Beskriva och bedöma hot mot IT-system

Varför är det svårt?

Jonas Hallberg, Johan Bengtsson och Henrik Karlzén

FOI-R--4330--SE

December 2016

FOI

Jonas Hallberg, Johan Bengtsson, Henrik Karlzén

Beskriva och bedöma hot mot IT-system

Varför är det svårt?

Bild/Cover: Johan Bengtsson

Titel	Beskriva och bedöma hot mot IT-system – Varför är det svårt?
Title	Describe and assess threats to IT systems – Why is is so difficult?
Rapportnr/Report no	FOI-R--4330--SE
Månad/Month	December
Utgivningsår/Year	2016
Antal sidor/Pages	75
ISSN	1650-1942
Kund/Customer	Försvarsmakten
Forskningsområde	4. Informationssäkerhet och kommunikation
FoT-område	Ledning och MSI
Projektnr/Project no	E72628
Godkänd av/Approved by	Christian Jönsson
Ansvarig avdelning	Ledningssystem

Sammanfattning

Att beskriva hot mot IT-system och bedöma dessa hots sannolikhet och konsekvens är en krävande uppgift. Under 2016 har FoT-projektet *Bedömning och analys av IT-system* genomfört studier avseende svårigheter med att beskriva och bedöma hot mot IT-system. Inom studierna genomfördes två experiment med fokus på aspekter som kunde tänkas påverka samstämmigheten mellan olika personers bedömning av sannolikhet och konsekvens för hot.

Det första experimentet fokuserades på om hotbeskrivningens utformning påverkar bedömningarna av sannolikhet och konsekvens. Hotbeskrivningarna utgjordes av strukturerade tabeller respektive löpande text. Inga statistiskt signifikanta skillnader kunde dock ses i resultatet från experimentet.

Det andra experimentet byggde vidare på ett experiment som genomfördes i projektet under 2015. Experimentet genomfördes för att identifiera eventuella skillnader i samstämmighet när bedömningar görs enligt metoden för säkerhetsanalys (H Säk Grunder) respektive en metod baserad på jämförelsebaserade bedömningar. När jämförelsebaserade bedömningarna genomfördes av experter, med såväl god kunskap inom området informationssäkerhet som erfarenhet av att bedöma informationssäkerhetsrisker, uppnåddes statistiskt signifikant högre samstämmighet för bedömningarna av både sannolikhet och konsekvens.

Rapportens avslutas med slutsatser att beakta vid Försvarsmaktens fortsatta arbete med säkerhetsanalyser. Dessa slutsatser relaterar till utformningen av hotbeskrivningar och förutsättningar för bedömning av hots sannolikhet och konsekvens.

Nyckelord: Säkerhetsanalys, hotbeskrivning, sannolikhet, konsekvens, AHP

Summary

It is a demanding task to describe the threats to IT systems and to assess the likelihood and consequence of these threats. This report describes studies related to the difficulties in describing and assessing threats to IT systems. Within the studies, two experiments were conducted focusing on aspects that could affect the correlation between different people's assessment of the likelihood and consequence of threats.

The first experiment focused on whether the presentation of the threat description affects the assessments of likelihood and consequence. The threat presentations consisted of either structured tables or running text. No statistically significant differences could be found in the results from the experiment.

The second experiment was based on an experiment conducted in the project in 2015. The experiment was carried out to identify any differences in interrater consistency when assessments are made according to the method *Säkerhetsanalys* (Swedish Armed Forces) compared to a comparison-based assessment method. When comparison-based assessments were conducted by experts, having good knowledge in the field of information security as well as experience in assessing information security risks, the results showed statistically significant higher consistency of the assessments regarding both likelihood and consequence.

The report ends with findings related to the presentation of threat descriptions and prerequisites for the assessment of threat likelihood and consequence.

Keywords: Risk assessment, threat description, likelihood, consequence, AHP

Innehåll

1 Inledning

Det är av stor vikt för arbetet med informationssäkerhet att beskriva och bedöma de hot som riktas mot IT-system. Det har tagits fram ett flertal metoder och ramverk vars syfte är att stödja arbetet med informationssäkerhetsrisker (Pan & Tomlinson, 2016). Dessa metoder och ramverk skiljer sig åt med avseende på vilken information genomförandet av analyserna baseras på (Korman, Sommestad, Hallberg, Bengtsson, & Ekstedt, 2014). En aspekt som dock inte har studerats i någon större omfattning är på vilket sätt identifierade hot ska beskrivas för att bäst stödja bedömningen av dem.

Inom ramen för de säkerhetsanalyser som genomförs för Försvarsmaktens planerade IT-system bedöms informationssäkerhetsrisker. Dessa bedömningar görs enligt ett tillvägagångssätt som liknar merparten av andra föreslagna metoder och ramverk. Tillvägagångssättet innefattar bedömningar av med vilken sannolikhet identifierade hot kommer att realiseras och hur allvarlig konsekvensen skulle vara om hoten realiseras. Att ta fram adekvata bedömningar av sannolikheter och konsekvenser för hot mot informationssystem utgör en stor utmaning, bland annat på grund av avsaknaden av kvantitativa underlag (Fenz, Heurix, Neubauer, & Pechstein, 2014; Sommestad, Karlzén, Nilsson, & Hallberg, 2016).

I denna rapport beskrivs de studier som under 2016 har genomförts i FoT-projektet *Bedömning och analys av IT-system* (BAIT). Studierna under 2016 har haft två syften:

- att undersöka på vilket sätt hot ska beskrivas för att stödja bedömning av dessa
- att bygga vidare på de resultat kring jämförelsebaserade metoder för bedömning av sannolikhet och konsekvens som togs fram under 2015.

Utgående från dessa två syften har frågor kring beskrivning av hot och jämförelsebaserade metoder studerats. Inom ramen för dessa studier har två experiment utformats och genomförts.

1.1 Terminologi

I detta avsnitt presenteras begrepp som är centrala i rapporten samt de statistiska koncept som används i Kapitel 3–5.

1.1.1 Centrala begrepp

De centrala begreppen frekvens, hot, konsekvens, risk och sannolikhet används i rapporten med följande betydelser.

Frekvens definieras som *antalet händelser av en bestämd typ inom ett givet tidsintervall*.

Hot definieras som *möjlig, oönskad händelse med negativa konsekvenser för verksamheten* (Försvarsmakten, 2013; SIS, 2007, 2015).

Konsekvens definieras som ett *resultat av en händelse med negativ inverkan* (Försvarsmakten, 2013; SIS, 2007).

Risk definieras som *kombinationen av sannolikhet för att ett givet hot realiseras och därmed uppkommande skadekostnad* (SIS, 2007).

Sannolikhet används i denna rapport med en övergripande betydelse som inkluderar såväl kvalitativa som kvantitativa beskrivningar. Denna användning av begreppet överensstämmer med det engelska begreppet *likelihood* och dess definition enligt ISO Guide 73:2009 (ISO International Organization for Standardization, 2009), *chance of something happening*, med en kommentar om att *likelihood* kan beskrivas i generella termer eller matematiskt.

1.1.2 Korrelationsmått

Centralt inom de statistiska resonemang som förs i rapporten är följande mått på korrelation (samstämmighet). I enlighet med andra studier inom områden där det saknas korrekta svar att jämföra gjorda bedömningar med studeras istället samstämmigheten hos de avgivna svaren, eftersom samstämmighet är en förutsättning för korrekthet (Shanteau, 2015). I denna rapport används samstämmighet i två olika avseenden. I det första fallet avses hur väl enskilda respondenters bedömning av samma hot vid olika tillfällen stämmer med varandra. I det andra fallet avses hur väl olika respondenters bedömningar av samma hot stämmer med varandra.

Pearsons produktmomentkorrelationskoefficient är ett mått på det linjära beroendet mellan två variabler. Värdet är i intervallet -1 till +1, där +1 innebär perfekt positivt linjärt beroende, 0 innebär avsaknad av linjärt beroende och -1 innebär perfekt negativt linjärt beroende.

Spearmans rangkorrelationskoefficient är ett statistiskt mått på styrkan av ett monotont förhållande mellan två variabler. Ett perfekt förhållande (+1) innebär att när den ena variabeln växer så växer även den andra, dock inte nödvändigtvis linjärt. Värden över 0,80 anses allmänt visa på en mycket stark korrelation[1].

Intra-class correlation (ICC) är en uppsättning mått som kan användas för att beskriva likheten mellan element som har grupperats tillsammans[2]. Dessa mått används bland annat för att värdera bedömares samstämmighet.

[1] http://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf
[2] ICC används i statistikprogrammet SPSS, https://www.ibm.com/software/se/analytics/spss/.

1.2 Läsanvisning

Återstoden av denna rapport disponeras enligt följande. I Kapitel 2 ges en diskussion kring olika svårigheter med att beskriva och bedöma hot mot IT-system. I Kapitel 3 presenteras resultaten från det första experimentet där två olika typer av sätt att beskriva hot jämförs. I Kapitel 4 analyseras olika aspekter av jämförelsebaserade bedömningar. I Kapitel 5 presenteras resultaten från det andra experimentet där bedömningar av sannolikhet och konsekvens enligt metoden för säkerhetsanalys (Försvarsmakten, 2013) ställs mot jämförelsebaserade bedömningar. I Kapitel 6 diskuteras resultaten från årets arbete. I Kapitel 7 presenteras slutligen ett antal slutsatser som är värda att beakta vid Försvarsmaktens fortsatta arbete med säkerhetsanalyser.

2 Svårigheter med att beskriva och bedöma hot

Identifiering och bedömning av hot mot Försvarsmaktens IT-system görs primärt inom arbetet med säkerhetsplanering. Säkerhetsplanering syftar till att systematiskt, utgående ifrån verksamhetens skyddsvärda tillgångar, identifiera och bedöma de risker som föreligger samt fastställa hur dessa risker ska hanteras (Försvarsmakten, 2013). Säkerhetsplanering är något som används för flera olika typer av verksamheter inom Försvarsmakten, bland annat vid övningsverksamhet och inför framtagandet av nya IT-system. I denna rapport fokuserar vi på användandet av metoden för säkerhetsanalys vid framtagandet av säkerhetsmålsättning för nya IT-system.

Säkerhetsplanering (Försvarsmakten, 2013) inleds med att genomföra en säkerhetsanalys som innefattar följande fem steg:

1. identifiera och prioritera skyddsvärda tillgångar
2. bedöm säkerhetshot
3. bedöm sårbarheter
4. bedöm risker
5. prioritera och hantera risker.

Sett till vad säkerhetsanalysen omfattar kan den jämföras med aktiviteter som i andra sammanhang beskrivs med begreppen *riskanalys*, *riskbehandling*, *riskhantering* eller *risk management* (SIS, 2015).

De resultat som produceras under steg 1–3 används i steg 4 som underlag för att bedöma de hot som anses föreligga mot det analyserade IT-systemet. Underlaget innefattar beskrivningar eller bedömningar i form av följande information:

- skyddsvärd tillgång
- oönskad händelse
- konsekvens med tillhörande konsekvensbeskrivning
- aktör (kapacitet, intention, tillfälle)
- hotnivå med tillhörande hotbeskrivning
- existerande skyddsåtgärder
- sårbarheter
- sårbarhetsnivå.

Värt att notera är att hotbeskrivning i denna kontext avser den textuella definitionen av en hotnivå, exempelvis att en bedömd hotnivå på 2 definieras som *lågt hot*. Konsekvensen av att ett hot realiseras bedöms redan i steg 1 enbart baserat på information om skyddsvärd tillgång och oönskad händelse. Mer detaljerad information om hotet, såsom vilken aktör som bedöms kunna realisera hotet, anses alltså inte ha någon inverkan på konsekvensen. Däremot nyttjas mer underlag när sannolikheten för att ett hot realiseras bedöms i steg 4, då all information i listan ovan är tillgänglig att beakta. När analyser genomförs enligt metoden för säkerhetsanalys är alltså det underlag som nyttjas vid bedömningen av sannolikhet och konsekvens olika. Detta är en stor skillnad metodmässigt jämfört med metoden för hot-, risk- och sårbarhetsanalys enligt H SÄK IT (Försvarsmakten, 2006) som tidigare använts för analys av IT-system. Med den tidigare metoden bedömdes både sannolikhet och konsekvens vid samma tillfälle baserat på samma informationsunderlag, vilket motsvarar det underlag som enligt metoden för säkerhetsanalys används vid bedömning av sannolikhet.

Under 2015 genomfördes en behovsanalys avseende behov relaterade till bedömning av sannolikhet och konsekvens för hot mot IT-system (Hallberg, Bengtsson, & Karlzén, 2015) och det identifierades då, bland annat, att det fanns behov av minskat personberoende vid bedömning av sannolikhet och konsekvens. Bedömarens kunskap och erfarenhet är viktiga faktorer när denna typ av bedömningar ska göras (Shanteau, 2015), men det förefaller troligt att även andra parametrar påverkar bedömningen, exempelvis vilket underlag som finns tillgängligt när bedömningen ska göras och hur denna information presenteras.

Att en bedömning är personberoende innebär att resultaten från olika bedömare inte är samstämmiga. Vilket resultat som uppnås beror därav på vem som har gjort bedömningarna. Vad som är orsaken till bristande samstämmighet vid bedömning av sannolikhet och konsekvens är oklart. I följande avsnitt presenteras faktorer som kan tänkas påverka samstämmigheten vid dessa bedömningar.

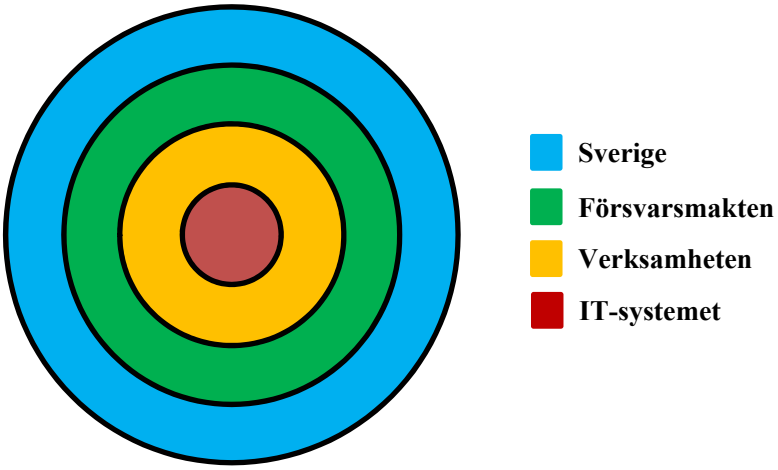
2.1 Identifiering av tillgångar

Det första som genomförs i analysarbetet är att avgöra vilka skyddsvärda tillgångar som finns. Det är oönskade händelser, och i förlängningen hot, som påverkar dessa skyddsvärda tillgångar som analysarbetet ska fokusera på. Om någon tillgång förbises under identifieringen, eller tas med även om den inte är skyddsvärd, påverkas det resterande analysarbete och därmed även slutresultatet.

En svårighet vid identifieringen av skyddsvärda tillgångar under analysen av ett IT-system är att avgöra vilket omfång som ska nyttjas. En tolkning av omfång kan vara att fokusera på tillgångar som ingår i det analyserade IT-systemet. En annan kan vara att även titta på tillgångar i den verksamhet som IT-systemet ska

stödja. Då dagens IT-system är mer sammankopplade än tidigare kan det även vara problematiskt att fastställa systemgränser (Bengtsson, Lundholm, & Hallberg, 2012) och därmed vilka tillgångar som ska ingå i den aktuella analysen och vilka tillgångar som ska ingå vid analys av andra IT-system.

En annan svårighet vid identifieringen är att avgöra ur vems perspektiv tillgången ska vara skyddsvärd för att vara relevant att ta med i analysen (Figur 1). Den mest systemnära och snäva tolkningen är att det ska vara tillgångar som är skyddsvärda för det aktuella IT-systemets funktion. En annan tolkning är att det ska vara tillgångar som är skyddsvärda ur den stödda verksamhetens perspektiv. Ytterligare möjliga tolkningar är att det kan vara tillgångar som är skyddsvärda sett ur hela Försvarsmaktens perspektiv eller för hela riket.



Figur 1: Visualisering av olika omfång av perspektiv.

2.2 Konkretisering av hot

När konkreta hot mot de skyddsvärda tillgångarna ska identifieras finns det olika tillvägagångssätt som kan användas. Att förutsättningslöst genomföra en brainstorming under en workshop kan vara ett sätt att formulera konkreta hot. Ett annat sätt kan vara att låta arbetet utgå från en sedan tidigare framtagen hotlista som kan användas som inspiration till vilka typer av hot som kan tänkas vara aktuella för IT-systemet som ska analyseras. I H SÄK IT Hot (Försvarsmakten, 2001) finns exempelvis en katalog över hot mot IT-system som kan nyttjas som inspiration. Internationellt finns det mer omfattande hotkataloger, som exempelvis IT-Grundschutz (Bundesamt für Sicherheit in der Informationstechnik (BSI), 2005) som omfattar över 2 500 sidor.

I dagsläget är ett vanligt tillvägagångssätt inom Försvarsmakten att låta analysarbetet utgå från hot som har identifierats i tidigare analyser för andra IT-system av liknande karaktär. På detta sätt bygger bedömaren upp en egen hotkatalog som nyttjas som grund vid varje ny analys. Samtidigt som en hotkatalog kan snabba upp arbetsprocessen, då analysarbetet inte börjar om från noll varje gång, så kan det även hämma identifierandet av nya hot eller till och med nya typer av hot som inte återfinns i den befintliga hotkatalogen. Att olika bedömare utgår från olika underlag för att identifiera konkreta hot mot de skyddsvärda tillgångarna kan dessutom påverka den resulterande uppsättningen hot mot, och i slutändan de identifierade riskerna för, det analyserade systemet.

Den uppsättning krav som fås från KSF (Försvarsmakten, 2014) är obligatoriska att uppfylla, vilket resulterar i en grunduppsättning skyddsåtgärder som motsvarar en miniminivå som alla Försvarsmaktens IT-system ska uppfylla. Något som försvårar identifierandet av hot, eller snarare riskerar att bidra till dubbelarbete, är att det inte framgår vilka hot som redan hanteras genom att uppfylla de krav som fås från KSF. En svårighet vid genomförandet av säkerhetsanalysen är dock att information om vilka hot dessa skyddsåtgärder hanterar inte är tillgängligt.

2.3 Bedömningsunderlag

Även om det underlag som har tagits fram och finns tillgängligt vid bedömningstillfället skulle bli likvärdigt för två bedömare så kan deras tidigare erfarenheter påverka både hur informationen i underlaget tillgodogörs och hur den tolkas. Bedömarens kunskap om och erfarenheter av verksamheten som IT-systemet ska stödja kan ha inverkan på både bedömning av sannolikhet och risk. Bättre kunskap om hur verksamheten fungerar kan ge bättre insikt i hur realisering av specifika hot skulle kunna påverka verksamheten.

Hur underlaget presenteras kan också tänkas påverka hur bedömaren tolkar informationen. Att presentera underlaget i form av tabeller är något som görs både i H Säk Grunder (Försvarsmakten, 2013) och i H Säk IT Hot (Försvarsmakten, 2001) när metod och exempel presenteras. Tabellformatet kan tänkas vara fördelaktigt när olika hot ska jämföras då de olika parametrarna (exempelvis tillgång och aktör) presenteras i separata kolumner.

Ett alternativ till att presentera informationen i form av tabeller kan vara att presentera den i form av löpande text, vilket kan upplevas som enklare att läsa, men som samtidigt kan lämna ett visst utrymme för tolkning till läsaren. I tabellformat är det redan någon som gjort en tolkning av vad som exempelvis är den oönskade händelsen eller vem som är aktören i det beskrivna hotet.

Om det är en och samma person som genomför hela analysen så bör inte presentationen av underlaget vara ett problem då personen rimligen väljer att

presentera underlaget på det sätt som personen själv tycker fungerar bäst vid bedömningen. Om det är flera personer som deltar i analysarbetet och är delaktiga i olika delar så kan dock presentationen av underlaget påverka bedömningarna och därmed hela analysresultatet. Trots en gedigen sökinsats har inga relevanta forskningsresultat relaterade till presentation av underlag för hotbedömning hittats.

