontiers*, vol. 10, 2008, pp. 573-588.

[46] X. Chen, J. Li, and S. Zhang, “Study of generating attack graph based on privilege escalation for computer networks,” *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on*, 2008, pp. 1213 -1217.

[47] E. Hulitt and R. Vaughn, “Information system security compliance to FISMA standard: A quantitative measure,” *Computer Science and Information Technology, 2008. IMCSIT 2008. International Multiconference on*, 2008, pp. 799 -806.

[48] E. Hulitt and R. Vaughn, “Information system security compliance to FISMA standard: a quantitative measure,” *Telecommunication Systems*, 2009.

[49] M. Ahmed, E. Al-Shaer, and L. Khan, “A Novel Quantitative Approach For Measuring Network Security,” *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, 2008, pp. 1957 -1965.

[50] R.W. Hamming, “Error Detecting and Error Correcting Codes,” *The Bell System Technical Journal*, vol. XXIX, Apr. 1950.

[51] T. Sundmark, K. Lundholm, J. Bengtsson, and J. Hallberg, *Test of Security Assessment Relevance - Demonstrator*, Linköping, Sverige: FOI, 2010.

Bilaga A Attribut

Innehållet i denna bilaga är en kopia av Appendix A i metodbeskrivningen av TSAR-proceduren[1]. Den återges här i sin helhet för att underlätta för läsaren.

This appendix presents a set of characteristics which can be used to perform tests using the TSAR procedure. An overview of the characteristics is presented as a tree in Figure 1.

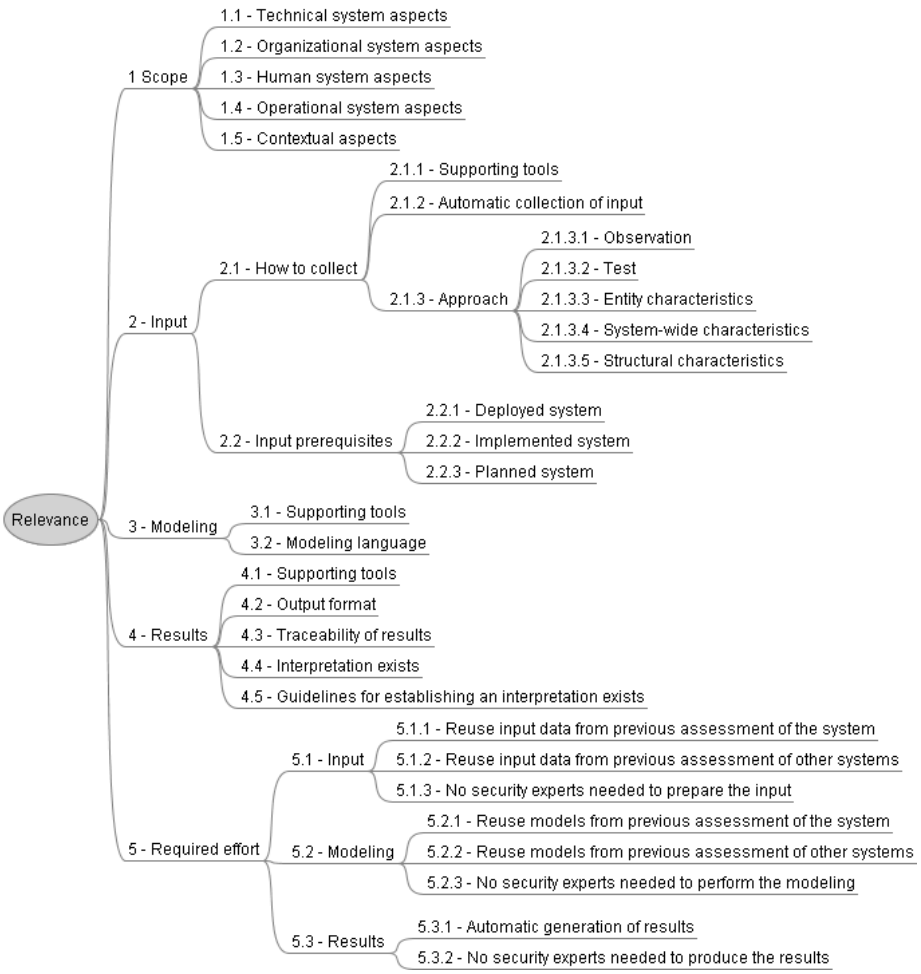


Figure 1: Overview of the proposed set of characteristics.

[1] Bengtsson et al, The TSAR procedure rev. 1: Test of Security Assessment Relevance, Swedish Defence Research Agency, FOI, 2010

1 Scope

This is a compound characteristic concerning what the scope for the security assessment is.