Vilken detaljnivå som valts vid formulering av oönskade händelser påverkar sannolikheten, eftersom det strikt talat handlar om olika händelser. Om exempelvis den skyddsvärda tillgången är *datorutrustning* och den oönskade händelsen enbart beskrivs som *”förlust av datorutrustning”* så förutsätts inte något specifikt tillvägagångssätt. Den oönskade händelsen kan då omfatta förlust på grund av exempelvis inbrott, brand eller översvämning. Om den oönskade händelsen istället formuleras som *”inbrott i datorlektionssal och stöld av datorutrustning”* så innefattar den oönskade händelsen enbart ett specifikt tillvägagångssätt. Vilken av de två detaljnivåerna som väljs för händelsen påverkar dock inte konsekvensen då den i båda fallen är att tillgången förloras. Däremot påverkas sannolikheten då den första formuleringen avser alla möjliga tillvägagångssätt som kan leda till förlust av tillgången, medan den andra formuleringen enbart avser händelsen att tillgången stjäls vid inbrott i datorlektionssalen. Ett tydliggörande av hur de olika begreppen ska används i bedömningsunderlaget skulle kunna leda till ökad samstämmighet mellan olika bedömare. Avseende formuleringen av den oönskade händelsen i exemplet ovan kan ett tillägg av begreppet *tillvägagångssätt* i bedömningsunderlaget vara fördelaktigt.

2.4 Bedömning av konsekvens och sannolikhet

Bedömningen av konsekvens görs betydligt tidigare när en analys genomförs enligt metoden för säkerhetsanalys jämfört med den äldre metoden för hot-, risk- och sårbarhetsanalys. Detta leder till att det är konsekvensen av att en oönskad händelse inträffar och påverkar en specifik tillgång som bedöms. Konsekvensen blir således kopplad till kombinationen av en tillgång och en oönskad händelse. Andra faktorer, som exempelvis vem aktören är eller vilket tillvägagångssätt som används, har inte någon inverkan på den bedömda konsekvensen.

En svårighet med att bedöma konsekvens är att, på samma sätt som vid identifiering av skyddsvärda tillgångar, avgöra vilket omfång som är relevant att beakta. Med omfång avses om det är konsekvenser för exempelvis IT-systemet, verksamheten som IT-systemet ska stödja, hela Försvarsmakten eller Sverige som ska tas upp i analysen (Figur 1). Sammantaget kan valet av omfång både vid identifiering av skyddsvärda tillgångar och vid bedömning av konsekvens resultera i stora skillnader avseende analysens slutresultat.

Vid bedömning av sannolikheten är det exempelvis en svårighet att bedöma hot som inträffar väldigt sällan, eller som kanske aldrig ens har inträffat någon gång. Om det istället är hot som inträffar regelbundet kan det finnas statistik från tidigare år som kan användas som underlag och därmed underlätta bedömningen av sannolikheten. För hot som inträffar ofta kan det dock vara tydligare att prata om frekvens snarare än om sannolikhet.

Det är otydligt vilken tidshorisont som ska användas vid bedömningen av sannolikheten för att hot realiseras. Skillnaden i tidshorisont skulle exempelvis kunna vara att en person gör sannolikhetsbedömningar för det kommande året, medan en annan person gör bedömningar som avser de kommande tre åren och en tredje person gör bedömningar som avser hela systemets livslängd. Denna otydlighet i hur bedömningarna ska genomföras är en svårighet som kan påverka personberoendet för sannolikhetsbedömningar.

3 Experiment med hotbeskrivningar

Konsekvens- och sannolikhetsbedömningar för hot har visat sig variera beroende på vem som bedömer (Hallberg et al., 2015; Sommestad et al., 2016) och det är ett konstaterat behov i Försvarsmakten att nå fram till mer objektiva bedömningar med minskat personberoende (Hallberg et al., 2015). Det är därför relevant att undersöka vilka faktorer som gör att bedömningarna varierar och om det går att minska skillnaderna genom att justera bedömningsprocessen. Samtidigt är det inte ett självändamål att nå bedömningar som är samstämmiga mellan olika bedömare, om det inte samtidigt gör att bedömningarna närmar sig den faktiska sanningen för vad sannolikheten för att ett hot inträffar är och hur allvarliga konsekvenserna faktiskt kan bli. En förbättrad metod som visar högre samstämmighet och ligger närmare sanningen får heller inte vara alltför kostsam att genomföra, exempelvis med avseende på tidsåtgång och ansträngning.

Mot bakgrunden av diskussionen i 2.3 kring presentation av underlag vid hotbedömning genomfördes ett enkätbaserat experiment som jämförde bedömningar av hot som presenterades i strukturerad tabellform respektive med naturligt språk. Inför experimentet formulerades sju hypoteser indelade i tre grupper:

Hypotes 1. Konsekvensbedömningar

a. Samstämmigheten mellan olika respondenters konsekvensbedömningar är högre när hoten beskrivs i strukturerad tabell än när de beskrivs med naturligt språk.
b. Respondenterna är säkrare på sina konsekvensbedömningar när hoten beskrivs i strukturerad tabell än när de beskrivs med naturligt språk.

Hypotes 2. Sannolikhetsbedömningar

a. Samstämmigheten mellan olika respondenters sannolikhetsbedömningar är högre när hoten beskrivs i strukturerad tabell än när de beskrivs med naturligt språk.
b. Respondenterna är säkrare på sina sannolikhetsbedömningar när hoten beskrivs i strukturerad tabell än när de beskrivs med naturligt språk.

Hypotes 3. Kognitiv belastning för hotbedömningar som helhet

a. Mental ansträngning är lägre när hoten beskrivs i strukturerad tabell än när de beskrivs med naturligt språk.
b. Svårigheten är lägre när hoten beskrivs i strukturerad tabell än när de beskrivs med naturligt språk.
c. Tidsåtgången är lägre när hoten beskrivs i strukturerad tabell än när de beskrivs med naturligt språk.

3.1 Enkätupplägg

För att jämföra de två sätten att presentera hot togs två olika enkäter fram. Den ena enkäten beskrev hoten med naturligt språk, det vill säga i löpande text. Den andra enkäten hade samma hot, men presenterade i en strukturerad tabell nedbrutna i beståndsdelarna aktör, sårbarhet, tillvägagångssätt, tillgång, oönskad händelse samt konsekvensbeskrivning. Båda enkäterna besvarades av samtliga respondenter.

För att inte bedömningarna på respondentens första enkät skulle påverka den andra enkäten besvarades enkäterna med omkring en veckas mellanrum. Dessutom besvarade hälften av respondenterna enkäten med naturligt språk först medan andra hälften besvarade enkäten med strukturerade tabeller först. På detta sätt går det att minska eventuell effekt av att ett sätt att presentera hoten ger en större inlärningseffekt än det andra.

3.1.1 Hoten

För att komma fram till realistiska hot som respondenterna kunde relatera till utvecklades ett scenario (Bilaga 1) där ett fiktivt medelstort svenskt företag inom den finansiella sektorn och med omfattande internationella relationer ska införa ett nytt IT-system för fakturahantering. I enlighet med gängse metodik i Försvarsmakten (Försvarsmakten, 2013) beskrevs kortfattat verksamheten som systemet var avsett att stödja, systemets direkta tillgångar, potentiella oönskade händelser och konsekvenser, hotaktörer, samt möjliga tillvägagångssätt. Antalet hot sattes, baserat på erfarenheter från tidigare experiment (exempelvis (Hallberg et al., 2015; Sommestad et al., 2016)), till 23 stycken för att inte överväldiga respondenterna, men samtidigt uppnå tillräcklig statistisk signifikans. Enkäterna inleddes med en beskrivning av scenariot på ungefär en sida.

I Figur 2 visas ett exempel på hot beskrivet med naturligt språk, medan Figur 3 visar samma hot beskrivet med den strukturerade tabellmetoden. Figur 4 visar de tillhörande frågorna och skalorna som var desamma för båda enkäterna. Stegen i skalan för sannolikhet var inte lika stora utan skalan var mer finmaskig för lägre sannolikheter, i enlighet med H Säk Grunder (Försvarsmakten, 2013).

19. **Stora mängder inloggningsuppgifter läcker ut från en stor webbtjänst. En missnöjd kund som vill straffa Truzio känner igen mejladressen för en anställd på Truzios ekonomiavdelning och provar vid ett besök hos Truzio att logga in i ZAP med uppgifterna, vilket går eftersom den anställde har återanvänt sitt lösenord. Kunden tar därefter bort betalningsorder vilket gör att en del utbetalningar uteblir och bokföringsunderlaget blir felaktigt.**

Figur 2: Exempel på hot beskrivet med naturligt språk.

19.	
Aktör	Missnöjd kund
Sårbarhet	Anställd använder sina inloggningsuppgifter för ZAP även i stor webbtjänst
Tillvägagångssätt	Återanvänt lösenord, som spridits på internet genom läckage från stor webbtjänst, används för att vid besök hos Truzio logga in i ZAP och ta bort betalningsorder.
Tillgång	Betalningsorder
Oönskad händelse	Borttagning av betalningsorder
Konsekvensbeskrivning	Betalning uteblir och bokföringsunderlag blir felaktigt

Figur 3: Exempel på hot beskrivet med den strukturerade tabellmetoden.

Hur stor skulle konsekvensen bli om ett sådant hot skulle inträffa?

1	**2**	**3**	**4**	**5**
Försumbar	Lindrig	Kännbar	Allvarlig	Synnerligen allvarlig

Hur stor är sannolikheten att detta hot realiseras under de kommande 12 månaderna?

1	**2**	**3**	**4**	**5**
Ingen identifierad <1%	Låg 1-5%	Förhöjd 6-25%	Hög 26-50%	Mycket hög >50%

Figur 4: De frågor med tillhörande skalor som användes i båda enkäterna.

3.1.2 Kognitiv belastning och visshet

Kognitiv belastning undersöktes dels med två frågor (baserade på (Paas, 1992) och (Marcus, Cooper, & Sweller, 1996)) där respondenterna fick indikera hur svårt och hur mentalt ansträngande det var att besvara varje enkät, dels genom att som i (Fink & Neubauer, 2001) ta tid på hur lång tid besvarandet tog.

Dessutom mättes hur säkra respondenterna kändes sig på sina svar genom att avsluta varje enkät med en fråga om visshet rörande konsekvensbedömningarna

och en likadan fråga för sannolikhetsbedömningarna. Detta är kopplat till hypoteserna 1b och 2b som visar på en uppfattning inför studien om att mer stöd i form av den strukturerade metoden borde ge ökad visshet.

Skalan för visshetsfrågorna gick från extremt osäker (1) till extremt säker (7). Frågorna och skalan inspirerades av forskning som rör den diskrepans som har funnits mellan enkätsvar (intention) och faktiskt beteende när det gäller benägenheten att spendera pengar och därmed att värdera nyttan av det som köps (Champ, Moore, & Bishop, 2009).

3.1.3 Respondenterna

Alla fjorton respondenter besvarade båda enkäterna. Respondenterna valdes slumpmässigt ut bland de arbetande på tre olika forskningsenheter på FOI. För att undersöka om de specifika respondenterna hade några särskilda egenskaper som skulle kunna påverka resultatet mättes deras kognitiva stil och expertis. Åtta frågor ställdes i respondenternas första enkät angående kognitiv stil där ett högt resultat i medel indikerade en mer logisk och faktabaserad stil medan ett lågt resultat motsvarade en mer intuitiv och känslobaserad stil. I första enkäten ställdes dessutom tre frågor om respondentens eventuella erfarenhet av riskbedömningar och expertis på informationssäkerhetsområdet.

3.2 Resultat

I detta avsnitt beskrivs de resultat som experimentet gav. För att öka läsbarheten har uppgifter om vilka statistiska test som har använts och deras resultat till stor del lagts i fotnoter. Informationen om vilka statistiska test som har använts har inkluderats för att öka spårbarheten och göra det möjligt att reproducera experimentet.

3.2.1 Konsekvensbedömningarna

Konsekvensbedömningarna var någorlunda jämnt fördelade över nivåerna 2–5, medan den lägsta nivån (1) knappt användes alls. Detta är dock naturligt med tanke på att hot med triviala konsekvenser sorterats bort redan vid enkätframtagningen.

Varje respondents svar på enkäten som använde strukturerade tabeller var ganska samstämmigt[3] med samma respondents svar på enkäten som använde naturligt språk. Avsaknaden av full samstämmighet mellan de två enkäterna som varje respondent fyllde i indikerar dock att de två beskrivningssätten till viss del ger olika resultat. En viss del av den bristande samstämmigheten kan även höra

[3] Pearsons korrelationskoefficient ligger i intervallet 0,355–0,882 med ett medel på 0,610

samman med att respondenternas bedömningar helt enkelt varierar över tiden, men det förklarar troligen bara en mindre del (Sommestad et al., 2016).

Samstämmigheten för konsekvensbedömningarna var för respondenterna sammantaget högre för naturligt språk än för tabeller, dock inte statistiskt signifikant högre[4]. Om alla fjorton respondenters bedömningar läggs samman fås en hög samstämmighet för de 23 hoten[4], men väljs istället en slumpmässig respondent ut blir samstämmigheten låg[5] även om det beaktas att respondenterna börjar sina inre skalor på olika ställen[6]. Exempelvis kanske en respondent på grund av generell enkätifyllnadsstil har regelbundet förskjutna bedömningar jämfört med andra respondenter. Denna typ av förskjutning verkar vara låg, men ändå en betydande faktor[7].

3.2.2 Sannolikhetsbedömningarna

Avseende sannolikhetsbedömningar beräknades samstämmigheten baserat på de procentsatser som ligger till grund för sannolikhetsskalan[8]. Hela 79 % av sannolikhetsbedömningarna var i nedre delen av skalan (1:or eller 2:or), trots att skalan var betydligt mer finmaskig där 1:or och 2:or tillsammans var avsedda att täcka in endast sannolikhetsintervallet 0–5 %. Å andra sidan var hoten avsedda att vara detaljerade nog för att få mer samstämmiga tolkningar, vilket samtidigt gör att sannolikheten för ett specifikt hot minskar avsevärt mot mer generella beskrivningar, som exempelvis de som användes i det experiment som genomfördes 2015 (Hallberg et al., 2015), där inte denna koncentration av bedömningsresultat med låga sannolikheter syntes.

Jämförs samma respondent mellan de två olika beskrivningssätten fås en låg intern samstämmighet[9]. Beskrivningssättet med naturligt språk gav generellt sett högre samstämmighet än tabeller, men detta är inte statistiskt signifikant[10] och samstämmigheten är klart lägre än för konsekvensbedömningarna[11].

[4] Intra-class correlation (ICC) consistency agreement average measures 95% CI är 0,877–0,965 för naturligt språk jämfört med 0,782–0,938 för tabeller.
[5] ICC absolute agreement single measures 95% CI är 0,215–0,548 för naturligt språk och 0,149–0,440 för tabeller.
[6] ICC consistency agreement single measures 95% CI är 0,338–0,663 för naturligt språk och 0,204–0,518 för tabeller.
[7] ICC absolute measures ligger ungefär 0,1 lägre än consistency measures. Tar man även med en multiplikativ effekt som i Pearson-korrelationer fås medel på 0,509 för naturligt språk respektive 0,365 för tabeller.
[8] De värden som användes utgjordes av medelvärdet för respektive intervallen, exempelvis motsvaras 2 på sannolikhetsskalan av medelvärdet (0,01 + 0,05) / 2 = 0,03.
[9] Pearson-korrelationerna ligger mellan -0,350 och 0,853 med ett medel på 0,431.
[10] Konfidensintervallen överlappar, t.ex. fås efter konvertering till en monoton skala ICC consistency agreement average measures 95% CI 0,459–0,846 för naturligt språk och 0,365–0,819 för tabeller.
[11] Pearson-korrelationerna är i medel 0,204 för naturligt språk och 0,171 för tabeller.

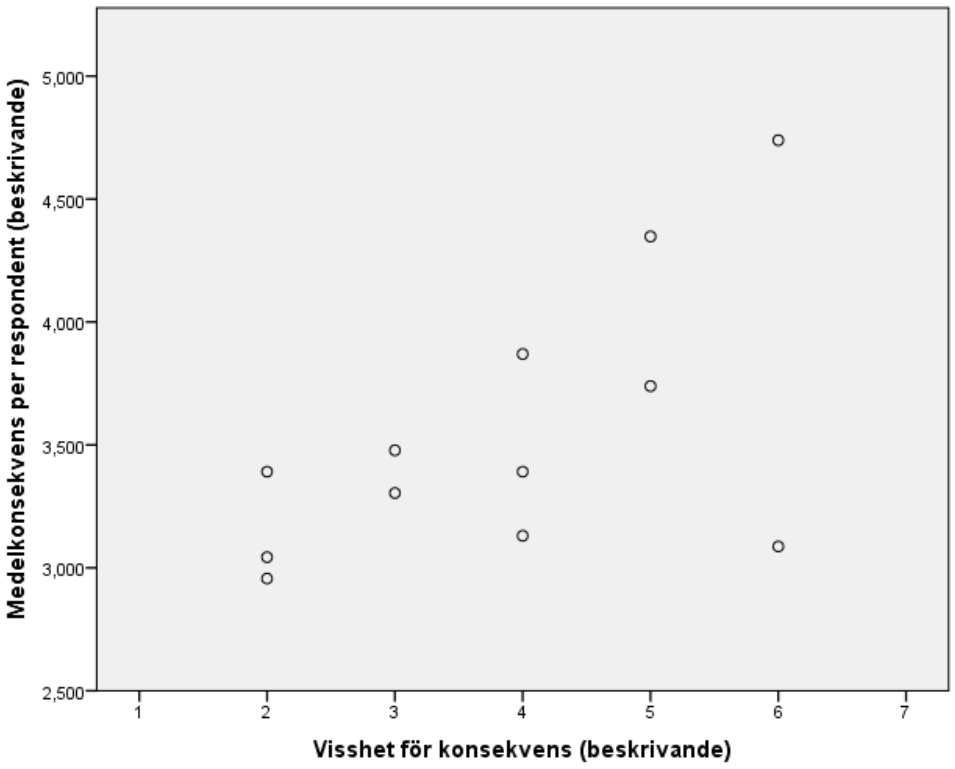
3.2.3 Kognitiv belastning

Hypoteserna 3a, 3b och 3c innebar att de tre typerna av kognitiv belastning (tid, svårighet, och mental ansträngning) var för sig skulle vara lägre, det vill säga mindre belastande, för enkäten med tabeller jämfört med enkäten med naturligt språk. Detta kunde dock inte ses i analysen av svaren[12]. Därmed föll hypoteserna 3a, 3b och 3c.

3.2.4 Visshet

Respondenterna som fyllde i att de kände sig ganska osäkra på sina konsekvensbedömningar för enkäten med naturligt språk bedömde generellt konsekvenserna lägre än övriga[13], och närmare mitten. Dessutom fanns en större spridning bland bedömd medelkonsekvens hos de mer säkra respondenterna[14] (Figur 5). Detta kan tolkas som att den som känner sig osäker håller sig runt mitten på skalan, medan de med mer självförtroende använder hela skalan. Inför studien fanns ingen uttalad uppfattning (hypotes) om vad mer försiktigt skulle innebära varför denna del rörande visshet är explorativ.

[12] Det fanns inga statistiskt signifikanta samband, inte ens vid $p < 0{,}100$.
[13] Säkrare halvan hade medelkonsekvensvärden som var 0,008–1,282 (95% CI) högre i snitt, enligt oberoende t-test tolerant för olika varians (s.k. Welch-test)
[14] $p = 0{,}078$ med Levenes test av lika varians (dvs. homoskedasticitet)



Figur 5: Relationen mellan varje respondents medelkonsekvens och visshet i enkäten med naturligt språk. Diagrammet innehåller 12 resultat eftersom två fick plockas bort på grund av ofullständiga svar gällande visshet.

Generellt såg Champ et al. (2009) att respondenter som svarar att de är osäkra på sina bedömningar i en enkät senare var mer benägna, än de som var säkra på sina bedömningar, att spendera mindre pengar än vad bedömningen hade förutsett. Det specifika fallet hotbedömningar är dock inte utrett. Hotbedömningar är normalt avsedda att leda fram till en riskbedömning som i sin tur används för att avgöra om skydd behöver införas. Det som köps är därmed skydd och att spendera mindre pengar än förutsett innebär alltså att skyddets effektivitet alternativt risken undervärderas jämfört med det förutsedda. Bortses från skyddseffektivitet – som har att göra med vilka alternativa skydd som finns att tillgå – innebär att spendera mindre pengar just att risken undervärderas vilket innebär att ett hots potentiella konsekvens, sannolikhet eller både och undervärderas. Detta innebär att en respondent som sätter lägre visshet för bedömningen av konsekvensen troligen även sätter lägre siffra för konsekvensen än en respondent med högre visshet gör. Det är också möjligt att en respondents ovisshet tar sig uttryck i att respondenten försöker bedöma konsekvens mer

försiktigt som en snittkonsekvens. Givet en för ändamålet anpassad skala bör många av konsekvensbedömningarna hamna nära mitten av skalan, så för osäkra respondenter är det statistiskt vettigt att lägga sig i mitten. I detta experiment innebär det att ovisshet bör ge en tendens att svara runt medel av skalan (3) justerat för de specifika hoten i enkäten (som gav medel på ca 3,5). Figur 5 visar just att lägre visshet leder till både lägre konsekvensmedel samt konsekvensmedel närmare snittet.

Det fanns inga andra kopplingar mellan konsekvens- eller sannolikhetsbedömningarna och visshet. Varken för konsekvens- eller sannolikhetsbedömningarna, fanns det heller någon koppling mellan visshet och huruvida hoten beskrivs i strukturerad tabell eller med naturligt språk[15], vilket innebär att hypotes 1b och 2b inte fann stöd.

3.2.5 Kognitiv stil

Frågorna om kognitiv stil (skala 1–5 där 1 är mest intuitiv och 5 är mest logisk) betraktas utifrån varje respondents snitt för dessa frågor. Respondenternas snitt låg nära varandra (i intervallet 2,25–4,25) med ett ganska högt medel (3,58).

Respondenternas kognitiva stil hade ingen statistiskt signifikant korrelation med samstämmigheten bland respondenterna, varken avseende hotbeskrivningar med tabeller eller naturligt språk. Det fanns dock en indikation att de mer logiska respondenterna hade ganska mycket lägre samstämmighet med respondenterna som helhet[16] avseende sannolikhetsbedömningarna i enkäten med tabeller. Det är möjligt att respondenter med intuitiv kognitiv stil inte på allvar försökte bedöma sannolikheten utan satte samma siffror på allt. Sannolikhetsbedömningarna var generellt väldigt låga vilket ger en indikation på att skalan inte var lämplig för hoten ifråga.

3.2.6 Expertis

Medelvärdena för de enskilda respondenternas svar på frågorna om expertis var jämfört med kognitiv stil mer spridda över skalan med ett medel för alla respondenter på 2,5 av 5. De flesta respondenterna hade låga medelvärden kring 1–2 medan tre respondenter hade höga medelvärden på 4,67; 4,67; respektive 5.

Respondenternas expertis hade betydelse för samstämmigheten för konsekvensbedömningarna i enkäten med tabeller. Ökad expertis innebar högre samstämmighet[17]. Experterna verkar inbördes vara än mer samstämmiga med

[15] p = 0,821 för t-testet för de två beskrivningssättens medelvärden för konsekvens respektive p = 0,863 för sannolikhet.
[16] p = 0,085; 95% CI -0,511–0,038 för beta i en linjär regressionsmodell.
[17] P < 0,05; 95% CI 0,024–0,157 för beta i en linjär regressionsmodell.

varandra när tabeller används än när naturligt språk används, men det kan inte statistiskt säkerställas baserat på insamlade data[18].

Omvänt verkar icke-experter mer benägna att vara inbördes samstämmiga när naturligt språk används än när tabeller används, men inte heller den skillnaden är statistiskt signifikant[19]. Detta innebär att det kan vara bra att anpassa metoden efter den utförandes expertis, även om mer forskning krävs. Saknas tillgång till experter är det möjligen bättre med hot beskrivna i naturligt språk. En intressant aspekt är också att experterna var mer benägna att sätta maxvärden på konsekvensen när tabeller användes än när naturligt språk användes, vilket också gav högre medelvärden (4,07 jämfört med 3,55).

3.2.7 Ordningseffekter

Det andra enkättillfället tog lite kortare tid[20], troligen på grund av att respondenterna redan hade läst scenariot vid första tillfället. För övrigt hittades inga skillnader mellan första och andra enkättillfället med avseende på kognitiv belastning eller visshet.

Hälften av respondenterna började med enkäten som hade tabeller, andra hälften med enkäten som hade naturligt språk. De senare rapporterade att de var säkrare på sina konsekvensbedömningar, både på första och andra enkäten[21]. Kanske detta är förknippat med att de flesta respondenter (icke-experterna) verkar föredra beskrivande form och därmed känner sig säkrare om de får använda den, en säkerhet som lever kvar även för nästa enkät. Detta kan dock inte sägas med bestämdhet eftersom det, som noterats i avsnittet om visshet, inte fanns någon statistiskt signifikant koppling mellan visshet och huruvida hoten beskrivs i strukturerad tabell eller med naturligt språk.

3.3 Resultatdiskussion

I detta avsnitt diskuteras resultaten från experimentet och återkoppling görs till de hypoteser som formulerades innan experimentet genomfördes.

[18] ICC consistency agreement single measures 95% CI 0,420–0,811 för tabeller jämfört med 0,282–0,740 för naturligt språk.
[19] ICC consistency agreement single measures 95% CI 0,141–0,449 för tabeller jämfört med 0,355–0,686 för naturligt språk.
[20] p = 0,013; 95% CI 0,998–6,859 minuter kortare i snitt, enligt beroende t-test.
[21] p = 0,01; 95% CI 0,440–2,760 steg högre (av maximala 7), enligt oberoende t-test tolerant för olika varians (s.k. Welch-test), efter minimal bortfallsjustering.

3.3.1 Konsekvens och sannolikhet

Lägsta nivån (1 av 5) för konsekvensbedömningarna användes sällan, vilket troligen beror på att triviala hot naturligt sorterats bort vid framtagningen av enkäten. Den näst högsta nivån (4 av 5) användes något mer flitigt än övriga nivåer, vilket kan bero på att det bland en del respondenter finns ett motstånd mot att sätta högsta möjliga konsekvens. Samstämmigheten mellan respondenterna var för konsekvensbedömningarna generellt sett ganska låg. En mindre del av variationen kan förklaras med att en del respondenter verkar börja sin inre skala lite högre eller lägre än nivå 1. Hypotesen att användandet av strukturerade tabeller skulle ge högre samstämmighet för bedömningarna av konsekvens än när naturligt språk används (hypotes 1a) fann inget generellt stöd.

De allra flesta sannolikhetsbedömningarna sattes till någon av de två lägsta nivåerna (nivå 1 eller 2), trots att de nivåerna bara motsvarade sannolikhetsintervallet 0–5 %. Med andra ord var hoten väldigt specifika och en lärdom kan vara att, åtminstone för sannolikhetsbedömningar, ha mindre specifika hot alternativt en än mer detaljerad skala för de låga värdena. Samstämmigheten för sannolikhetsbedömningarna var generellt mycket låg mellan respondenterna. Hypotesen att användandet av strukturerade tabeller skulle ge högre samstämmighet för bedömningarna av sannolikhet än vid användandet av naturligt språk (hypotes 2a) fann inget generellt stöd.

3.3.2 Kognitiv belastning och visshet

Den kognitiva belastningen, med avseende på mental ansträngning, svårighet samt tid, skilde inte mellan de två beskrivningssätten. Hypoteserna att användandet av strukturerade tabeller skulle ge lägre kognitiv belastning än vid användandet av naturligt språk (hypotes 3a, 3b och 3c) fann inget stöd. Inte heller hypoteserna att användandet av strukturerade tabeller skulle ge högre visshet i respondenternas svar än vid användandet av naturligt språk (hypotes 1b och 2b) fann något stöd.

3.3.3 Kognitiv stil och expertis

Respondenternas självuppskattade expertis inom informationssäkerhetsområdet och rörande riskbedömningar var generellt låg, undantaget tre respondenter som uppgav sig vara mycket väl bevandrade på området. Dessa tre respondenter hade betydligt högre samstämmighet än övriga när det gällde konsekvens-bedömningarna vid användandet av tabeller. Det fanns alltså en indikation att experterna var mer inbördes samstämmiga vid användandet av tabeller än vid användandet av naturligt språk, medan det fanns en indikation för det motsatta för icke-experterna. Alltså kan det vara av värde att anpassa beskrivningssättet efter nivån av expertis hos de personer som ska genomföra bedömningarna, även

om det verkar vara lämpligast att eftersträva så hög nivå av expertis som möjligt. En intressant aspekt är dessutom att det finns en indikation på att användning av naturligt språk gör att respondenterna uttrycker en högre grad av visshet.

4 Jämförelsebaserade bedömningar

Inom Försvarsmakten används, i enlighet med H Säk Grunder (Försvarsmakten, 2013), absoluta bedömningar där varje hot bedöms individuellt, ofta med hjälp av de föreslagna skalorna för sannolikhet och konsekvens. I H Säk Grunder nämns dock att bedömning av bland annat sannolikhet och konsekvens ska ses som relativa bedömningar. Det kan då ifrågasättas om inte bedömningarna istället borde göras enligt en metod som tydligare ger stöd för genomförandet av just relativa bedömningar.

Under 2015 genomfördes en enkätundersökning för att undersöka hur stor samstämmigheten är mellan olika bedömare när sannolikhet och konsekvens ska bedömas för en uppsättning IT-säkerhetsrelaterade hot (Hallberg et al., 2015). I undersökningen testades både absoluta bedömningar enligt metoden i H Säk Grunder (Försvarsmakten, 2013) och jämförelsebaserade bedömningar baserat på Analytic Hierarchy Process (AHP) (Saaty, 1990). Resultaten från enkätundersökningen 2015 visade inte på att de jämförelsebaserade bedömningarna gav bättre samstämmighet, snarare tvärt om. Det fanns dock indikationer på att annat val av bedömningsskala, strategi för vilka jämförelser som ska göras samt beräkningsmodell för att få fram absoluta värden skulle kunna ge mer samstämmiga bedömningar.

AHP används för att prioritera alternativa lösningar genom att jämföra dem baserat på en uppsättning med kriterier (Forman & Selly, 2001). Exempelvis kan pris och prestanda vara två sådana kriterier och olika bilmodeller kan utgöra alternativen. Innan alternativen jämförs ska de kriterier som ligger till grund för jämförelserna prioriteras. På så sätt kan exempelvis priset ges en större betydelse än prestandan när prioritetsvärdena för de alternativa lösningarna beräknas. I AHP används således samma tillvägagångssätt för att jämföra kriterier såväl som alternativa lösningar. När detta tillvägagångssätt ska beskrivas används det mer generella begreppet *element* för att benämna både *kriterier* och *alternativa lösningar* i exemplet ovan.

I den ursprungliga versionen av AHP jämförs alla möjliga par med element direkt med varandra och resultaten förs in i en så kallad bedömningsmatris. Om N stycken element ska jämföras med varandra kommer bedömningsmatrisen att innehålla N gånger N stycken värden. I Figur 6 återfinns en bedömningsmatris för tio stycken element (E_1 till E_{10}). Det innebär att det finns tio gånger tio stycken värden som ska fyllas i.

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
E_1	?	?	?	?	?	?	?	?	?	?
E_2	?	?	?	?	?	?	?	?	?	?
E_3	?	?	?	?	?	?	?	?	?	?
E_4	?	?	?	?	?	?	?	?	?	?
E_5	?	?	?	?	?	?	?	?	?	?
E_6	?	?	?	?	?	?	?	?	?	?
E_7	?	?	?	?	?	?	?	?	?	?
E_8	?	?	?	?	?	?	?	?	?	?
E_9	?	?	?	?	?	?	?	?	?	?
E_{10}	?	?	?	?	?	?	?	?	?	?

Figur 6: Bedömningsmatris för tio element (E_1 till E_{10}) vars relativa vikter ska beräknas.

Eftersom alla element (E_1 till E_{10}) är likvärdiga när de jämförs med sig själva kommer diagonalen alltid att innehålla ettor. Dessutom är de enskilda elementens relativa vikter oberoende av i vilken ordning de jämförs, dvs. om jämförelsen av element E_3 med E_8 ger det relativa värdet v så ger E_8 jämfört med E_3 det relativa värdet $1/v$. Dessa egenskaper hos bedömningsmatrisen illustreras i Figur 7, där alla värden i diagonalen är satta till 1 samt att v och $1/v$ har förts in för de värden som gäller förhållandet mellan E_3 och E_8. Det förhållande som därmed finns mellan de värden som återfinns ovanför diagonalen (markerade med mörkgrått i figuren) och de värden som återfinns under diagonalen gör att det är tillräckligt att ange hälften av de värden som återfinns utanför (ovan eller under) diagonalen. Vid användning av AHP nyttjas detta faktum ofta så att värdena ovanför diagonalen anges, medan värdena under diagonalen beräknas utifrån de ovan diagonalen införda värdena.

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
E_1	1	?	?	?	?	?	?	?	?	?
E_2	?	1	?	?	?	?	?	?	?	?
E_3	?	?	1	?	?	?	?	*v*	?	?
E_4	?	?	?	1	?	?	?	?	?	?
E_5	?	?	?	?	1	?	?	?	?	?
E_6	?	?	?	?	?	1	?	?	?	?
E_7	?	?	?	?	?	?	1	?	?	?
E_8	?	?	1/*v*	?	?	?	?	1	?	?
E_9	?	?	?	?	?	?	?	?	1	?
E_{10}	?	?	?	?	?	?	?	?	?	1

Figur 7: Bedömningsmatris där diagonalens värden har satts till ett, den del av matrisen som återfinns ovanför diagonalen har markerats samt *v* och 1/*v* har förts in som de värden som gäller förhållandet mellan E_3 och E_8.

Sammanfattningsvis behöver långt färre än alla N^2 värden som ska återfinnas i en fullständig bedömningsmatris anges; eftersom diagonalen (markerad med blått i Figur 7) endast innehåller ettor återstår N^2-N värden att ange; på grund av förhållandet mellan värden ovanför och under diagonalen behöver slutligen (N^2-N)/2 värden anges, vilket motsvaras av det rosa området i Figur 7. Även om det är betydligt färre än N^2 värden som i slutändan behöver anges ökar det totala antalet jämförelser snabbt. Exempelvis ger en uppsättning med tio element (som i Figur 6 och Figur 7) 45 par att jämföra, medan en uppsättning med 100 element ger 4950 par att jämföra. (Hallberg et al., 2015)

Tabell 1 innehåller viktiga begrepp från AHP som används i resonemangen i fortsättningen av detta kapitel och i följande kapitel.

Tabell 1: Begrepp som används i samband med AHP.

Begrepp	**Beskrivning**
Bedömningsmatris	Innehåller resultaten av de jämförelser som har gjorts
Prioritetsvärde	De värden som beräknas utifrån bedömningsmatrisen
Rangordning	Sortering av element utifrån prioritetsvärde
Skala	Bedömningsalternativ med tillhörande vikter
Vikt	Beskriver förhållandet mellan två element

Under 2016 har studierna i projektet *Bedömning och analys av IT-system* fokuserats på att fortsatt utvärdera hur jämförelsebaserade bedömningar skulle kunna användas vid bedömning av sannolikhet och konsekvens för risker relaterade till Försvarsmaktens IT-system. I avsnitten som följer beskrivs och analyseras olika skalor som kan användas vid jämförelser av hot. Dessutom beskrivs en strategi för att välja vilka par av hot som ska jämföras samt hur förslag på absoluta värden tas fram för de hot som har jämförts. Kapitlet avslutas med rekommendationer avseende fortsatt utvärdering av jämförelsebaserade bedömningar.

4.1 Val av skala för jämförelser

Det finns tre väsentliga frågor att besvara gällande den skala som ska användas vid jämförelse av sannolikhet och konsekvens kopplad till olika hot.

1. Hur många steg ska skalan ha?
2. Vilka vikter ska associeras med de olika stegen?
3. Ska skalorna vara desamma för konsekvens respektive sannolikhet?

Vid den praktiska användningen av skalorna är det också väsentligt att avgöra ifall skalorna ska beskrivas med de vikter som associeras med de olika stegen alternativt värdeord.

4.1.1 Antal steg på skalan

En skala med få steg har fördelen att det blir enklare att genomföra bedömningarna då jämförelsen inte behöver vara lika specifik. Ett exempel på en skala med få steg kan vara en som vid jämförelse av alternativ A och B enbart har de tre alternativen *A är större*, *Lika stora* och *B är större*. Bedömaren behöver med denna 3-gradiga skala inte avgöra exakt hur mycket större alternativ A eller B är, utan det räcker att avgöra om något av alternativen är större än det andra eller om de två är ungefär lika stora.

En skala med många steg har å andra sidan fördelen att den möjliggör en högre precision i jämförelserna. Möjligheten till högre precisionen beror på att det går att uttrycka mer specifikt hur stor skillnaden mellan två alternativ är. Istället för att bara kunna ange att A är större än B så kan skalan exempelvis innehålla steg såsom *A är lite större*, *A är större* eller *A är betydligt större*.

För att prova hur jämförelser baserat på olika skalor skiljer sig åt genomfördes ett mindre test där 13 hot (Tabell 2) bedömdes genom att alla par med hot (78 stycken) jämfördes med varandra. Testet utfördes med hjälp av skalor med 3 respektive 9 steg. De resulterande bedömningsmatriserna återfinns i Bilaga 2.

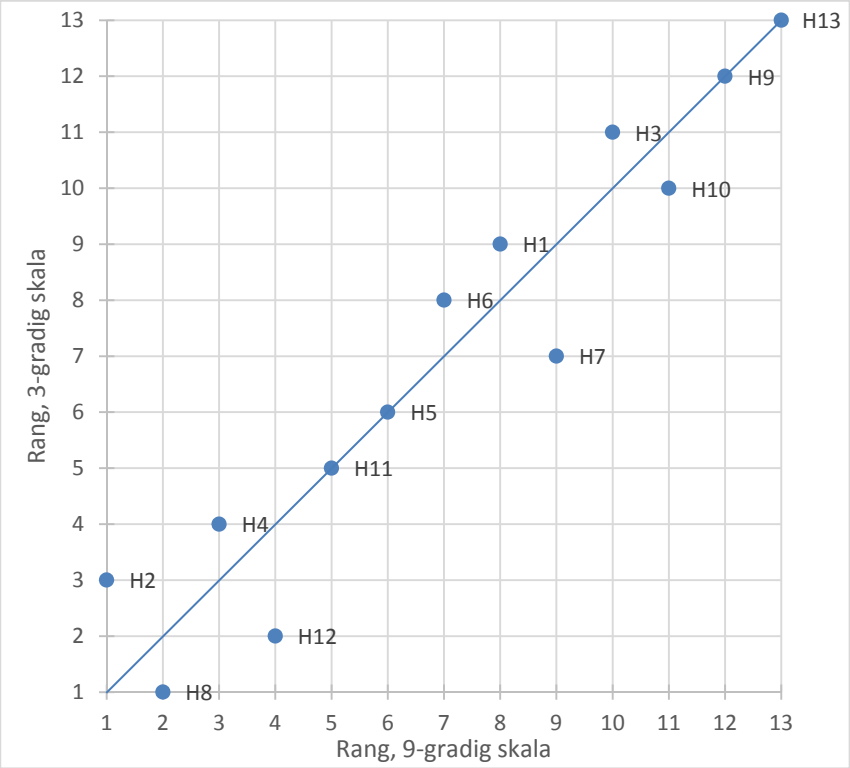
Tabell 2: De 13 hot som jämfördes med varandra avseende sannolikhet och konsekvens. Hotbeskrivningarna inspirerades av två publikationer från Mcaffee (Beek et al., 2016) och AT&T[22] som publicerades under 2015 och där de beskriver vilka typer av IT-säkerhetshot de trodde skulle öka under 2016.