The scope of the system security assessment can be limited to one or more system aspects. For example, focusing on technical system aspects will result in a radically different assessment than an assessment based on purely organizational aspects.

1.1 Technical system aspects

This characteristic specifies whether the scope of the assessment includes technical system aspects. Technical system aspects regard if the impact of technical artifacts on the IT security is considered when the assessment is performed.

Technical artifacts include, for example, the choice and configuration of:

- firewalls
- computers
- routers
- web services
- communication protocols

1.2 Organizational system aspects

This characteristic specifies whether the scope of the assessment includes organizational system aspects. Organizational system aspects regard if the impact of organizational artifacts on the IT security is considered when the assessment is performed.

Organizational artifacts include policies, processes, procedures, and routines for information security work.

1.3 Human system aspects

This characteristic specifies whether the scope of the assessment includes human factors. Human system aspects regard whether the impact on the IT security resulting from the interaction between humans and the technical system is considered.

Human factors include the usability of security functions, level of security awareness and training, and employee satisfaction.

1.4 Operational system aspects

This characteristic specifies whether the scope of the assessment includes operational system aspects. Operational system aspects regard how the IT security of the system is affected by the fact that it is in operation.

Operational factors include how information security work is actually carried out, like for example the time it takes until security updates are installed.

The difference to organizational factors is that the operational factors consider the actual work, while the organizational factors consider the routines for said work.

1.5 Contextual aspects

This characteristic specifies whether the scope of the assessment includes contextual aspects. This regards if the impact of contextual artifacts on the possibility to achieve certain levels of IT security is considered when the assessment is performed.

Contextual artifacts refer to factors, which are not part of the system, but still indirectly affect the IT security of the actual system. This can for example include physical protection, laws and regulations.

2 Input

This is a compound characteristic for specifying demands on the input to the security assessment method.

2.1 How to collect

This is a compound characteristic concerning how the input data for the assessment method is collected.

2.1.1 Supporting tools

This characteristic specifies whether the assessment method has explicitly mentioned tools for collecting the input. Supporting tools for the collection of input includes:

- Software for semi-automatic data collection
- Documented procedures for data collection in a structured way

An example of a semi-automatic data collection is when the collected data needs to be processed in Excel before being used for the actual security assessment. Vulnerability scanners, Log analysis tools, Debuggers and AHP are other examples of semi-automatic data collection.

2.1.2 Automatic collection of input

This characteristic specifies whether there exist tools for automatic collection of input.

This characteristic regards reoccurring assessments where the input is automatically collected after an initial configuration. This requires software with fully automated collection of the input data. The software can collect input directly from the system or via other tools such as vulnerability scanners.

2.1.3 Approach

This is a compound characteristic for how the input data to the assessment method is obtained.

2.1.3.1 Observation

This characteristic specifies whether input data is collected through observations. Examples of input data collected through observations are:

- Number of virus infections per time unit
- Number of intrusions per time unit
- Number of confidentiality violations per time unit
- System logs

2.1.3.2 Test

This characteristic specifies whether input data is collected through tests. Examples of tests that can be used are:

- Vulnerability scanners
- Red teams
- Code inspection

2.1.3.3 Entity characteristics

This characteristic specifies whether input data is collected based on entity characteristics.

Entities are subjects, objects or subsystems that perform tasks in a system and the tasks themselves. These can be described by performance characteristics, interfaces etc.

Examples of entities are:

- Computers
- Implemented security functions

2.1.3.4 System-wide characteristics

This characteristic specifies whether input data is collected based on system-wide characteristics.

Examples of system-wide characteristics are:

- Systems ability to withstand attacks

- Update policies and their implications
- Number of computers
- Number of users

2.1.3.5 Structural characteristics

This characteristic specifies whether input data includes structural characteristics.

Examples of structural characteristics are:

- Interactions between entities
- Network routing

2.1.4 Specification of how to collect

The assessment method should state how the input data is to be collected.

Just stating that input will be gathered through brainstorming or by using vulnerability scanners is not sufficient for this characteristic to be fulfilled.

If brainstorming is to be used there should be a definition of what the topic for the brainstorming should be, what the session is supposed to produce, how the produced data should be represented, and how to use the results.

For methods using vulnerability scanners there should be a description of how the output from the scanner should be converted to input for the method.

2.2 Input prerequisites

This is a compound characteristic concerning what stage of development the assessed system must be in for the assessment method to be applicable.

2.2.1 Deployed system

This characteristic concerns whether it is possible to collect the input data needed by the assessment method from a deployed system. That is, the assessment considers aspects of deployed systems affecting the security.

Deployed systems are those that produce services in a live environment. That is, systems that are relied upon to perform a service.

2.2.2 Implemented system

This characteristic concerns whether it is possible to collect the input needed by the assessment method from an implemented system. That is, the assessment considers aspects of the system, which can be partially or fully implemented.