Id	Hot
H1	Infektion av s.k. ransomware leder till att data krypteras och krav på lösensumma ställs för att krypterade data ska återställas.
H2	Sårbarheter i hårdvara nyttjas av extern aktör för att komma åt information.
H3	Sårbarheter i mjukvara nyttjas av extern aktör för att komma åt information.
H4	Betalningssystem attackeras av kriminell organisation med syfte att komma över stora summor med pengar.
H5	Molntjänster attackeras för att störa verksamheten och komma över känslig information.
H6	Inloggningsuppgifter till kontorssystem säljs via nätbaserade kriminella nätverk.
H7	Informationsstöld genomförs av kriminella för att kunna idka utpressning alternativt sälja information till högstbjudande.
H8	Statsfinansierade hackers attackerar IT-system som led i underrättelseverksamhet.
H9	DDoS-attack som stör verksamheten genomförs av hacktivistgrupp (t.ex. Anonymous).
H10	Mobila enheter attackeras för att komma åt information.
H11	Attacker riktas och anpassas mot användare (spear fishing) av informationssystem för att lura dessa att ge ifrån sig information alternativt installera skadlig mjukvara.
H12	Industriella informations- och styrsystem, s.k. SCADA-system, i kritisk infrastruktur attackeras med syfte att störa samhället.
H13	Skadlig kod som cirkulerar på Internet infekterar dator i kontorsnät.

Spearmans rangkorrelation (ibland kallad Spearmans rho) var 0,95 respektive 0,85 för samstämmigheten för konsekvensbedömningarna respektive sannolikhetsbedömningarna. Bedömningarna av konsekvens var därmed mer samstämmiga än bedömningarna av sannolikhet. Samstämmigheten illustreras i

[22] https://networkingexchangeblog.att.com/enterprise-business/top-11-security-threats-of-2016/

sambandsdiagrammet i Figur 8 som illustrerar förhållandena mellan rangen som de olika hoten erhöll med 3-gradig respektive 9-gradig skala för jämförelse av konsekvens. Idealt skulle rangen inte bero på antalet steg i jämförelseskalan och därmed ligga längs diagonalen från (1, 1) till (13, 13). Så är inte fallet, men inget av hoten ligger längre än två steg ifrån linjen, därav det höga värdet på Spearmans rangkorrelation. Detta antyder att det inte har någon större betydelse för utfallet av jämförelserna huruvida det är den 3-gradiga eller den 9-gradiga skalan som används, förutsatt att det endast är ordningen mellan hoten som är relevant.



Figur 8: Sambandsdiagram med 13 hot (H1 till H13) införda baserat på den rang avseende konsekvens som erhölls med 3-gradig respektive 9-gradig skala för jämförelser.

Ett mått som brukar användas i samband med AHP är konsistensen hos de gjorda jämförelserna. Konsistensen kan användas för att identifiera möjliga logiska fel i de bedömningar som gjorts. Exempelvis skulle följande kombination av bedömningar kunna leda till en låg konsistens:

- B är större än A
- C är större än B
- A är större än C

Konsistensen blir låg i exemplet då den första och andra bedömningen gemensamt ger att C är större än A samtidigt som den tredje bedömningen ger att A är större än C.

Hur konsistensen påverkas av huruvida det är en 3- eller 9-gradig skala som används är också en parameter att beakta vid valet av antalet steg på skalan. Eftersom AHP baseras på att värden sätts på de relativa skillnaderna mellan de enheter som jämförs leder en skala med fler steg rimligen till förbättrade möjligheter att uppnå en hög konsistens mellan jämförelserna. Konsistensmått påverkas dock även av vilka vikter som används på skalorna varför även detta behöver beaktas.

4.1.2 Vikter på skala för jämförelser

Baserat på underlaget från testet prövades en uppsättning med skalor för jämförelser som har föreslagits i olika forskningspublikationer. Ishizaka och Labib (2011) presenterar en sammanställning med åtta skalor. Franek och Kresta (2014) presenterar en uppdaterad version av sammanställningen och beräkningar av underlag för att beräkna konsistensen hos jämförelser baserade på de olika skalorna. Baserat på de föreslagna skalorna genomfördes vidare utvärdering med de tre typer av skala som i litteraturen refereras till som linjär, geometrisk och logaritmisk. Det finns en skillnad gällande vad begrepp innebär avseende skalor för storheter och skalor för jämförelser av storheter. Skalor för jämförelser benämns i allmänhet efter den funktion som används för att bestämma vilka vikter de olika stegen på skalorna ska ha (Franek & Kresta, 2014; Ishizaka & Labib, 2011). En jämförelseskala med exponentiellt växande värden kallas för geometrisk (eftersom värdena på skalan utgör en geometrisk talföljd), medan en skala för ett sambandsdiagram med värden som växer exponentiellt kallas logaritmisk.

I Tabell 3 återfinns för var och en av skalorna den formel som används för att beräkna vikterna på respektive skala samt de värden (vikter) som används i denna utvärdering. De mått som används för att utvärdera skalorna är Spearmans rangkorrelation och konsistensmått för bedömningsmatriser. Spearmans rangkorrelation ger en uppfattning om hur stor skillnad det blir avseende vilken rang de olika hoten får med avseende på sannolikhet och konsekvens. Konsistensmått ger en uppfattning om i vilken grad skalan påverkar utfallet av jämförelserna. Ji och Jiang (2003) beskriver hur den använda skalan kan påverka konsistensen hos jämförelserna. Exempelvis begränsas den maximala skillnaden som kan uttryckas av skalans största respektive minsta värde, dvs. om

förhållandet mellan A och B är 5 och förhållandet mellan B och C är 3 leder det till att förhållandet mellan A och C är 15, men det går inte att uttrycka på en skala där det maximala värdet är 9 och leder till försämrad konsistens. Ytterligare en begränsning hos skalor är att de har ett begränsat antal värden. Om förhållandet mellan A och B är 5 och förhållandet mellan B och C är 1/3 leder det till att förhållandet mellan A och C är 5/3, men om det värdet inte finns med på skalan leder även det till försämrad konsistens.

Eftersom vi använder samma matris med jämförelser vid beräkningarna av prioritetsvärden baserade på de olika skalorna ger konsistensmåtten en indikation på i vilken omfattning den använda skalan bidrar till försämrad konsistens.

I denna studie användes det mått som är närmast förknippat med AHP-metoden (Saaty, 1990), vilket kallas CR (eng. Consistency Ratio). För att kunna sätta en gräns för vilken storlek ett CR-värde ska ha för att motsvarande bedömningsmatris ska anses var tillräckligt konsistent baseras CR inte endast på konsistensen för den aktuella bedömningsmatrisen utan även på den genomsnittliga konsistensen i slumpvis framtagna bedömningsmatriser (1).

$$CR = CI/RI \tag{1}$$

CI (eng. Consistency Index) är konsistensen för aktuell matris och RI (eng. Random Index) är genomsnittligt CI för slumpvis framtagna bedömningsmatriser.

CI för en bedömningsmatris beror såväl på matrisens storlek som den skala som används för jämförelserna. Detta medför att beräkningen av CR kräver att det finns RI framtaget för den storlek på matris och skala som används i det aktuella fallet. En begräsning med avseende på beräkning av konsistens är att det endast finns en begränsad uppsättning med RI-värden för olika skalor och storlekar på bedömningsmatriser redovisade i litteraturen. Alonso och Lamata (2006) har härlett en generell ekvation som kan användas för att räkna ut konsistens för en godtycklig storlek på bedömningsmatris dock endast för en typ av skala med ett specifikt antal steg på skalan. Franek och Kresta (2014) har dock tagit fram RI-värden för åtta skalor och matriser upp till och med storlek 15.

De värden som presenterades av Franek och Kresta (2014) kan inte användas i det test som beskrivs i detta kapitel då deras beräkningar genomfördes med 17-gradiga skalor. Testet som beskrivs i detta kapitel använde en 3-gradig skala och en 9-gradig skala samt bestod av 13 hot som jämfördes, vilket ger en matrisstorlek på 13. Avsaknaden av RI-värden för 3- och 9-gradiga skalor löstes genom att beräkna de specifika RI-värdena med hjälp av en algoritm som implementerades i Excel.

För att validera den implementerade algoritmen jämfördes resultaten med de resultat som presenterades av Franek och Kresta (2014). Genom att beräkna

genomsnittligt CI för 500 000 bedömningsmatriser innehållande slumpmässiga värden, erhölls de värden som återfinns i Tabell 4. Resultaten visar att metoderna producerar likvärdiga resultat, där den enda märkbara skillnaden uppstod för den geometriska skalan. Den relativa skillnaden är dock mindre än 0,02 %. En möjlig anledning till att det uppstod en skillnad för just den geometriska skalan är att Franek och Kresta (2014) redovisar sina resultat med 3 decimaler, vilket medför att det finns fler värdesiffror att jämföra för den geometriska skalan som har större absoluta värden.

Efter valideringen nyttjades metoden för att beräkna RI-värden för de 9-gradiga skalorna som återfinns i Tabell 3. Dessa värden blir något högre än motsvarande för 17-gradiga skalor.

Tabell 3: Tre skalor som kan användas vid jämförelser enligt AHP. Formlerna beskriver hur de vikter som är större än eller lika med 1 tas fram. Resulterande vikter som används i detta test listas i kolumnen Vikter. De vikter som är mindre än 1 ges av inversen av de vikter som är större än 1. RI (eng. random index) utgörs av det genomsnittliga värdet för konsistensindex för 500 000 slumpmässiga bedömningsmatriser av storlek 13. Spearmans rangkorrelation anger det genomsnittliga värdet vid jämförelse med de två andra skalorna för sannolikhet (S) respektive konsekvens (K). Konsistensen (CR) avser resultaten för bedömningsmatriserna för sannolikhet (S) respektive konsekvens (K).

Skala	**Formel, i = 1, 3, 5, 7, 9**	**Vikter**	**RI**	**Spearman**		**CR**	
				K	**S**	**K**	**S**
Linjär	$v_i = i$	1; 3; 5; 7; 9	1,655	0,992	0,955	0,081	0,099
Geometrisk	$v_i = 2^{i-1}$	1; 4; 16; 64; 256	31,97	0,984	0,901	0,0095	0,020
Logaritmisk	$v_i = \log_2(i + 1)$	1; 2; 2,58; 3; 3,32	0,4269	0,992	0,948	0,116	0,131

Tabell 4: Värden från litteraturen och den metod som användes för att beräkna nya RI-värden. Skalorna är 17-gradiga. De vikter som är mindre än 1 ges av inversen av de vikter som är större än 1. Värdena avser bedömningsmatriser med storleken 13. Såväl de tidigare resultaten, från (Franek & Kresta, 2014), som de nya beräkningarna baseras på medelvärden för 500 000 matriser.

Skala	**Formel, i = 1, 2, …, 9**	**Vikter**	**Tidigare beräkningar av RI**	**Nya beräkningar av RI**
Linjär	$v_i = i$	1; 2; 3; 4; 5; 6; 7; 8; 9	1,555	1,555
Geometrisk	$v_i = 2^{i-1}$	1; 2; 4; 8; 16; 32; 64; 128; 256	25,117	25,112
Logaritmisk	$v_i = \log_2(i + 1)$	1; 1,58; 2; 2,2; 2,58; 2,81; 3; 3,17; 3,32	0,410	0,410

De värden som har beräknats för Spearmans rangkorrelation visar på en hög samstämmighet avseende vilken rang de olika skalorna resulterar i för de jämförda hoten. Detta gäller speciellt med avseende på konsekvens där korrelationen är nära 1. Den geometriska skalan utmärker sig dock något genom lägre korrelation med de andra två. Då korrelation fortfarande är mycket stark ger den i huvudsak samma resultat som de två andra skalorna.

Vad gäller konsistensen hos bedömningsmatriserna är dock skillnaderna större. Den geometriska skalan leder till betydligt bättre konsistens. Detta kan förklaras med att den ger större utrymme för att uttrycka stora skillnader mellan hoten med avseende på sannolikhet och konsekvens. Det är dock vanskligt att dra alltför stora slutsatser baserat på det begränsade underlaget. En annan begränsande faktor är att konsistensen beräknas med olika parametrar (RI).

För att kunna jämföra konsistensen hos bedömningsmatriserna för den 3-gradiga och den 9-gradiga skalan är det nödvändigt att beräkna konsistensen hos ett stort antal bedömningsmatriser med en 3-gradiga skala. En parameter som måste väljas är storleken på de två vikter som inte är lika med ett. I Tabell 5 återfinns resultaten för fem olika val av den vikt (v) som är större än 1, vilket också ger den minsta vikten som blir 1/v. Genom att beräkna konsistensindex för 500 000 slumpmässiga bedömningsmatriser baserade på dessa 3-gradiga skalor erhålls RI-värden för dessa. Baserat på de erhållna RI-värdena kan CR-värden beräknas för de två bedömningsmatriserna för konsekvens respektive sannolikhet som togs fram under testet. RI-värdena ökar med variansen hos skalan. Denna ökning

motsvaras dock i hög grad av ökande CI-värden för de framtagna bedömningsmatriserna, vilket leder till att CR-värdena är relativt likvärdiga.

Tabell 5: Konsistensvärden för den 3-gradiga skalan. Vikt anger storleken på den vikt som är större än 1. RI anger beräknade värden baserade på 500 000 matriser. CR anger det resulterande konsistensvärdet för den bedömningsmatris för 13 hot som togs fram under testet.

Vikt	RI	CR konsekvens	CR sannolikhet
2	0,1425	0,236	0,271
3	0,3843	0,225	0,258
5	0,9359	0,206	0,238
7	1,5164	0,193	0,224
9	2,1070	0,254	0,295

CR-värdena i Tabell 5 är betydligt högre än de värden som återfinns i Tabell 3 och ger vid handen att 3-gradiga skalor inte ger tillräckligt bra konsistens för att användas vid bedömningarna. Konsistensen påverkas dock både av de genomförda bedömningarna och den använda skalan.

4.1.3 Rekommendationer gällande val av skala

Det finns ett antal faktorer som behöver beaktas vid val av vilken skala som ska användas vid jämförelser av sannolikhet respektive konsekvens kopplad till hot mot informationssäkerhet. Den ansträngning och tid som krävs för genomförandet av jämförelserna talar för en skala med få steg, medan möjligheten att konsistent uttrycka förhållandet mellan konsekvenser respektive sannolikheter kopplade till olika hot talar för en skala med fler steg.

I detta kapitel presenteras ett enkelt test av olika skalor. De skalor som testats utgörs av två grupper med tre respektive nio steg. Skalorna med nio steg leder konsekvent till bättre intern konsistens för de två bedömningsmatriser som skapades genom ett test baserat på jämförelser av sannolikhet respektive konsekvens för 13 hot. Vilken mängd steg på skalan som är bäst beror sannolikt på med vilken noggrannhet jämförelserna behöver genomföras.

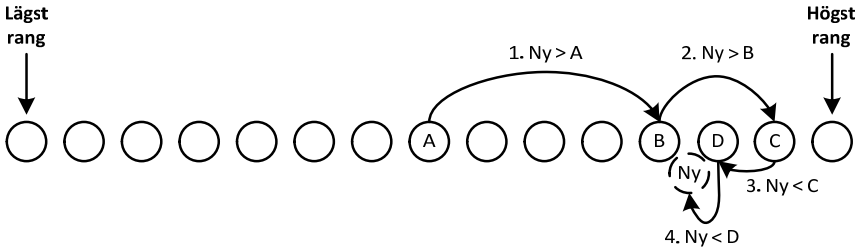
För 9-gradiga skalor verkar det finnas en klar fördel, med avseende på konsistens mellan jämförelserna, att nyttja en geometrisk skala, vilket framgår av CR-värdena i Tabell 3. Detta kan bero på att en sådan skala ger större möjligheter att uttrycka såväl relativt små som stora skillnader avseende sannolikhet och konsekvens.

4.2 Val av par med hot som ska jämföras

Det finns många tänkbara sätt att välja ut vilka par av hot som ska jämföras. Under 2015 provades en strategi som gick ut på att använda det minsta möjliga antalet jämförelser som behövs för att kunna beräkna prioritetsvärden för en uppsättning med hot (Hallberg et al., 2015). Genom att nyttja indirekta jämförelser, exempelvis att förhållandet mellan B och C kan avgöras genom att beakta förhållandet mellan A och B respektive A och C, kan hoten rangordnas med enbart n-1 direkta jämförelser för n stycken hot. Ett rättvisande resultat förutsätter dock att de jämförelser som genomförs beskriver det exakta förhållandet mellan de jämförda hoten. När precisionen i jämförelserna är begränsad medför dock indirekta jämförelser problem genom att hoten inte nödvändigtvis får samma rang som de skulle ha fått genom direkta jämförelser. Om det finns en korrekt rangordning är det inte säkert att direkta jämförelser mellan alla hot leder till just denna rangordning när jämförelsernas precision är begränsad och det dessutom kan förekomma jämförelser som inte är konsistenta med varandra. En av tankarna med AHP är dock att redundansen mellan jämförelserna ska minska problemet med jämförelsernas begränsade precision och förekomsten av icke-konsistenta jämförelser.

En slutsats av det under 2015 genomförda experimentet är att det behövs strategier som leder till fler än det minimala antalet jämförelser. En möjlighet är att jämföra alla par med hot som kan formas utifrån den aktuella uppsättningen. En annan möjlighet är att nyttja en strategi som ger tillräcklig redundans.

I detta avsnitt beskrivs en strategi för att rangordna hot och beräkna deras prioritetsvärden baserat på en delmängd av alla möjliga jämförelser samt ett test av denna strategi. Målet med strategin är att med så få direkta jämförelser som möjligt få samma rangordning av hot som skulle ges av att jämföra alla par med hot, dvs. minska antalet jämförelser som behöver genomföras utan att det påverkar rangordningen av hoten. Strategin har likheter med sorteringsalgoritmen *insertion sort* (Sedgewick & Wayne, 2011). En sådan strategi kan fungera enligt exemplet i Figur 9 som illustrerar hur hot placeras in i en ordnad lista. Exemplet illustrerar också att en stor del av jämförelserna görs med hot som ligger nära det nya hotet i uppsättningen med rangordnade hot.



Figur 9: Exempel som illustrerar hur hot placeras in i en ordnad lista. Hot Ny som ska placeras in i listan jämförs med det mellersta hotet i listan, hot A. eftersom hot Ny har högre rang än hot A jämförs det därefter med hot B, som ligger mitt i listan med hot som har högre rang än hot A. Återigen har hot Ny högre rang och jämförs därför med hot C som ligger mitt i listan med hot som har högre rang än hot B. Då hot Ny har lägre rang än hot C jämförs det sedan med hot D som är det enda (och därmed mellersta) hot som har lägre rang än hot C (i listan med hot som har högre rang än hot B). Då hot Ny har lägre rang än hot D placeras det slutligen till vänster om hot D i listan.

Rangordningen av hot utgår ifrån en uppsättning med oordnade hot (U_O), vilka ett i taget placeras in i en uppsättning med rangordnade hot (U_R). När rangordningen inleds återfinns alla hot i U_O och när rangordningen är klar återfinns alla hot i U_R. För att avgöra var i U_R som ett hot ska placeras in används en tillfällig uppsättning U_J som återspeglar de hot som fortfarande kan vara relevanta att jämföra med. Strategin genomförs i följande steg.

1. Välj slumpmässigt ett hot från uppsättningen med oordnade hot (U_O) och lägg in det i uppsättningen med rangordnade hot (U_R).
2. Om det inte finns några hot kvar i U_O är rangordningen klar. Annars välj slumpmässigt ett nytt hot (H_{NY}) från U_O.
3. Sätt U_J till U_R.
4. Jämför H_{NY} med det hot (H_A) som ligger mitt i U_J.
    a. Om H_{NY} har samma rang som H_A, placera H_{NY} direkt efter H_A i U_R och återgå till steg 2.
    b. Om H_{NY} har lägre rang än H_A, ta bort H_A och de hot som har högre rang än H_A från U_J.
        i. Om nya U_J är tom, placera H_{NY} direkt efter H_A i U_R och återgå till steg 2.
        ii. Annars återgå till steg 4.
    c. Om H_{NY} har högre rang än H_A, ta bort H_A och de hot som har lägre rang än H_A från U_J.

i. Om nya U_J är tom, placera H_{NY} direkt innan H_A i U_R och återgå till steg 2.

ii. Annars återgå till steg 4.

Strategin medför att i de fall då bedömningen är att det nya hotet har en lägre eller högre rang än det hot som det jämförs med (4b respektive 4c) så fortsätter strategin med en uppsättning som innehåller maximalt hälften så många hot. Det innebär att det maximala antalet jämförelser som krävs för att placera in ett hot i den rangordnade uppsättningen med hot blir in funktion av logaritmen med bas 2 av antalet hot som redan finns in den rangordnade uppsättningen och växer relativt långsamt eftersom en dubblering av antalet hot endast kräver ytterligare en jämförelse. I praktiken krävs färre jämförelser eftersom de fall där de jämförda hoten bedöms vara lika leder till att det nya hotet placeras in i den rangordnade uppsättningen med hot utan ytterligare jämförelser. Det faktum att de fall när hoten bedöms vara lika leder till att det nya hotet placeras efter det hot som redan var inplacerat i uppsättningen med rangordnade hot leder till att den slutliga rangordningen av hot som bedöms vara lika beror på slumpen.

När hoten har rangordnats med den föreslagna strategin finns det en viss redundans mellan de gjorda jämförelserna. Med hjälp av Harkers metod (Harker, 1987) kan den extra kunskap som redundansen ger infogas i beräkningar av prioritetsvärden för hoten. Baserat på de därigenom erhållna prioritetsvärdena kan en modifierad rangordning av hoten tas fram.

Ett test genomfördes för att erhålla tentativa svar på följande frågor.

1. Hur stor blir skillnaden mellan rangordningar av hot baserade på en fullständig bedömningsmatris respektive den föreslagna strategin för sannolikhet respektive konsekvens?
2. Blir skillnaden mindre om rangordningarna modifieras baserat på prioritetsvärden beräknade med Harkers metod för sannolikhet respektive konsekvens?

Testet baserades på samma uppsättning med jämförelser för alla möjliga kombinationer av hot som användes för testet av 3- och 9-gradiga skalor i föregående avsnitt. Eftersom strategin innebär att hoten rangordnas i slumpmässig ordning är den inte deterministisk utan kan resultera i olika rangordningar beroende på i vilken ordning hoten väljs. Därför användes strategin för att generera 1 000 rangordningar. Därefter beräknades genomsnittliga värden för antalet jämförelser, som behövde nyttjas för att rangordna hoten, samt genomsnittliga värden för Spearmans rangkorrelation med den rangordning som gavs av den fullständiga bedömningsmatrisen. Beräkningarna med Harkers metod använde den 3-gradiga skalan med vikterna 1/3, 1 och 3 respektive den 9-gradiga geometriska skalan, som återfinns i Tabell 3.

Tabell 6 innehåller resultaten för konsekvens och Tabell 7 innehåller resultaten för sannolikhet. Av tabellerna framgår att vid jämförelser med den 3-gradiga skalan minskade antalet jämförelser med 71% respektive 69% för konsekvens respektive sannolikhet i förhållande till den fullständiga bedömningsmatrisens 78 jämförelser. Vid jämförelser med den 9-gradiga skalan var minskningen 68% respektive 69% för konsekvens respektive sannolikhet. Spearmans rangkorrelation var, för såväl den 3- som den 9-gradiga skalan, mycket hög mellan rangordningen baserad på den fullständiga bedömningsmatrisen och den initiala respektive den modifierade rangordningen. Dessa resultat uppnåddes för såväl konsekvens som sannolikhet. Avseende skillnaden mellan den initiala respektive den modifierade rangordningen visar tre av fyra resultat på en högre rangkorrelation för den modifierade rangordningen. Endast för jämförelserna av sannolikhet baserade på den 9-gradiga skalan erhålls en högre rangkorrelation för den initiala rangordningen.

Tabell 6: Resultat för jämförelser med avseende på konsekvens. Antalet jämförelser är ett genomsnitt av vad som krävdes för att rangordna de 13 hot som jämfördes. Spearmans rangkorrelation avser samstämmigheten med den rangordning av hoten som ges av den fullständiga ordningsmatrisen.

Metod	Antal jämförelser	Spearmans rangkorrelation
Urval av jämförelser, 3-gradig skala	22,9	0,894
Urval av jämförelser, 9-gradig skala	24,7	0,875
Harkers metod, 3-gradig skala	-	0,919
Harkers metod, 9-gradig skala	-	0,917

Tabell 7: Resultat för jämförelser med avseende på sannolikhet. Antalet jämförelser är ett genomsnitt av vad som krävdes för att rangordna de 13 hot som jämfördes. Spearmans rangkorrelation avser samstämmigheten med den rangordning av hoten som ges av den fullständiga ordningsmatrisen.

Metod	Antal jämförelser	Spearmans rangkorrelation
Urval av jämförelser, 3-gradig skala	24,3	0,930
Urval av jämförelser, 9-gradig skala	24,4	0,908
Harkers metod, 3-gradig skala	-	0,938
Harkers metod, 9-gradig skala	-	0,845

4.3 Beräkning av sannolikhets- och konsekvensvärden

En absolut skala innebära att hoten grupperas enligt värdena på skalan, vilket exempelvis är fallet med de femgradiga skalor för sannolikhet och konsekvens som vanligen används inom Försvarsmakten (Försvarsmakten, 2013). När hot istället prioriteras med en jämförelsebaserad metod såsom AHP kan relativa prioritetsvärden beräknas, men inga absoluta värden erhålls för sannolikheten att hoten realiseras respektive deras eventuella konsekvens.

I detta avsnitt beskrivs en metod som kan användas för att omvandla relativa prioritetsvärden till förslag på absoluta värden. Metoden bygger på antagandena att hot vars prioritetsvärden är relativt lika också har samma absoluta värde samt att de hot som bedömts täcker in hela den absoluta skalan från 1 till 5. Målsättningen är därmed att hitta kluster av prioritetsvärden. Om det går att hitta lika många kluster som det finns steg på den absoluta skalan antas ett tentativt förslag till absoluta värden ges av att associera ett av dessa kluster med vart och ett av stegen på den absoluta skalan.

Metoden bygger på hierarkisk klustring (Murtagh, 1983) och inleds med att skapa ett kluster med de hot som har den minsta relativa skillnaden mellan sina prioritetsvärden. Därefter skapas nya kluster med de hot (eller kluster) som har den minsta relativa skillnaden mellan sina prioritetsvärden. Ett nytt kluster kan skapas genom att lägga samman två hot, lägga till ett hot till ett befintligt kluster eller slå ihop två befintliga kluster. Varje kluster som skapas tilldelas ett prioritetsvärd som utgörs av medelvärdet av de ingående hotens prioritetsvärden. Sammanslagningen av hot och kluster avbryts när antalet kluster och antalet enskilda hot, som inte ingår i något kluster, tillsammans är lika stort som antalet steg på den absoluta skala som används. Därefter tilldelas hoten absoluta värden baserat på prioritetsvärdet för klustret de ingår i alternativt baserat på deras eget prioritetsvärde, om de inte ingår i något kluster.

För att utvärdera den föreslagna metoden användes de data som tidigare har använts för att utvärdera olika skalor samt strategin för val av vilka hot som ska jämföras. Dessa data utgörs av fyra fullständiga bedömningsmatriser för 13 hot, en matris med jämförelser enligt en 3-gradig och en matris med jämförelser enligt en 9-gradig skala för sannolikhet respektive konsekvens. Dessutom används de urval av hotpar att jämföra som togs fram i avsnitt 4.2. Dessa urval utgör delvis ifyllda bedömningsmatriser.

Tabell 8 innehåller resultaten för hoten med avseende på konsekvens och Tabell 9 innehåller motsvarande för sannolikhet när jämförelserna gjordes med den 3-gradiga skalan. I båda tabellerna har raderna sorterats efter prioritetsvärdet för

den fullständiga bedömningsmatrisen. Avseende konsekvens, för de 13 hoten jämförda med den 3-gradiga skalan (Tabell 8) har algoritmen för att tilldela absoluta värden grupperat hoten så att gränserna mellan de olika absoluta värdena sammanfaller med de största skillnaderna mellan prioritetsvärdena, vilket inte är givet på förhand eftersom sammanslagningen av hot till kluster av hot inte bara beror på skillnaden mellan intilliggande prioritetsvärden utan också på skillnaden mellan prioritetsvärden inom ett kluster. Detta illustreras av resultaten i Tabell 9 där gränsen mellan de absoluta värdena 1 och 2 har lagts mellan hoten H6 och H4 snarare än mellan hoten H12 och H2 fast den relativa skillnaden är större mellan det senare paret av hot. Anledningen till detta är att den grupp av hot som motsvarar det absoluta värdet 2 då skulle ha fått en större relativ skillnad mellan lägsta och högsta prioritetsvärde (50,3%) än vad gruppen som motsvaras av värdet 1 fick (44,7 %).

Avseende tilldelning av absoluta värden baserat på de delvis ifyllda bedömningsmatriserna framgår det av Tabell 8 och Tabell 9 att de allra flesta hot tilldelas olika absoluta värden baserat på de 1 000 olika delvis ifyllda bedömningsmatriser som skapades under testet. Undantaget är endast hotet H13 som alltid tilldelas värdet 5 avseende sannolikhet, vilket beror på att H13 i alla direkta jämförelser med andra hot bedöms ha högre sannolikhet och därmed alltid rangordnas som det hot som har högst sannolikhet. Genomsnittet för Spearmans rangkorrelation mellan absolutvärdena baserade på den fullständiga bedömningsmatrisen och de absoluta värdena baserade på de delvis ifyllda bedömningsmatriserna är dock höga, 0,907 för konsekvens och 0,924 för sannolikhet. Detta tyder på en mycket stark samstämmighet mellan de värden som erhålls med den fullständiga respektive de delvis ifyllda bedömningsmatriserna.

Tabell 8: Prioritets- och absolutvärden, avseende konsekvens, för de 13 hoten jämförda med den 3-gradiga skalan med vikterna 1/3, 1 och 3. Tabellen innehåller två uppsättningar med värden baserade på fullständigt respektive delvis ifylld bedömningsmatris. Kolumnen *skillnad* anger den relativa skillnaden mellan prioritetsvärdena för intilliggande hot. Kolumnen *spann* anger hur stor den relativa skillnaden är mellan det lägsta och högsta prioritetsvärdet som har resulterat i samma absolutvärde. Prioritets- och absolutvärdena för delvis ifylld bedömningsmatris är de medelvärden som erhölls för 1 000 delvis ifyllda bedömningsmatriser i Avsnitt 4.2.

Hot	**Fullständig bedömningsmatris**				**Delvis ifylld bedömningsmatris**	
	Prioritets-värde	Skillnad (%)	Absolut-värde	Spann	Prioritets-värde	Absolutvärde
H8	0,1554	-	5	34,4%	0,1842	4,93
H12	0,1384	12,3	5		0,1532	4,70
H2	0,1192	16,1	5		0,1256	4,38
H4	0,1156	3,1	5		0,1202	4,30
H11	0,0829	39,5	4	26,5%	0,0745	3,59
H5	0,0772	7,3	4		0,0703	3,51
H7	0,0662	16,7	4		0,0619	3,32
H6	0,0655	1,0	4		0,0604	3,28
H1	0,0435	50,7	3	11,6%	0,0392	2,54
H10	0,0404	7,7	3		0,0367	2,44
H3	0,0390	3,6	3		0,0330	2,23
H9	0,0318	22,4	2	-	0,0261	1,89
H13	0,0249	27,7	1	-	0,0147	1,02

Tabell 9: Prioritets- och absolutvärden, avseende sannolikhet, för de 13 hoten jämförda med den 3-gradiga skalan med vikterna 1/3, 1 och 3. Tabellen innehåller två uppsättningar med värden baserade på fullständigt respektive delvis ifylld bedömningsmatris. Kolumnen *skillnad* anger den relativa skillnaden mellan prioritetsvärdena för intilliggande hot. Kolumnen *spann* anger hur stor den relativa skillnaden är mellan det lägsta och högsta prioritetsvärdet som har resulterat i samma absolutvärde. Prioritets- och absolutvärdena för delvis ifylld bedömningsmatris är de medelvärden som erhölls för 1 000 delvis ifyllda bedömningsmatriser i Avsnitt 4.2.

Hot	**Fullständig bedömningsmatris**				**Delvis ifylld bedömningsmatris**	
	Prioritets-värde	Skillnad (%)	Absolut-värde	Spann	Prioritetsvärde	Absolut-värde
H13	0,1853	-	5	-	0,2615	5,00
H3	0,1253	47,9	4	10,4%	0,1281	4,05
H9	0,1253	0,0	4		0,1278	4,03
H1	0,1135	10,4	4		0,1134	3,91
H10	0,0838	35,3	3	7,7%	0,0756	3,32
H11	0,0778	7,7	3		0,0685	3,20
H7	0,0536	45,1	2	19,3%	0,0445	2,62
H8	0,0500	7,3	2		0,0413	2,54
H5	0,0489	2,2	2		0,0396	2,46
H6	0,0449	8,8	2		0,0371	2,40
H4	0,0357	26,0	1	44,7%	0,0271	1,98
H12	0,0313	14,1	1		0,0228	1,79
H2	0,0246	26,9	1		0,0128	1,01

Tabell 10 innehåller resultaten för konsekvens och Tabell 11 innehåller resultaten för sannolikhet när jämförelserna gjordes med den 9-gradiga skalan.

Avseende tilldelning av absoluta värden baserat på de delvis ifyllda bedömningsmatriserna framgår det att även när den 9-gradiga geometriska skalan används så tilldelas de allra flesta hot olika absoluta värden baserat på de 1 000 olika delvis ifyllda bedömningsmatriser som skapades under testet i Avsnitt 4.2. Undantaget är endast hotet H13 som alltid tilldelas värdet 1 avseende konsekvens respektive värdet 5 avseende sannolikhet, vilket beror på att H13 i alla direkta jämförelser med andra hot bedöms ha lägre konsekvens och högre sannolikhet

och därmed alltid rangordnas som det hot som har lägst konsekvens respektive högst sannolikhet. Genomsnittet för Spearmans rangkorrelation mellan absolutvärdena baserade på den fullständiga bedömningsmatrisen och de absoluta värdena baserade på de delvis ifyllda bedömningsmatriserna är dock höga, 0,945 för konsekvens och 0,885 för sannolikhet. Detta tyder på en mycket stark samstämmighet mellan de värden som erhålls med den fullständiga respektive de delvis ifyllda bedömningsmatriserna.

Tabell 10: Prioritets- och absolutvärden, avseende konsekvens, för de 13 hoten jämförda med den 9-gradiga geometriska skalan. Tabellen innehåller två uppsättningar med värden baserade på fullständig respektive delvis ifylld bedömningsmatris. För delvis ifylld bedömningsmatris är prioritetsvärdena (beräknade med Harkers metod (Harker, 1987)) och absolutvärdena de medelvärden som erhölls för 1 000 delvis ifyllda bedömningsmatriser skapade enligt den strategi som beskrivs i Avsnitt 4.2.

Hot	**Fullständig bedömningsmatris**		**Delvis ifylld bedömningsmatris**	
	Prioritetsvärde	Absolutvärde	Prioritetsvärde	Absolutvärde
H2	0,2590	5	0,2865	4,97
H8	0,2563	5	0,2749	4,98
H12	0,1817	5	0,1813	4,85
H4	0,1673	5	0,1613	4,83
H5	0,0331	4	0,0205	3,49
H11	0,0284	4	0,0245	3,69
H6	0,0176	3	0,0118	3,20
H1	0,0169	3	0,0120	3,20
H3	0,0144	3	0,0087	2,88
H7	0,0133	3	0,0096	3,07
H10	0,0091	3	0,0070	2,86
H9	0,0021	2	0,0015	1,84
H13	0,0008	1	0,0005	1,00

Tabell 11: Prioritets- och absolutvärden, avseende sannolikhet, för de 13 hoten jämförda med den 9-gradiga geometriska skalan. Tabellen innehåller två uppsättningar med värden baserade på fullständig respektive delvis ifylld bedömningsmatris. För delvis ifylld bedömningsmatris är prioritetsvärdena (beräknade med Harkers metod (Harker, 1987)) och absolutvärdena de medelvärden som erhölls för 1 000 delvis ifyllda bedömningsmatriser skapade enligt den strategi som beskrivs i Avsnitt 4.2.

Hot	**Fullständig bedömningsmatris**		**Delvis ifylld bedömningsmatris**	
	Prioritetsvärde	Absolutvärde	Prioritetsvärde	Absolutvärde
H13	0,7925	5	0,8459	5,00
H9	0,0734	4	0,0448	3,81
H1	0,0495	4	0,0397	3,92
H3	0,0445	4	0,0343	3,86
H10	0,0082	3	0,0064	2,54
H7	0,0068	3	0,0057	2,51
H11	0,0068	3	0,0055	2,51
H6	0,0064	3	0,0053	2,50
H5	0,0057	3	0,0049	2,47
H12	0,0027	2	0,0037	2,39
H8	0,0025	2	0,0031	2,31
H4	0,0006	1	0,0003	1,17
H2	0,0004	1	0,0002	1,03

5 Experiment med jämförelsebaserade bedömningar

Under de två endagskonferenserna *IT-säkerhetsdagen* och *Informationsinfrastrukturdagen* som arrangerades i Stockholm 2016-11-02 – 2016-11-03 gavs möjligheten att genomföra en ny enkätundersökning. Båda dagarna hade över 200 anmälda deltagare som alla bedömdes ha generellt hög expertis inom informationssäkerhetsområdet, vilket gav bra förutsättningar för både hög svarsfrekvens och ett tydligt resultat. Deltagarna kom primärt från svenska försvarsrelaterade myndigheter. Det fanns inte någon specifikt avsatt tid för deltagarna att besvara enkäten, men det gick att besvara enkäten under lunchpaus och fikapauser under respektive dag.

Att genomföra en enkätundersökning under pauser i konferenser ger vissa ramar för hur enkäten behöver utformas för att få in tillräckligt många svar. Framförallt bedömdes tidsåtgången för enkätifyllnad vara den viktigaste faktorn. Pauserna under konferenserna var på 20–60 minuter och det bedömdes som rimligt att enkätifyllnad inte fick ta mer än fem minuter att genomföra för att nå en rimlig svarsfrekvens.

Följande ramar sattes upp för utformningen av enkätundersökningen.

- **Ett enkelt scenario**
  Med tidsramen på fem minuter krävs en enkel och tydlig kontext som konferensbesökarna kan relatera till.
- **Få och generella hot**
  De hot som används i enkätundersökningen kan inte vara för många till antalet. De behöver även vara av en mer generell art för att kunna vara relevanta för flertalet av bedömarna.
- **Enkla bedömningsskalor**
  De skalor som används vid bedömningarna ska antingen vara redan etablerade skalor eller enklare skalor som inte kräver omfattande instruktioner före användande.

I de två avsnitt som följer presenteras utformningen av enkäter följt av resultatet från den genomförda enkätundersökningen.

5.1 Utformning av enkäter

Målsättningen med enkätundersökningen var att jämföra samstämmigheten mellan respondenter utifrån två olika metoder för att bedöma hot. Eftersom undersökningen omfattade två olika metoder krävdes två separata enkäter. Den ena enkäten utgick från metoden för absoluta bedömningar enligt H Säk Grunder (Försvarsmakten, 2013) medan den andra utgick från jämförelsebaserade

bedömningar enligt AHP (Kapitel 4). Hälften av respondenterna förväntades svara på den ena enkäten och hälften på den andra.