A fully implemented system is a system which is ready for deployment.

Partially implemented systems can be prototypes of systems or functional sub-components of a system. An example of this is systems under development.

2.2.3 Planned systems

This characteristic concerns whether it is possible to collect the input needed by the assessment method from descriptions of systems that do not yet exist but are planned to be implemented.

Planned systems can be described as, for example:

- System requirements documentation
- Network topology
- System specification

3 Modeling

This is a compound characteristic about how to perform the modeling required for the assessment method.

3.1 Supporting tools

This characteristic concerns the availability of tools that facilitate the modeling required by the assessment method. The modeling is part of the actual assessment. Thus, the tools used for modeling should be specified in the description of the assessment method.

Examples of modeling tools are Microsoft Visio, MooD and Sparx.

3.2 Modeling language

This characteristic specifies whether created models are described using a standardized or de-facto standard modeling language.

Established modeling languages can for example be:

- Unified Modeling Language (UML)
- Finite state machines

4 Results

This compound characteristic concerns how the result from the assessment method is produced as well as how this result is presented. Further, the interpretation of the result is considered.

4.1 Supporting tools

This characteristic specifies whether tools exist for producing the results of the assessment method. The tools used for producing the assessment results should be specified in the description of the assessment method.

4.2 Output format

This characteristic concerns whether the results from supporting tools are available in an export-friendly format.

Examples of export-friendly formats are:

- XML
- Text file
- Excel

4.3 Traceability of results

This characteristic concerns whether the results produced by the assessment method can be traced back to the factors that caused the results. In order to fulfill this characteristic it should be possible to identify the reasons for the outcome of the assessment.

In a method without traceability of results the assessment states the security level, but not why the security is at this level. A method with traceability is able to motivate the results.

4.4 Interpretation exists

This characteristic concerns whether there is a predefined way to interpret the result from the assessment.

An assessment method can be considered to have a predefined interpretation of the results if there is a mapping from the possible results to qualitative statements about the security.

For example, a method producing a security value X in the range $0 \leq X \leq 1$ is supplemented with the following predefined interpretations of the result:

$0 \leq X < 0.6$: Needs immediate attention

$0.6 \leq X < 0.9$: Should be further checked up on

$0.9 \leq X \leq 1$: No action required

Another type of predefined interpretation is when an assessment tool has the interpretation as output. The result from the assessment method can for example be a list of actions that should be performed to reach an adequate level of security.

4.5 Guidelines for establishing interpretation exists

This characteristic concerns whether there are guidelines for establishing an interpretation of the assessment results.

Guidelines for establishing an interpretation can for example be an algorithm for producing the interpretation based on the assessment results and additional data. Examples of additional data are:

- Statistics on assessment results from other systems
- Previous assessment results

- Security needs of system stakeholders

The difference to a predefined interpretation is that the guidelines do not tell the user how to interpret the assessment results, but rather helps the user to establish a basis for the interpretation.

5 Required effort

This compound characteristic concerns how much effort is required to perform the security assessment. The term effort includes both work hours needed as well as direct financial costs.

5.1 Input

This compound characteristic concerns the effort required to collect the input data.

5.1.1 Reuse input data from previous assessments of the system

This characteristic specifies whether it is possible to reuse input data from previous security assessments of the system. That is, only changes to the system needs to be considered.

Reused data is static properties in the system that has not changed since the last assessment. This data can for example describe:

- Characteristics of system hardware
- Classification of information

5.1.2 Reuse input data from previous assessments of other systems

This characteristic specifies whether it is possible to reuse input data from previous assessments of other systems.

Reused data could for example describe:

- Components that the systems have in common
- Rules or regulations that apply to the systems

5.1.3 No security experts needed to prepare the input

This characteristic specifies whether preparation of the input to the security assessment can be performed *without* consulting experts.

The characteristic should be considered fulfilled if the assessment method does not require experts to prepare the input to the method.

5.2 Modeling

This compound characteristic concerns the effort required to create the models required by the assessment method.

5.2.1 Reuse models from previous assessments of the system

This characteristic specifies whether it is possible to reuse models from previous security assessments of the system.

Reused data could for example be:

- System models from previous assessments which can be modified to reflect the current system
- Prioritizations of security relevant characteristics influencing the assessment results

5.2.2 Reuse models from previous assessments of other systems

This characteristic specifies whether it is possible to reuse models from previous security assessments of other systems.