Något som var gemensamt för de båda enkäterna var uppsättningen av hot att bedöma. Hoten som valdes var av mer generell art som kan tänkas vara vanligt förekommande i riskanalyser hos de flesta organisationer. Följande fem hot valdes från listan med hot som presenterades i Avsnitt 4.1.1.

1. Infektion av s.k. ransomware leder till att data krypteras och krav på lösensumma ställs för att krypterade data ska återställas.
2. Attacker riktas och anpassas mot användare (spear fishing) av informationssystem för att lura dessa att ge ifrån sig information alternativt installera skadlig mjukvara.
3. Sårbarheter i hårdvara nyttjas av extern aktör för att komma åt information.
4. DDoS-attack som stör verksamheten genomförs av hacktivistgrupp (t.ex. Anonymous).
5. Skadlig kod som cirkulerar på Internet infekterar dator i kontorsnät.

Även om antalet hot blev lågt bedömdes dessa vara tillräckligt representativa för att det skulle kunna vara möjligt att dra slutsatser från resultatet. En ytterligare anledning till valet av dessa mer allmänna hot var att det för konferensdeltagarna dessutom kan tänkas finnas ett visst intresse av att få se de resulterande rangordningarna.

Avseende scenario, som syftar till att ge en kontext till de bedömningar som skulle göras, angavs för båda enkäterna att bedömningarna skulle göras avseende *alla svenska statliga myndigheter tillsammans under 2017*. Vad gäller konsekvensen skulle bedömningen avse ett *genomsnittligt värde för alla svenska statliga myndigheter under 2017*.

Med tanke på den valda uppsättningen av hot och valt scenario var det troligt att sannolikheten för att hoten realiseras under 2017 skulle vara nära 1 för samtliga hot. Därför valdes att enkäten istället skulle efterfråga bedömning av frekvens, det vill säga hur många gånger hotet bedöms realiseras under 2017.

Enkäten med absoluta bedömningar använde den föreslagna skalan för konsekvensbedömningar från H Säk Grunder (Tabell 12).

Tabell 12: Absolut skala för bedömning av konsekvens.

Konsekvensvärde	Beskrivning
5	Synnerligen allvarlig
4	Allvarlig
3	Kännbar
2	Lindrig
1	Försumbar

Eftersom bedömning av sannolikhet ersatts av bedömning av frekvens kunde inte den föreslagna sannolikhetsskalan i H Säk Grunder användas. Istället skapades egenhändigt en bedömningsskala för frekvens (Tabell 13).

Tabell 13: Absolut skala för bedömning av frekvens.

Frekvensvärde	Beskrivning	
5	Mycket ofta	Minst 1/vecka
4	Ofta	Minst 1/månad
3	Återkommande	Minst 1/kvartal
2	Sällan	Minst 1/år
1	Mycket sällan	Inte varje år

För enkäten som utgår från jämförelsebaserade bedömningar användes en 3-gradiga skala, i enlighet med Avsnitt 4.1.3, som enbart omfattade de tre alternativen *A är större*, *lika stora*, *B är större*. Denna skala användes både för bedömning av frekvens och konsekvens. Genom att använda den 3-gradiga skalan var tanken att se om jämförelsebaserade bedömningar, i sin enklaste form, kan vara en möjlig väg för att nå bättre samstämmighet mellan olika bedömare.

Utöver bedömning av de fem hoten fanns det även med två frågor på båda enkäterna som avsåg fånga respondenternas erfarenhet inom området informationssäkerhet samt av att bedöma informationssäkerhetsrisker. Dessa frågor syftade till att möjliggöra identifiering av experter bland respondenterna.

Totalt sett innefattade enkäten med absoluta bedömningar tolv frågor – fem bedömningar av frekvens, fem bedömningar av konsekvens samt två bedömningar av erfarenhet. För enkäten med jämförelsebaserade bedömningar gjordes en komplett uppsättning med jämförelser, vilket resulterade i 22

bedömningar – 10 jämförelser avseende frekvens, 10 jämförelser avseende konsekvens samt 2 bedömningar av erfarenhet.

Innan enkätundersökningen genomfördes formulerades följande fyra hypoteser kring resultatet.

Hypotes 1. Samstämmigheten mellan olika bedömares konsekvensbedömningar är högre vid jämförelsebaserade bedömningar än vid absoluta bedömningar.

Hypotes 2. Samstämmigheten mellan olika bedömares frekvensbedömningar är högre vid jämförelsebaserade bedömningar än vid absoluta bedömningar.

Hypotes 3. Vid användandet av jämförelsebaserade bedömningar fås högre samstämmighet mellan olika bedömares konsekvensbedömningar när experter gör bedömningarna.

Hypotes 4. Vid användandet av jämförelsebaserade bedömningar fås högre samstämmighet mellan olika bedömares frekvensbedömningar när experter gör bedömningarna.

5.2 Resultat

30 stycken komplett ifyllda enkäter mottogs för enkäten med absoluta bedömningar och 25 stycken komplett ifyllda enkäter mottogs för enkäten med jämförelsebaserade bedömningar. Totalt motsvarade detta en svarsfrekvens på cirka 25%. Fördelningen av svar för båda enkäterna återges i Bilaga 3. Konferensdeltagarna ombads att ta den enkät som låg överst i högen av enkäter som var sorterad så att varannan enkät var enligt absoluta metoden och varannan enligt den jämförelsebaserade. Trots sorteringen så valde en del konferensdeltagare att leta i högen tills de fick den som använde absoluta jämförelser då denna enkät var på en sida istället för två. Efterhand justerades dock högen för att få en jämnare fördelning, vilket lyckades delvis.

5.2.1 Konsekvensbedömningarna

Bedömningen av konsekvens enligt de två olika metoderna resulterade i två rangordningar som skilde sig åt avseende vilket hot som ansågs ha störst konsekvens respektive näst störst konsekvens (Tabell 14). För övrigt var de två rangordningarna identiska. Resultatet visar inte på någon statistiskt signifikant[23]

[23] För $p < 0,05$.

skillnad avseende samstämmigheten vid användandet av de två metoderna[24]. Därmed föll hypotes 1.

Tabell 14: Resulterande rangordning av hotens konsekvens baserat på samtliga respondenters bedömningar för respektive metod. För metoden med absoluta bedömningar (ABS) anges även snittvärdet på konsekvensen för varje hot. För metoden med jämförelsebaserade bedömningar (JMF) anges även snittet av hur ofta varje hot angavs ha större konsekvens än ett annat hot, borträknat hur många gånger hotet angavs ha mindre konsekvens än ett annat hot.

ABS			**JMF**	
Hot	Värde		Hot	Värde
H3	3,83		H2	1,60
H2	3,77		H3	1,12
H1	3,70		H1	-0,24
H5	3,20		H5	-0,80
H4	2,93		H4	-1,68

5.2.2 Frekvensbedömningarna

Bedömningen av frekvens enligt de två olika metoderna resulterade i två rangordningar som skilde sig åt avseende alla placeringar förutom andraplaceringen som i båda rangordningarna innehades av H5 (Tabell 15). Resultatet visar på en statistiskt signifikant[25] skillnad avseende samstämmigheten vid användandet av de två metoderna[26] till fördel för metoden med absoluta bedömningar. Därmed föll hypotes 2.

[24] Spearman totalt i snitt för metoden med absoluta bedömningar 0,209 och Spearman totalt i snitt för jämförelsebaserade bedömningar 0,211. Metoden med absoluta bedömningar gav även Intra-class correlation på 95% CI för single measures på 0,060–0,0699.

[25] Vid p = 0,01 < 0,05 med 95% konfidensintervall för skillnaden på 0,081–0,287 till fördel för metoden med absoluta bedömningar.

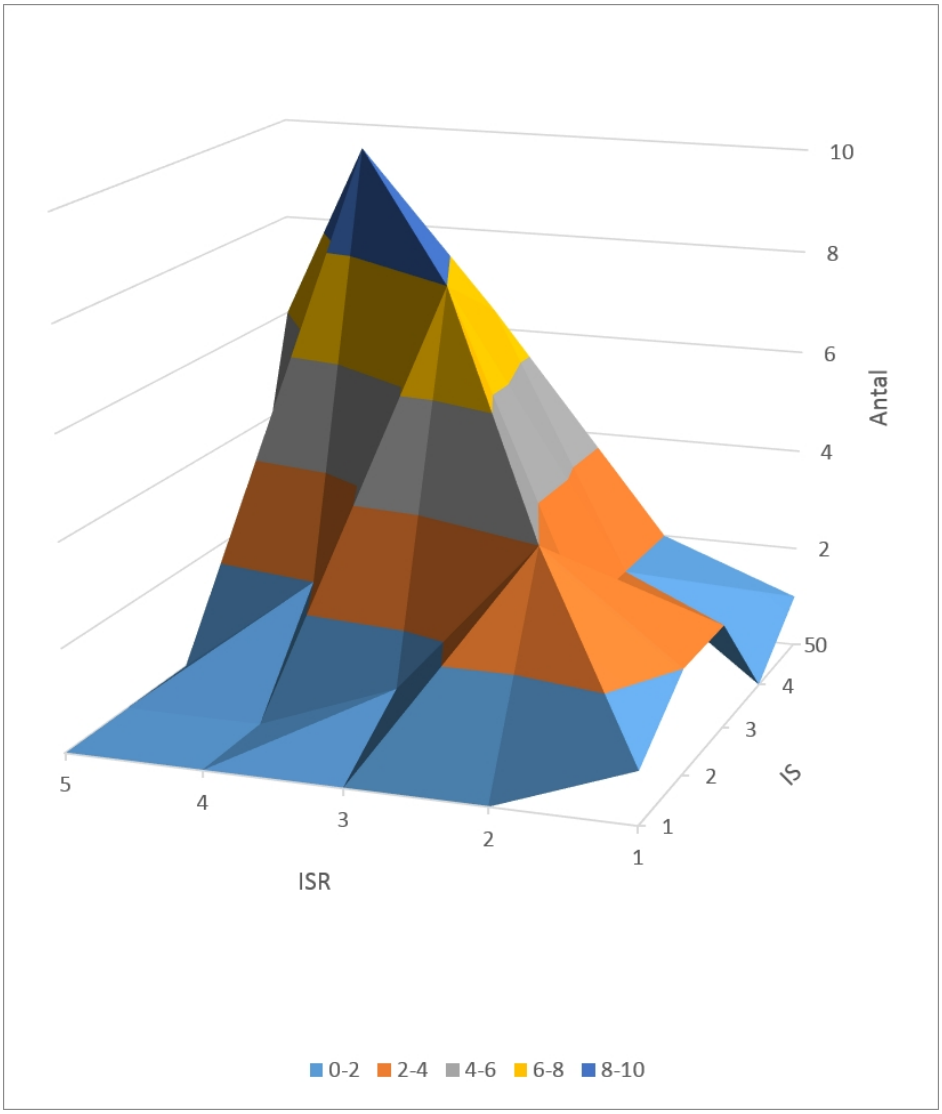
[26] Spearman totalt i snitt för metoden med absoluta bedömningar 0,249 och Spearman totalt i snitt för jämförelsebaserade bedömningar 0,065. Metoden med absoluta bedömningar gav även Intra-class correlation på 95% CI för single measures på 0,076–0,735.

Tabell 15: Resulterande rangordning av hotens frekvens baserat på samtliga respondenters bedömningar för respektive metod. För metoden med absoluta bedömningar (ABS) anges även snittvärdet på frekvens för varje hot. För metoden med jämförelsebaserade bedömningar (JMF) anges även snittet av hur ofta varje hot angavs ha högre frekvens än ett annat hot, borträknat hur många gånger hotet angavs ha lägre frekvens än ett annat hot.

ABS			**JMF**	
Hot	Värde		Hot	Värde
H2	3,43		H4	1,04
H5	3,40		H5	0,96
H4	2,93		H2	-0,28
H3	2,70		H1	-0,68
H1	2,17		H3	-1,04

5.2.3 Expertis

I båda enkäterna fanns en fråga om huruvida respondenten arbetar med att bedöma informationssäkerhetsrisker (ISR) samt en fråga om huruvida respondenten arbetar med informationssäkerhet (IS). Dessa två frågor resulterade i en självuppskattad expertisnivå som återges i Figur 10. Z-axeln (höjdled) visar antalet svar med den kombination av svar som X-axeln (IS) och Y-axeln (ISR) anger.



Figur 10: Fördelningen av svar på frågorna rörande expertis.

För att vidare analysera resultatet delades respondenterna in i grupper baserat på deras svar på expertisfrågorna. Expertisfrågorna besvarades på en 5-gradig skala (1–5) och ett svar som var 3–5 tolkades som högt medan ett svar på 1–2 tolkades som lågt. Resultatet presenteras i Tabell 16.

Tabell 16: Gruppindelning av respondenter för båda enkäterna utifrån svar på expertisfrågor. Kombinationen högt på ISR och lågt på IS förbisågs då den kombinationen ansågs irrationell, vilket resulterade i ett bortfall på en respondent.

Expertis bedömning av informationssäkerhetsrisker (ISR)	Expertis informationssäkerhet i allmänhet (IS)	Grupp	**Antal**
Högt	Högt	Expertis 3	38
Lågt	Högt	Expertis 2	9
Lågt	Lågt	Expertis 1	7

Expertisgrupp 3 var klart större än övriga i antal (Figur 10) vilket indikerar att respondenterna som helhet var mer informationssäkerhetsexperter än IT-arbetande i allmänhet. Detta var också förväntat med tanke på att enkäten genomfördes i samband med en IT-säkerhetskonferens riktad mot svenska försvarsrelaterade myndigheter. Det är dock tänkbart att en del potentiella respondenter valde att inte delta på grund av att hoten faktiskt är relevanta även i skarpt läge och att information om hur dessa hot bedöms var något de inte kunde delge i publika sammanhang.

5.2.3.1 Konsekvens

För enkäten med absoluta bedömningar hade expertisgrupp 3 en intern samstämmighet som var statistiskt signifikant lägre än expertisgrupp 1[27]. Vid jämförelse mellan expertisgrupp 3 och expertisgrupp 2 var den interna samstämmigheten även då statistiskt signifikant lägre för expertisgrupp 3[28]. För enkäten med den jämförelsebaserade metoden hade expertisgrupp 3 en intern samstämmighet som var statistiskt signifikant högre både jämfört med expertisgrupp 1[29] och med expertisgrupp 2[30]. Övriga skillnader var ej statistiskt signifikanta.

Om respondenterna istället delades in i expertisgrupper baserat på att *högt* definieras som 4–5 och *lågt* som 1–3 resulterade gruppindelningen för den jämförelsebaserade metoden i en expertisgrupp 1 och en expertisgrupp 3 som vardera innefattade 10 respondenter av totalt 25 respondenter. På så sätt blev resultatet två grupper med större skillnad i uppskattad expertisnivå än tidigare indelning. Även med denna gruppindelning var samstämmigheten statistiskt

[27] p = 0,007 med en skillnad på 0,085–0,580 (95% CI)
[28] p = 0,020 med en skillnad på 0,035–0,457 (95% CI)
[29] p < 0,001 med en skillnad på 0,301–0,850 (95% CI)
[30] p < 0,001 med en skillnad på 0,511–1,060 (95% CI)

signifikant högre för expertisgrupp 3 jämfört med expertisgrupp 1[31]. Därmed bekräftades hypotes 3.

Generellt sett hade expertisgrupp 3 högre samstämmighet än övriga grupper när den jämförelsebaserade bedömningar användes, men samtidigt lägre samstämmighet vid användandet av absoluta bedömningar. Skillnaden i samstämmighet verkar dock vara klart större när jämförelsebaserade bedömningar användes.

5.2.3.2 Frekvens

För enkäten med absoluta bedömningar hade expertisgrupp 3 en intern samstämmighet som var statistiskt signifikant lägre än expertisgrupp 1[32]. För enkäten med den jämförelsebaserade metoden hade expertisgrupp 3 en intern samstämmighet som var statistiskt signifikant högre än expertisgrupp 1[33] och även statistiskt signifikant högre än expertisgrupp 2[34]. Övriga skillnader var ej statistiskt signifikanta.

Även om respondenterna istället delades in i expertisgrupper baserat på att *högt* definierades som 4–5 och *lågt* som 1–3, på samma sätt som i Avsnitt 5.2.3.1, var samstämmigheten för expertisnivå 3 statistiskt signifikant högre än för expertisgrupp 1[35]. Därmed bekräftades hypotes 4.

Generellt sett hade expertisgrupp 3 högre samstämmighet än övriga grupper när den jämförelsebaserade bedömningar användes, men samtidigt lägre samstämmighet vid användandet av absoluta bedömningar. Skillnaden i samstämmighet verkar dock vara klart större när jämförelsebaserade bedömningar användes.

5.3 Rekommendationer

Experimentet visar att jämförelsebaserade bedömningar kan fungera bättre än absoluta bedömningar under förutsättning att bedömningarna görs av experter. Den i experimentet använda varianten av jämförelsebaserad metod ger dock ganska grova bedömningar, så frågan är om skalan måste – och kan – göras mer finmaskig för att metoden ska vara användbar vid riskbedömningar.

[31] $p < 0{,}001$ med en skillnad på 0,358–0,634 (95% CI)
[32] $p = 0{,}043$ med en skillnad på 0,006–0,454 (95% CI)
[33] $p < 0{,}001$ med en skillnad på 0,261–0,784 (95% CI)
[34] $p < 0{,}001$ med en skillnad på 0,455–0,977 (95% CI)
[35] $p = 0{,}001$ med en skillnad på 0,167–0,536 (95% CI)

6 Diskussion

Det är en formidabel uppgift att beskriva och bedöma hot mot IT-system. Uppgiften kan lösas, men det krävs kompromisser avseende säkerhetsanalysernas upplösning och precision. De experiment som har genomförts inom det aktuella projektet under 2015 och 2016 har visat att bedömningar av övergripande beskrivna hot leder till mycket låg samstämmighet avseende sannolikhet och konsekvens. Under 2016 har projektet fokuserat på två centrala aspekter av säkerhetsanalyserna: hur de hot som utgör grunden för säkerhetsanalysens slutresultat ska beskrivas samt hur de beskrivna hoten ska bedömas.

Avseende beskrivning av hot är en utmaning att det finns få, om ens några, explicita och validerade specifikationer av vad som ska ingå i en hotbeskrivning och hur dessa ingående delar ska specificeras för konkreta hot. Det är viktigt att det är tydligt vad de olika delarna av en hotbeskrivning ska innehålla. I Avsnitt 2.3 nämns att detaljnivån i bedömningsunderlaget kan vara av varierande art. Det nämns som exempel att hur den oönskade händelsen beskrivs kan få påverkan på främst sannolikhetsbedömningen. I Tabell 2.4 på sidan 51 i H Säk Grunder (Försvarsmakten, 2013) visas exempel på hur underlaget kan se ut efter att steg 1 av säkerhetsanalysen genomförts. Exemplet visar tydligt på hur beskrivningen av en oönskad händelse kan innehålla information om både tillvägagångssätt och konsekvens.

Ett möjligt tydliggörande av hotbeskrivningens olika delar är att dela upp beskrivningen av oönskad händelse genom att tillföra hotbeskrivningsdelen *tillvägagångssätt*. Tabell 17 innehåller ett exempel som illustrerar hur införandet av hotbeskrivningsdelen tillvägagångssätt kan renodla beskrivningen av oönskad händelse. Införandet av tillvägagångssätt löser dock inte upp alla svårigheter med att få hotbeskrivningar som är konsistenta med varandra med avseende på vilken typ av information som förs in i de olika delarna av hotbeskrivningen. Vidare studier av hur de i hotbeskrivningen ingående delarna ska specificeras är nödvändiga för att minska förekomsten av annan information än vad som är tänkt i hotbeskrivningens olika delar. Minskade möjligheter till variation i framtaget underlag skulle leda till minskat personberoende vid framtagande av hotbeskrivningar, öka spårbarheten hos säkerhetsanalyserna och ge ökad konsistens hos underlaget för bedömning av hots sannolikhet och konsekvens. Tydliga instruktioner för beaktande av underlag har inom andra områden visats ha positiva effekter även när det är erfarna personer som genomför analyserna (Knapp & Knapp, 2001).

Tabell 17: Exempel där kolumnen *Oönskad händelse* har delats upp genom att införa den nya kolumnen *Tillvägagångssätt*.

Tillgång	**Oönskad händelse**	**Tillväga-gångssätt**	**Konsekvensbeskrivning**	**Konsekvens-bedömning**
Dator-utrustning i lektionssal	Förlust av dator-utrustning	Inbrott i dator-lektionssal	Värdet på utrustningen uppgår till 500kkr. Vid förlust kommer införandet av det nya verksamhets-ledningssystemet att försenas med ca 6 månader.	4
		Brand i dator-lektionssal		
		Över-svämning i dator-lektionssal		

I denna rapport presenterades resultaten från ett experiment där användandet av strukturerade tabeller alternativt naturligt språk för att beskriva hot testades. Sammanfattningsvis föll alla sju hypoteserna och det kunde inte *med statistisk signifikans* påvisas att det finns några fördelar med att använda sig av strukturerade tabeller istället för naturligt språk vid beskrivning av hot vars sannolikhet och konsekvens ska bedömas. Det finns dock indikationer på att användningen av strukturerade tabeller kan ha fördelar när det är experter som genomför bedömningarna.

Avseende bedömning av sannolikheter och konsekvenser kopplade till beskrivna hot har projektet under 2015 och 2016 studerat möjligheten att använda bedömningsmetoder baserade på jämförelser av hot snarare än att hoten bedöms direkt på en absolut skala. Det finns skäl att anta att det skulle ge flera fördelar att nyttja jämförelsebaserade metoder för att bedöma sannolikheter och konsekvenser kopplade till hot mot IT-system. Tidiga forskningsresultat visade på att det är enklare att bedöma relativa skillnader än att sätta absoluta värden (Thurstone, 1927; Yokoyama, 1921). Vidare är prioritering av risker en viktig uppgift vid genomförandet av säkerhetsanalyser och jämförelsebaserade metoder genererar just prioritetsvärden för de bedömda hoten. Beroende på hur en jämförelsebaserad metod realiseras ger den också olika möjligheter att beräkna konsistensvärden för de genomförda bedömningarna. Mot bakgrund av svårigheterna förknippade med att bedöma sannolikheter och konsekvenser kopplade till hot kan det utgöra ett verktyg för förbättringsarbete med avseende på själva bedömningarna. Samtidigt finns det nackdelar med jämförelsebaserade metoder. Antalet bedömningar (genom jämförelser) som behöver genomföras ökar med kvadraten på antalet hot om alla hot ska jämföras direkt med varandra.

Varje bedömning kräver samtidigt beaktande av egenskaperna hos två stycken hot, vilket kan öka den mentala ansträngningen även om det totala antalet hot som måste beakta är detsamma. Om säkerhetsanalysen i slutändan ska leda till absoluta bedömningar av risker måste de prioritetsvärden som jämförelsebaserade metoder ger på något sätt omvandlas till absoluta värden.

I Kapitel 4 presenterades centrala aspekter av jämförelsebaserade metoder för bedömning av sannolikheter och konsekvenser kopplade till hot. Först och främst gällande vilken skala som ska användas vid genomförandet av jämförelser. Vidare studerades en strategi för urval av vilka hot som ska jämföras för att minska antalet jämförelser utan att det påverkar den slutliga bedömningen av sannolikheter och konsekvenser. Dessutom föreslogs en metod för att ge förslag på absoluta värden för sannolikheter och konsekvenser kopplade till hot utgående från deras prioritetsvärden. En generell slutsats baserat på resultaten i Kapitel 4 är att i och med de svårigheter som finns med att bedöma sannolikheter och konsekvenser bör det vara lämpligt att nyttja en förenklad skala vid jämförelser av hot.

I Kapitel 5 beskrivs ett experiment som har genomförts med en 3-gradig skala för jämförelsebaserade bedömningar samt en 5-gradig skala för absoluta bedömningar av hot. Baserat på det genomförda experimentet går det inte att visa att jämförelsebaserade metoder leder till högre samstämmighet. Däremot visar resultaten att experter är mer samstämmiga än övriga vid bedömning av hot när de använder jämförelsebaserade bedömningar. Resultaten pekar således på att jämförelsebaserade metoder kan vara att föredra när det är experter som genomför jämförelserna. Oavsett använd metod finns det många svårigheter kopplade till bedömning av hots sannolikhet (alternativt frekvens) och konsekvens. Dessa svårigheter kan påverka resultaten, vilket exempelvis diskuteras i Kapitel 2, och resulterar i generellt sett låg samstämmighet.

7 Slutsatser

Det krävs omfattande insatser för att genomföra säkerhetsanalyser som ger heltäckande beskrivningar av samtliga relevanta hot. Dessa hot ska dessutom bedömas, med beaktande av validitet såväl som precision. Givet dessa övergripande krav är det ytterst relevant att klargöra vad resultaten från dessa analyser ska användas till. Adekvata resultat från säkerhetsanalyser förutsätter att det är tydligt vilka behov hos olika intressenter som ligger till grund för analyserna samt vilka krav på genomförande som dessa behov ger upphov till.

Baserat på det arbete som har presenterats i denna rapport dras följande slutsatser att beakta vid Försvarsmaktens fortsatta arbete med säkerhetsanalyser.

Använd strukturerade tabeller för att beskriva hot
Strukturerade tabeller framstår som ett bra alternativ när de som genomför bedömningar av sannolikheter och konsekvenser kopplade till hot baserat på dessa hotbeskrivningar är erfarna. För icke-experter är troligen hotbeskrivningar baserade på naturligt språk lämpligare.

Tydliggör specifikationen för hotbeskrivningar
Det bör finnas en tydlig specifikation av hur hot mot IT-system ska beskrivas. Detta för att minska personberoendet avseende utformningen av hotbeskrivningar och förbättra förutsättningarna för samstämmiga bedömningar av sannolikheter och konsekvenser kopplade till hot.

Tydliggör förutsättningar för bedömningar
När sannolikheter och konsekvenser kopplade till hot ska bedömas är det inte bara deras beskrivningar som är väsentliga. Andra förutsättningar bör också vara tydligt specificerade, vilket inkluderar att

- tydligt avgränsa vilket system som avses och dess omfattning
- tydligt avgränsa omfång vid bedömning av konsekvens
- tydligt specificera för vilken tidsperiod som sannolikheter, alternativt frekvenser, ska bedömas.

Tydliggör vilken precision som behövs
Det är svårt att bedöma sannolikheter och konsekvenser kopplade till hot. För att på bästa sätt använda de resurser som finns tillgängliga för genomförandet av säkerhetsanalyser är det viktigt med tydlighet avseende vilken precision dessa bedömningar behöver ha. Vilken precision som behövs beror på hur resultaten ska nyttjas. Om nödvändig precision enbart är en rangordning av hoten utgör jämförelsebaserade metoder ett alternativ till direkta bedömningar av sannolikheters och konsekvensers absoluta värden.

8 Referenser

Alonso, J. A., & Lamata, M. T. (2006). Consistency in the analytic hierarchy process: a new approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *14*(4), 445–459.

Beek, C., Castillo, C., Cochin, C., Hinchliffe, A., Jarvis, J., Li, H., … Zhu, S. (2016). McAfee Labs 2016 Threats Predictions McAfee Labs offers a. *McAfee Labs*, (2016), 1–40.

Bengtsson, J., Lundholm, K., & Hallberg, J. (2012). *Behov av stöd vid genomförande analyser för IT-system inom Försvarsmakten*. FOI-R--3452--SE. Totalförsvarets forskningsinstitut, FOI.

Bundesamt für Sicherheit in der Informationstechnik (BSI). (2005). *IT-Grundschutz Catalogues: Version 2005*.

Champ, P. A., Moore, R., & Bishop, R. C. (2009). A comparison of approaches to mitigate hypothetical bias. *Agricultural and Resource Economics Review*, *38*(2), 166–180.

Fenz, S., Heurix, J., Neubauer, T., & Pechstein, F. (2014). Current challenges in information security risk management. *Information Management & Computer Security*, *22*, 410–430.

Fink, A., & Neubauer, A. C. (2001). Speed of information processing, psychometric intelligenceand time estimation as an index of cognitive load. *Personality and Individual Differences*, *30*(6), 1009–1021.

Forman, E. H., & Selly, M. A. (2001). *Decision by Objectives: How to Convince Others that You are Right*. World Scientific Publishing Company.

Franek, J., & Kresta, A. (2014). Judgment Scales and Consistency Measure in AHP. *Procedia Economics and Finance*, *12*(March), 164–173.

Försvarsmakten. (2001). *Handbok för Försvarsmaktens Säkerhetstjänst, Informationsteknik Hotbeskrivning (H SÄK IT Hot)*. M7745-734051. Försvarsmakten.

Försvarsmakten. (2006). *Handbok för Försvarsmaktens Säkerhetstjänst, Informationsteknik (H SÄK IT)*. M7745-734062. Försvarsmakten.

Försvarsmakten. (2013). *Handbok Säkerhetstjänst Grunder (H Säk Grunder)*. M7745-734011. Försvarsmakten.

Försvarsmakten. (2014). *KSF: Krav på IT-säkerhetsförmågor hos IT-system, v3.1*. Försvarsmakten.

Hallberg, J., Bengtsson, J., & Karlzén, H. (2015). *Bedömning av sannolikhet och konsekvens för informationssäkerhetsrisker - En studie av vikt*. FOI-R--4152--

SE. Totalförsvarets forskningsinstitut, FOI.

Harker, P. T. (1987). Alternative modes of questioning in the analytic hierarchy process. *Mathematical Modelling*, *9*(3–5), 353–360.

Ishizaka, A., & Labib, A. (2011). Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications*, *38*(11), 14336–14345.

ISO International Organization for Standardization. (2009). *Guide 73 Risk management — Vocabulary*. *ISO Guide 73:2009*.

Ji, P., & Jiang, R. (2003). Scale transitivity in the AHP. *Journal of the Operational Research Society*, *54*(8), 896–905.

Knapp, C. A., & Knapp, M. C. (2001). The effects of experience and explicit fraud risk assessment in detecting fraud with analytical procedures. *Accounting, Organizations and Society*, *26*(1), 25–37.

Korman, M., Sommestad, T., Hallberg, J., Bengtsson, J., & Ekstedt, M. (2014). Overview of Enterprise Information Needs in Information Security Risk Assessment. *18th IEEE International Enterprise Distributed Object Computing Conference (EDOC)*, 42–51.

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*(1), 49–63.

Murtagh, F. (1983). A Survey of Recent Advances in Hierarchical Clustering Algorithms. *The Computer Journal*, *26*(4), 354–359.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*(4), 429–434.

Pan, L., & Tomlinson, A. (2016). A systematic Review of Information Security Risk Assessment. *International Journal of Safety and Security Engineering*, *6*(2), 270–281.

Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, *48*(1), 9–26.

Sedgewick, R., & Wayne, K. (2011). *Algorithms* (4th ed.). Addison-Wesley Professional.

Shanteau, J. (2015). Why task domains (still) matter for understanding expertise. *Journal of Applied Research in Memory and Cognition*, *4*(3), 169–175.

SIS. (2007). *SIS HB 550: Terminologi för informationssäkerhet, utgåva 3*.

SIS. (2015). *Terminologi för informationssäkerhet*. Teknisk rapport, SIS-TR 50:2015.

Sommestad, T., Karlzén, H., Nilsson, P., & Hallberg, J. (2016). An empirical test of

the perceived relationship between risk and the constituents severity and probability. *Information & Computer Security*, *24*(2), 194–2,4.

Thurstone. (1927). A law of comparative judgment. *Psychlogical Review*, *34*, 273–286.

Yokoyama, M. (1921). The Nature of the Affective Judgment in the Method of Paired Comparisons. *The American Journal of Psychology*, *32*(3), 357–369.

Bilaga 1 Scenario

Nedan återges det scenario som användes vid genomförandet av experimentet som återges i Kapitel 3 där bedömningar av sannolikhet och konsekvens baserat på olika hotbeskrivningssätt jämfördes.

Införande av ZAP hos Truzio

Truzio är ett svenskt företag som tillhandahåller tjänster för penningtransaktioner i en stor del av världens länder. Både privatpersoner och företag använder Truzios tjänster för att överföra pengar till och från aktörer i andra länder, speciellt de länder där det kan vara svårt att nå fram genom en vanlig bank. Truzio har cirka 500 anställda som alla jobbar på kontoret i Stockholm. På Truzio planeras för närvarade ett införande av affärssystemet ZAP för att stödja ekonomienhetens arbete. ZAP är tänkt att stödja hanteringen av inkommande leverantörsfakturor.

Då Truzio köper produkter och tjänster från ett stort antal leverantörer är den manuella hanteringen av inkommande leverantörsfakturor betungande för ekonomienheten. Truzio är dessutom hårt skuldsatt på grund av de senaste årens expansion av verksamheten, vilket tidvis resulterat i försenade betalningar av leverantörsfakturor. Därför har ledningen uttryckt viljan att rationalisera genom att i högsta möjliga grad automatisera hanteringen av leverantörsfakturor. ZAP förväntas även ge en bättre bild av pengaflödet genom att rapporter kan skapas direkt från alla registrerade leverantörsfakturor.

Ekonomienheten ansvarar för att Truzios ekonomiska åtaganden och skyldigheter sköts enligt lagar, avtal och rutiner. ZAP är tänkt att användas från kontorsdatorer över kontorsnätverket av personalen på Truzio. Det är endast möjligt att ansluta till ZAP från kontorsdatorer anslutna till det interna nätverket. Det interna nätverket är anslutet till Internet. All personal får ett personligt användarkonto (användarnamn och lösenord) i ZAP med rättigheter baserat på vilken roll de ska ha i systemet.

Leverantörer fakturerar Truzio genom att antingen skicka elektronisk faktura via epost eller skicka pappersfaktura som skannas in av ekonomienheten på Truzio. Den inkomna fakturan registreras i ZAP av ekonomienheten. Därefter skickas ett epostmeddelande ut till attesteraren om att det finns inkommen faktura att attestera i ZAP. Attesteraren attesterar fakturan i ZAP som då skapar en elektronisk betalningsorder. En gång per dygn skickas framtagna betalningsorder via en krypterad länk till banken Bankus som genomför betalningen med medel från Truzios bankkonto. Leverantörsfakturan arkiveras i elektroniskt format i ZAP då betalda fakturor utgör bokföringsunderlag som enligt lag måste sparas i minst 7 år. Den resulterande databasen används även vid framtagandet av ekonomiska rapporter.

I planeringsfasen av införandet av ZAP har leverantören av ZAP påbörjat en riskanalys för att utvärdera om en acceptabel risknivå kan nås med den av ledningen tilldelade budgeten. Riskanalysarbetet har än så länge resulterat i 23 konkreta hot som ännu inte har bedömts. Då du i rollen som informationssäkerhetschef på Truzio är ansvarig för att säkerställa att införandet av ZAP inte utsätter Truzio för risker som företaget inte är beredd att ta, vill ledningen nu att du fortsätter det påbörjade riskanalysarbetet genom att bedöma de identifierade hoten.

Ledningen har uttryckt farhågor för att eventuella informationssäkerhetsincidenter kan äventyra förtroendet för Truzios penningtransaktionstjänster, vilket vore förödande för verksamheten. Ett införande av ZAP får därför inte äventyra kundernas nuvarande förtroende för Truzio och dess tjänster.

Två av Truzios konkurrenter i penningtransaktionsbranschen, det ryska statsägda företaget Eastern Separation och det tyska börsnoterade företaget Xerof, har under de senaste åren lidit av att deras tjänster har sammankopplats med pengatvätt, organiserad brottslighet och finansiering av terrorism. Detta har varit gynnsamt för Truzio som tack vare detta lyckats överta en delmängd av konkurrenternas tidigare kunder som inte vill bli sammankopplade med brottslig verksamhet. Truzios ledning ser därför en förvärrad hotbild med revanschsugna konkurrenter och ökat intresse från brottslingar.

Bilaga 2 Testdata

I denna bilaga återfinns de fyra bedömningsmatriser som fylldes i för att kunna genomföra olika test i Kapitel 4.

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}	E_{13}
E_1	1	1/5	1	1/3	1/3	1	1	1/5	3	3	1	1/5	7
E_2	5	1	5	1	5	5	5	1	7	7	5	3	9
E_3	1	1/5	1	1/5	1/5	1/3	1/3	1/5	3	1	3	1/5	5
E_4	3	1	5	1	5	5	5	1	7	5	3	1	9
E_5	3	1/5	5	1/5	1	1	1	1/5	5	1	1/3	1/3	7
E_6	1	1/5	3	1/5	1	1	1	1/5	5	1	1/3	1/5	7
E_7	1	1/5	3	1/5	1	1	1	1/5	3	1	1/3	1/7	5
E_8	5	1	5	1	5	5	5	1	9	5	5	3	9
E_9	1/3	1/7	1/3	1/7	1/5	1/5	1/3	1/9	1	1/3	1/5	1/7	3
E_{10}	1/3	1/7	1	1/5	1	1	1	1/5	3	1	1/3	1/5	3
E_{11}	1	1/5	1/3	1/3	3	3	3	1/5	5	3	1	1/5	5
E_{12}	5	1/3	5	1	3	5	7	1/3	7	5	5	1	7
E_{13}	1/7	1/9	1/5	1/9	1/7	1/7	1/5	1/9	1/3	1/3	1/5	1/7	1

Figur 11: Jämförelser av de 13 hoten avseende konsekvens med 9-gradig skala. I denna bedömningsmatris används vikter från den klassiska AHP-skalan.

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}	E_{13}
E_1	1	7	3	5	5	5	3	5	1	5	5	7	1/7
E_2	1/7	1	1/7	1/3	1/5	1/7	1/7	1/5	1/9	1/5	1/5	1/3	1/9
E_3	1/3	7	1	7	5	5	5	5	1	5	3	7	1/5
E_4	1/5	3	1/7	1	1/5	1/5	1/5	1/3	1/7	1/5	1/5	1/7	1/9
E_5	1/5	5	1/5	5	1	1	1	3	1/5	1	1	5	1/9
E_6	1/5	7	1/5	5	1	1	1	3	1/7	1	1	5	1/9
E_7	1/3	7	1/5	5	1	1	1	3	1/7	1	1	5	1/9
E_8	1/5	5	1/5	3	1/3	1/3	1/3	1	1/7	1	1	3	1/9
E_9	1	9	1	7	5	7	7	7	1	3	3	5	1/7
E_{10}	1/5	5	1/5	5	1	1	1	1	1/3	1	1	5	1/5
E_{11}	1/5	5	1/3	5	1	1	1	1	1/3	1	1	5	1/7
E_{12}	1/7	3	1/7	7	1/5	1/5	1/5	1/3	1/5	1/5	1/5	1	1/9
E_{13}	7	9	5	9	9	9	9	9	7	5	7	9	1

Figur 12: Jämförelser av de 13 hoten avseende frekvens med 9-gradig skala. I denna bedömningsmatris används vikter från den klassiska AHP-skalan.

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}	E_{13}
E_1	1	1/7	1	1/7	1/7	1	1/7	1/7	7	1	1/7	1/7	7
E_2	7	1	7	1	7	7	7	1/7	7	7	1	1	7
E_3	1	1/7	1	1/7	1/7	1/7	1	1/7	1	1	1/7	1/7	7
E_4	7	1	7	1	1	7	7	1	7	7	1	1	7
E_5	7	1/7	7	1	1	1	1	1/7	7	7	1	1/7	7
E_6	1	1/7	7	1/7	1	1	1	1/7	7	7	1	1/7	7
E_7	7	1/7	1	1/7	1	1	1	1/7	7	7	1	1/7	7
E_8	7	7	7	1	7	7	7	1	7	7	7	1	7
E_9	1/7	1/7	1	1/7	1/7	1/7	1/7	1/7	1	1/7	1/7	1/7	7
E_{10}	1	1/7	1	1/7	1/7	1/7	1/7	1/7	7	1	1/7	1/7	7
E_{11}	7	1	7	1	1	1	1	1/7	7	7	1	1/7	7
E_{12}	7	1	7	1	7	7	7	1	7	7	7	1	7
E_{13}	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1

Figur 13: Jämförelser av de 13 hoten avseende konsekvens med 3-gradig skala. I denna bedömningsmatris används vikten 7 för att representera *mer än* på skalan och vikten 1/7 för att representera *mindre än*.