Reused data could for example be:

- System models from previous assessments which can be modified to reflect the current system
- Prioritizations of security relevant characteristics influencing the assessment results

5.2.3 No security experts needed to perform the modeling

This characteristic specifies whether the modeling required during the assessment can be performed *without* consulting experts. Examples of required modeling include:

- Modeling of systems to be assessed
- Specification of the computations of assessment results

The characteristic is not fulfilled if the assessor, that is, the person responsible for the assessment, has to resort to consulting experts in order to be able to perform the required modeling. In some cases, it may be difficult to draw the line between the processes of acquiring input and consulting security experts. For example, acquiring the number of virus infections during the last month would be considered input data. On the other hand, acquiring data regarding the prioritization of security-relevant system characteristics would be considered to be support from security experts.

5.3 Results

This compound characteristic concerns the effort required to produce and interpret the assessment results.

5.3.1 Automatic generation of results

This characteristic concerns whether the results from the security assessment are produced automatically, without any involvement of humans. The method may require initial, manual setup.

One example is a network scanner that automatically detects and reports on vulnerabilities that needs to be fixed.

5.3.2 No security experts needed to produce the results

This characteristic specifies whether the assessment results can be generated *without* consulting security experts. Examples of tasks that need to be performed without consulting security experts are:

- Computation of aggregated security values
- Interpretation of the computed results
- Presentation of the computed results

Bilaga B Mätvärden

I denna bilaga återfinns de mätvärden som erhölls vid test av de fyra metodgrupperna *attackgrafer* (Tabell B 1), *hantering av osäkerhet i underlag för värdering* (Tabell B 2), *tillståndsmodeller och simulering* (Tabell B 3) samt *aggregering av värden* (Tabell B 4).

Tabell B 1: Mätvärden för metodgruppen attackgrafer.

	Uppfyllnad	Täckning
1.1	1	1
1.2	0	1
1.3	0	0
1.4	0	1
1.5	0	1
2.1.1	1	1
2.1.2	0	0
2.1.3.1	1	1
2.1.3.2	1	1
2.1.3.3	1	1
2.1.3.4	0	0
2.1.3.5	1	1
2.2.1	1	1
2.2.2	1	1
2.2.3	0	0
3.1	1	1
3.2	0	0
4.1	1	1
4.2	0	0
4.3	1	1
4.4	0	0
4.5	1	1
5.1.1	1	1
5.1.2	1	1
5.1.3	0	1
5.2.1	1	1
5.2.2	0	1
5.2.3	0	1
5.3.1	1	1
5.3.2	0	1

Tabell B 2: Mätvärden för metodgruppen hantering av osäkerhet i underlag för värdering.

	Uppfyllnad	Täckning
1.1	1	1
1.2	0	0
1.3	0	0
1.4	1	1
1.5	0	0
2.1.1	1	1
2.1.2	0	1
2.1.3.1	1	1
2.1.3.2	0	1
2.1.3.3	1	1
2.1.3.4	1	1
2.1.3.5	1	1
2.2.1	1	1
2.2.2	1	1
2.2.3	1	1
3.1	1	1
3.2	0	0
4.1	0	0
4.2	0	0
4.3	1	1
4.4	0	1
4.5	1	1
5.1.1	1	1
5.1.2	1	1
5.1.3	0	1
5.2.1	1	1
5.2.2	1	1
5.2.3	0	1
5.3.1	0	0
5.3.2	0	1

Tabell B 3: Mätvärden för metodgruppen tillståndsmodeller och simulering.

	Uppfyllnad	Täckning
1.1	1	1
1.2	0	0
1.3	0	0
1.4	1	1
1.5	0	1
2.1.1	0	0
2.1.2	0	1
2.1.3.1	1	1
2.1.3.2	1	1
2.1.3.3	1	1
2.1.3.4	0	1
2.1.3.5	1	1
2.2.1	1	1
2.2.2	1	1
2.2.3	0	1
3.1	0	0
3.2	1	1
4.1	0	0
4.2	0	0
4.3	1	1
4.4	0	0
4.5	1	1
5.1.1	1	1
5.1.2	1	1
5.1.3	0	1
5.2.1	1	1
5.2.2	1	1
5.2.3	0	1
5.3.1	0	0
5.3.2	0	1

Tabell B 4: Mätvärden för metodgruppen aggregering av värden.

	Uppfyllnad	Täckning
1.1	1	1
1.2	0	0
1.3	0	0
1.4	1	1
1.5	0	0
2.1.1	1	1
2.1.2	0	0
2.1.3.1	1	1
2.1.3.2	0	0
2.1.3.3	1	1
2.1.3.4	0	0
2.1.3.5	1	1
2.2.1	1	1
2.2.2	1	1
2.2.3	0	0
3.1	0	0
3.2	0	0
4.1	0	0
4.2	0	0
4.3	1	1
4.4	0	0
4.5	1	1
5.1.1	1	1
5.1.2	1	1
5.1.3	0	1
5.2.1	1	1
5.2.2	1	1
5.2.3	0	1
5.3.1	0	0
5.3.2	0	1