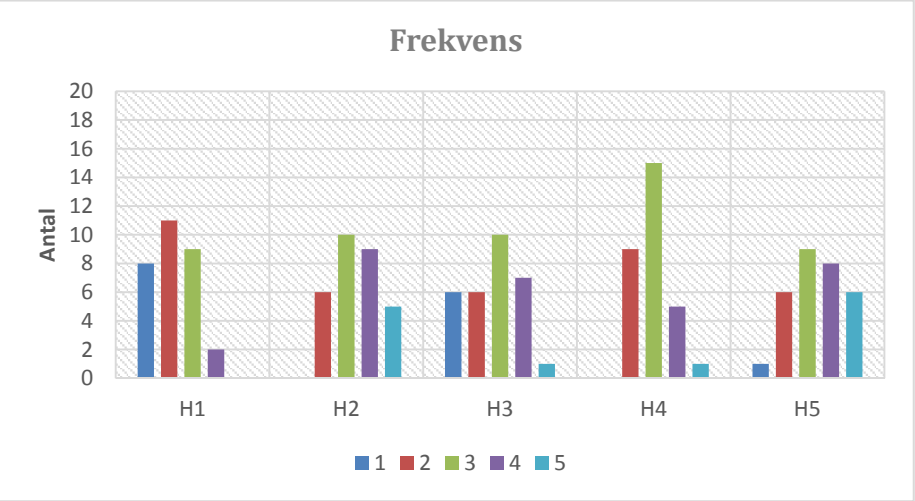
	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}	E_{13}
E_1	1	7	1	7	7	7	7	7	1	1	7	7	1/7
E_2	1/7	1	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7
E_3	1	7	1	7	7	7	7	7	1	7	7	7	1/7
E_4	1/7	7	1/7	1	1	1	1/7	1/7	1/7	1/7	1/7	1	1/7
E_5	1/7	7	1/7	1	1	1	1	1	1/7	1	1/7	7	1/7
E_6	1/7	7	1/7	1	1	1	1	1	1/7	1/7	1/7	7	1/7
E_7	1/7	7	1/7	7	1	1	1	1	1/7	1/7	1	7	1/7
E_8	1/7	7	1/7	7	1	1	1	1	1/7	1/7	1/7	7	1/7
E_9	1	7	1	7	7	7	7	7	1	7	7	7	1/7
E_{10}	1	7	1/7	7	1	7	7	7	1/7	1	1	7	1/7
E_{11}	1/7	7	1/7	7	7	7	1	7	1/7	1	1	7	1/7
E_{12}	1/7	7	1/7	1	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1	1/7
E_{13}	7	7	7	7	7	7	7	7	7	7	7	7	1

Figur 14: Jämförelser av de 13 hoten avseende frekvens med 3-gradig skala. I denna bedömningsmatris används vikten 7 för att representera *mer än* på skalan och vikten 1/7 för att representera *mindre än*.

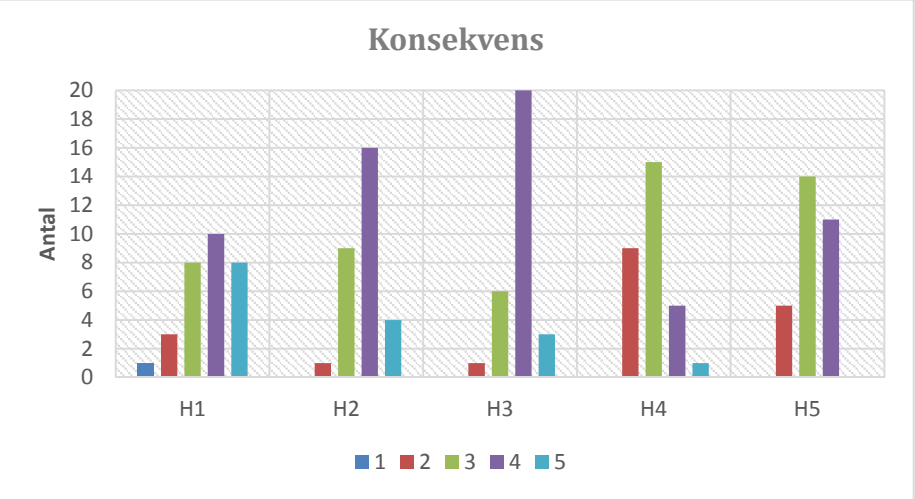
Bilaga 3 Statistik från experiment

I denna bilaga återfinns fördelningen av svar från det experiment som beskrivs i Kapitel 5.

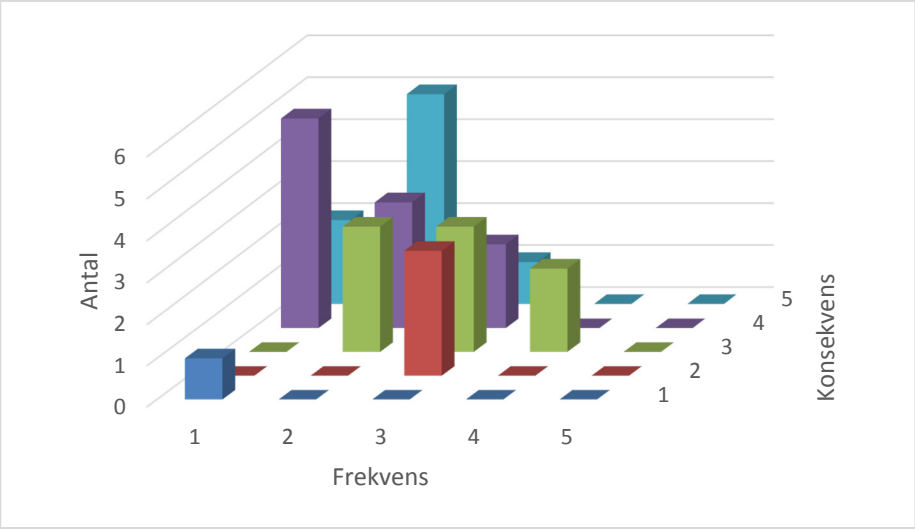
Absoluta bedömningar



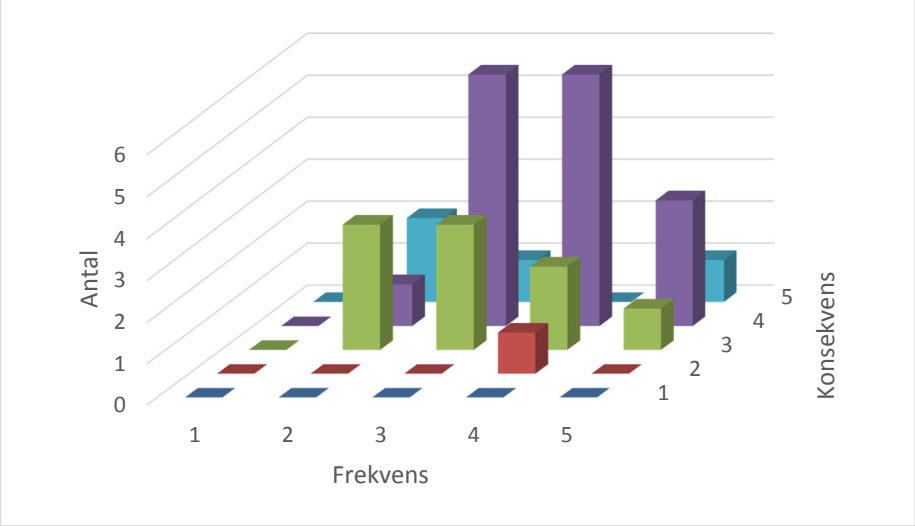
Figur 15: Fördelningen av svar avseende frekvens för de fem hoten.



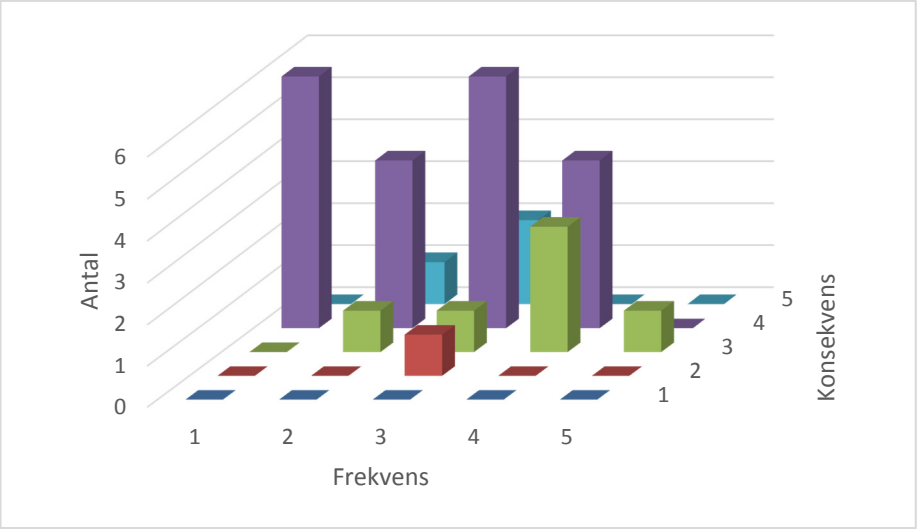
Figur 16: Fördelningen av svar avseende konsekvens för de fem hoten.



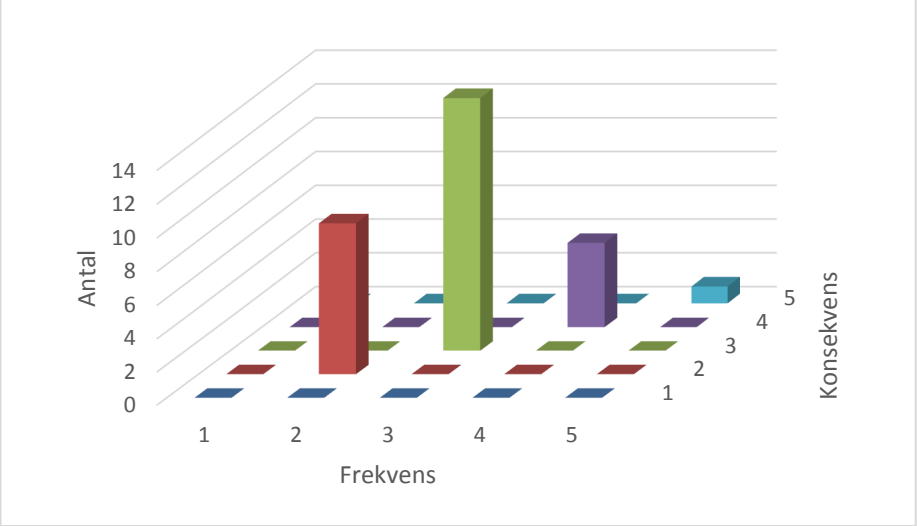
Figur 17: Fördelningen av svar avseende frekvens och konsekvens för hot 1. Staplarnas färg utgår från konsekvensvärdet.



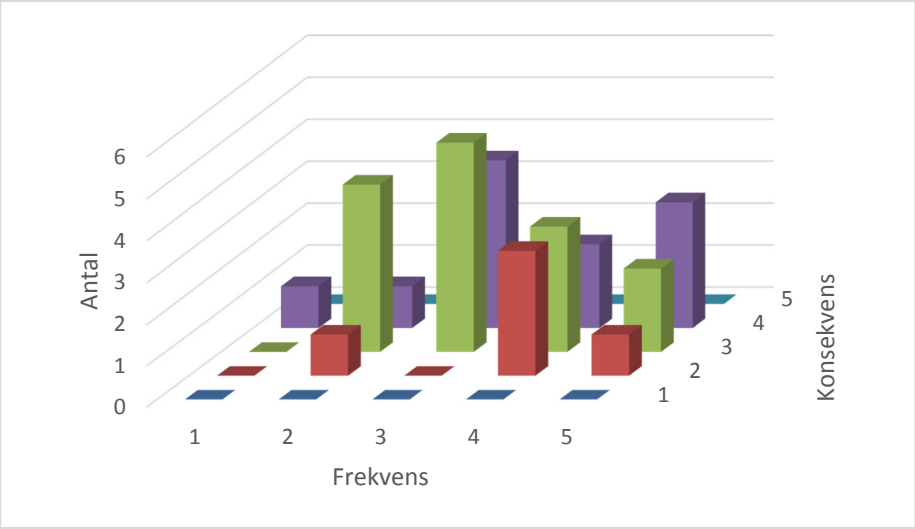
Figur 18: Fördelningen av svar avseende frekvens och konsekvens för hot 2. Staplarnas färg utgår från konsekvensvärdet.



Figur 19: Fördelningen av svar avseende frekvens och konsekvens för hot 3. Staplarnas färg utgår från konsekvensvärdet.

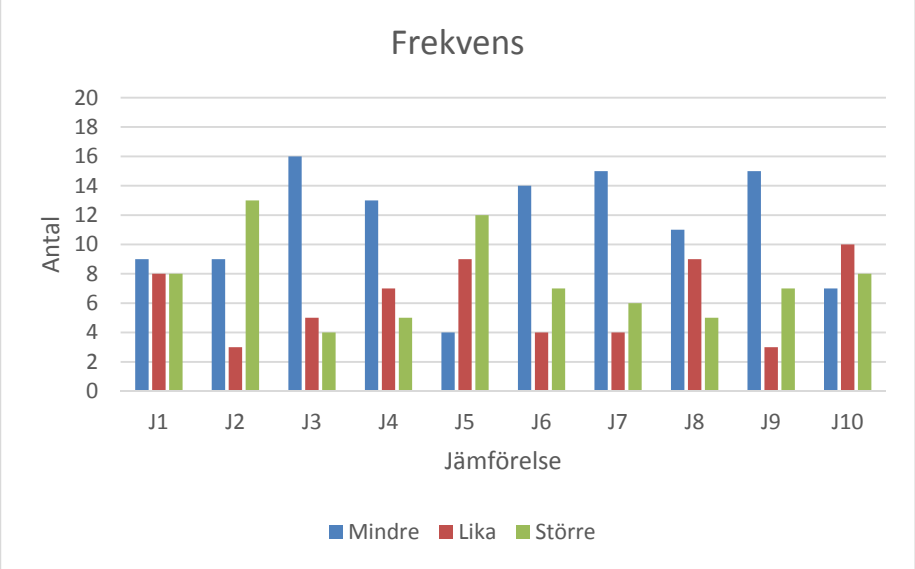


Figur 20: Fördelningen av svar avseende frekvens och konsekvens för hot 4. Staplarnas färg utgår från konsekvensvärdet.

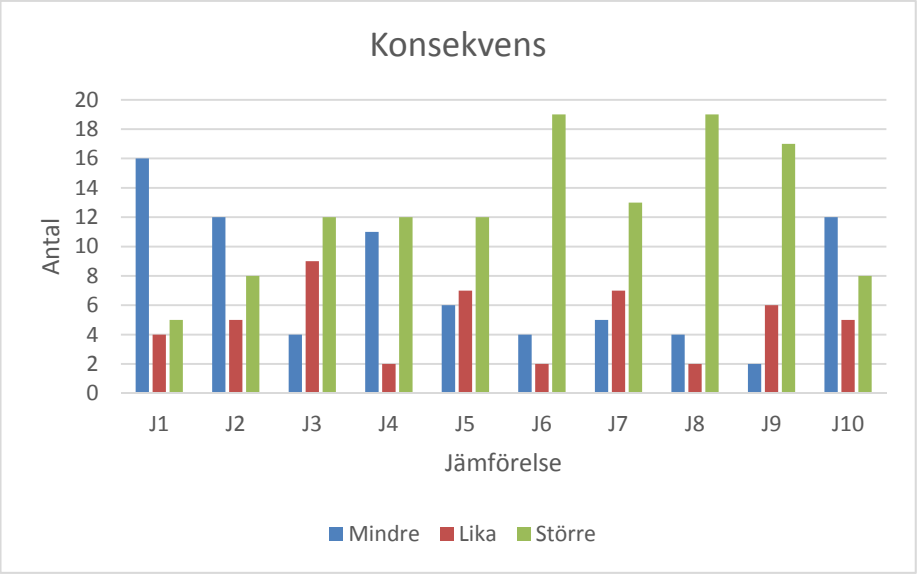


Figur 21: Fördelningen av svar avseende frekvens och konsekvens för hot 5. Staplarnas färg utgår från konsekvensvärdet.

Jämförelsebaserade bedömningar



Figur 22: Fördelning av svar för de tio jämförelserna avseende frekvens. Skalan (mindre, lika, större) utgår från *hot vänster* i jämförelse med *hot höger*, enligt Tabell 18.



Figur 23: Fördelning av svar för de tio jämförelserna avseende frekvens. Skalan (mindre, lika, större) utgår från *hot vänster* i jämförelse med *hot höger*, enligt Tabell 18.

Tabell 18: De tio jämförelserna som ingick i experimentet. Hot vänster jämförs med hot höger, vilket innebär att exempelvis J1 ger att hot 1 ska jämföras med hot 2. Antalet fall där frekvensen för hot 1 bedöms vara mindre än för hot 2

Jämförelse	**Hot vänster**	**Hot höger**
J1	Hot 1: Ransomware	Hot 2: Spear fishing
J2	Hot 1: Ransomware	Hot 3: Hårdvara
J3	Hot 1: Ransomware	Hot 4: DDoS-attack
J4	Hot 1: Ransomware	Hot 5: Skadlig kod
J5	Hot 2: Spear fishing	Hot 3: Hårdvara
J6	Hot 2: Spear fishing	Hot 4: DDoS-attack
J7	Hot 2: Spear fishing	Hot 5: Skadlig kod
J8	Hot 3: Hårdvara	Hot 4: DDoS-attack
J9	Hot 3: Hårdvara	Hot 5: Skadlig kod
J10	Hot 4: DDoS-attack	Hot 5: Skadlig kod

Att beskriva hot mot IT-system och bedöma dessa hots sannolikhet och konsekvens är en krävande uppgift. Under 2016 har FoT-projektet Bedömning och analys av IT-system genomfört studier avseende svårigheter med att beskriva och bedöma hot mot IT-system. Inom studierna genomfördes två experiment med fokus på aspekter som kunde tänkas påverka samstämmigheten mellan olika personers bedömning av sannolikhet och konsekvens för hot.

Det första experimentet fokuserades på om hotbeskrivningens utformning påverkar bedömningarna av sannolikhet och konsekvens. Hotbeskrivningarna utgjordes av strukturerade tabeller respektive löpande text. Inga statistiskt signifikanta skillnader kunde dock ses i resultatet från experimentet.

Det andra experimentet byggde vidare på ett experiment som genomfördes i projektet under 2015. Experimentet genomfördes för att identifiera eventuella skillnader i samstämmighet när bedömningar görs enligt metoden för säkerhetsanalys (H Säk Grunder) respektive en metod baserad på jämförelsebaserade bedömningar. När jämförelsebaserade bedömningarna genomfördes av experter, med såväl god kunskap inom området informationssäkerhet som erfarenhet av att bedöma informationssäkerhetsrisker, uppnåddes statistiskt signifikant högre samstämmighet för bedömningarna av både sannolikhet och konsekvens.

Rapportens avslutas med slutsatser att beakta vid Försvarsmaktens fortsatta arbete med säkerhetsanalyser. Dessa slutsatser relaterar till utformningen av hotbeskrivningar och förutsättningar för bedömning av hots sannolikhet och konsekvens.

FOI

ISSN1650-1942

www.foi.se