Försöksverksamhet inom logganalys för cybersäkerhet

PATRIK LIF, HANNES HOLM, TEODOR SOMMESTAD,
MAGDALENA GRANÅSEN & ERIK WESTRING

FOI-R--4328--SE
ISSN 1650-1942 November 2016

Patrik Lif, Hannes Holm, Teodor Sommestad,
Magdalena Granåsen & Erik Westring

Försöksverksamhet inom logganalys för cybersäkerhet

Titel	Försöksverksamhet inom logganalys för cybersäkerhet
Title	Exercise activities in log analysis for cybersecurity
Rapportnr/Report no	FOI-R--4328--SE
Månad/Month	November
Utgivningsår/Year	2016
Antal sidor/Pages	48
ISSN	1650-1942
Kund/Customer	FM
Forskningsområde	4. Informationssäkerhet och kommunikation
FoT-område	Ledning och MSI
Projektnr/Project no	E72629
Godkänd av/Approved by	Christian Jönsson
Ansvarig avdelning	Ledningssystem

Sammanfattning

Rapporten beskriver en studie inom cybersäkerhet med fokus på logganalys. Deltagarna tränade logganalys i rollerna chef, spanare och analytiker, både enskilt och i grupp. Studien nyttjades även för att studera cybersituationsmedvetande (CSA) och lärandeeffekter av återkoppling. Studien utfördes i FOI:s cyber range CRATE och innehöll webb-, e-post- och filservrar, nätverksutrustning och drygt 200 kontorsdatorer där även användaraktivitet simulerades. Deltagarnas uppgift var att med stöd av tillgänglig övervakningsutrustning upptäcka och analysera olika attacker. Första dagen arbetade deltagarna i grupp. Datainsamling genomfördes i syfte att studera arbetsmetoder samt värdera mätmetodik för CSA som utvecklades utifrån en litteraturstudie och intervjuer med logganalytiker inför försöket. Datainsamlingen inkluderade även prestationsmätning, kommunikationsanalys, hierarkisk uppgiftsanalys och subjektiva skattningar av lärandeeffekt, nytta och svårighetsgrad. Resultaten kring arbetsmetoder gav bra input till den deltagande organisationen avseende rollfördelning och arbetsbelastning. CSA-formuläret mottogs positivt och enbart mindre modifieringar har gjorts inför framtida studier. Andra dagen arbetade deltagarna enskilt med fyra uppgifter där hälften av deltagarna fick återkoppling. Datainsamlingen bestod av prestationsmätning och subjektiva skattningar, och syftade främst till att undersöka den eventuella nyttan med återkoppling. Resultaten indikerar att återkoppling har en betydelse för prestationen på efterföljande uppgift, men visade golveffekt till följd av uppgifternas svårighetsgrad och att de tillgängliga logganalysverktygen inte var optimalt konfigurerade. Tredje dagen bestod av utvärdering som visade att försöksupplägget och datainsamlingsverktygen i stort fungerar väl. Huvudsakliga åtgärder för kommande försök omfattar att förenkla komplexiteten i miljön och se över konfiguration av logganalysverktygen. Deltagarna visade stor uppskattning avseende genomförandet av de tre dagarna och det finns en bra grund att stå på inför kommande verksamhet.

Nyckelord: Logganalys, cyber situationsmedvetande, CRATE

Summary

The current report describes a cyber security log analysis exercise. The participants practiced log analysis in the roles of manager, scout and analyst, and practice individually as well as in teams. Furthermore, the exercise studied cyber situation awareness (CSA) and effects of receiving feedback. The exercise was conducted in FOI:s cyber range CRATE with web-, email-, and file servers, network equipment and more than 200 computer clients. The participants' task was to identify and analyze various attacks, including network scans from inside and outside the network, password guesses on network services, infected USB sticks and overload attacks. The first day of the exercise, the participants worked as a team and data was collected in order to study working methods and validate a measurement technique for CSA. The measurement instrument was developed by literature analysis as well as interviews with log analysts before the exercise. Also, data collection included performance measures, communication analysis, hierarchical task analysis and subjective assessments of learning effects, usefulness and complexity of the given tasks. The results concerning working methods gave a good input to the participating organization regarding division of labor and workload. The CSA instrument was well received and only minor changes are needed for future studies. During the second day, the participants worked individually, solving three different tasks. After each task, half of the participants received oral feedback. The data collection aimed to study the possible benefits of feedback. The results indicate that feedback had a positive effect on the accomplishment of the subsequent task but a drawback was a floor effect was obtained due to the level of difficulty of the tasks and that the available monitoring tools were not optimally configured for the tasks. Overall, the exercise setup and data collection tools proved successful. For future exercises, the complexity in the exercise environment should be reduced and configuration of monitoring tools revised. The exercise was very appreciated by the participants and forms a good basis for future project activities.

Keywords: Log analysts, cyber situation awareness, CRATE

Innehållsförteckning

1 Inledning

Denna rapport är framtagen som en del av projektet Övning och experiment för operativ förmåga i cybermiljön (ÖvExCy). Projektet ÖvExCy fokuserar på förmågan att analysera cybersäkerhetloggar och ska bland annat identifiera lämpliga utbildningsscenarier och experiment som är passande för att utveckla, pröva och utvärdera denna förmåga. Rapporten beskriver ett antal mätningar som gjorts kopplat till detta i samband med att ett antal logganalytiker tränades i FOI:s övnings- och experimentanläggning CRATE under november 2016. Denna inledning ger en kort bakgrund till området och därefter beskrivs genomförd studie övergripande i kapitel 2. De två genomförda försöken beskrivs mer detaljerat i kapitel 3 och 4. Slutligen sammanfattats studien i kapitel 5 och slutsatser presenteras.

1.1 Analys av cybersäkerhetsloggar

Begreppet cyberförsvar saknar en tydlig och etablerad definition, men det finns flera idéer om vad cybersäkerhetsarbete består av. En av de mer omfattande sammanställningarna heter National Cybersecurity Workforce Framework (NICE, 2015) och har producerats av det amerikanska institutet National Institute of Standards and Technology (NIST). Ramverket påvisar att cybersäkerhetsarbete innefattar många olika uppgifter och att det kräver ett stort antal olika kompetenser. Denna rapport fokuserar på logganalysförmåga, vilket enligt NIST:s ramverk främst innefattar arbete inom specialområdet nätverksanalys (eng. Computer Network Defense Analysis) (NICE, 2015). Enligt detta ramverk består nätverksanalys av 25 uppgifter, där flera handlar om det som kallas intrångsdetektion – att identifiera elakartad aktivitet utifrån systemloggar. Intrångsdetektion är ett aktivt forskningsområde, där över tusen nya artiklar presenteras årligen i vetenskapliga forum. Andra uppgifter inom intrångsdetektion handlar t.ex. om rapportering av incidenter och trendanalyser. Dessa är mer sällan studerade av forskare.

Det finns uppenbara kopplingar mellan de uppgifter som en logganalytiker utför och begreppet situationsmedvetande (eng. Situation Awareness - SA) (Endsley, 1995a, 1995b). Endsley (1995b) definierar situationsmedvetande på följande sätt:

> *”Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.”*

Trots det handlar merparten av forskningen inom området om rena teknikfrågor där system byggs för att upptäcka angrepp. Det är tydligt att metoder för att mäta cybersituationsmedvetande behöver utvecklas för logganalytiker.

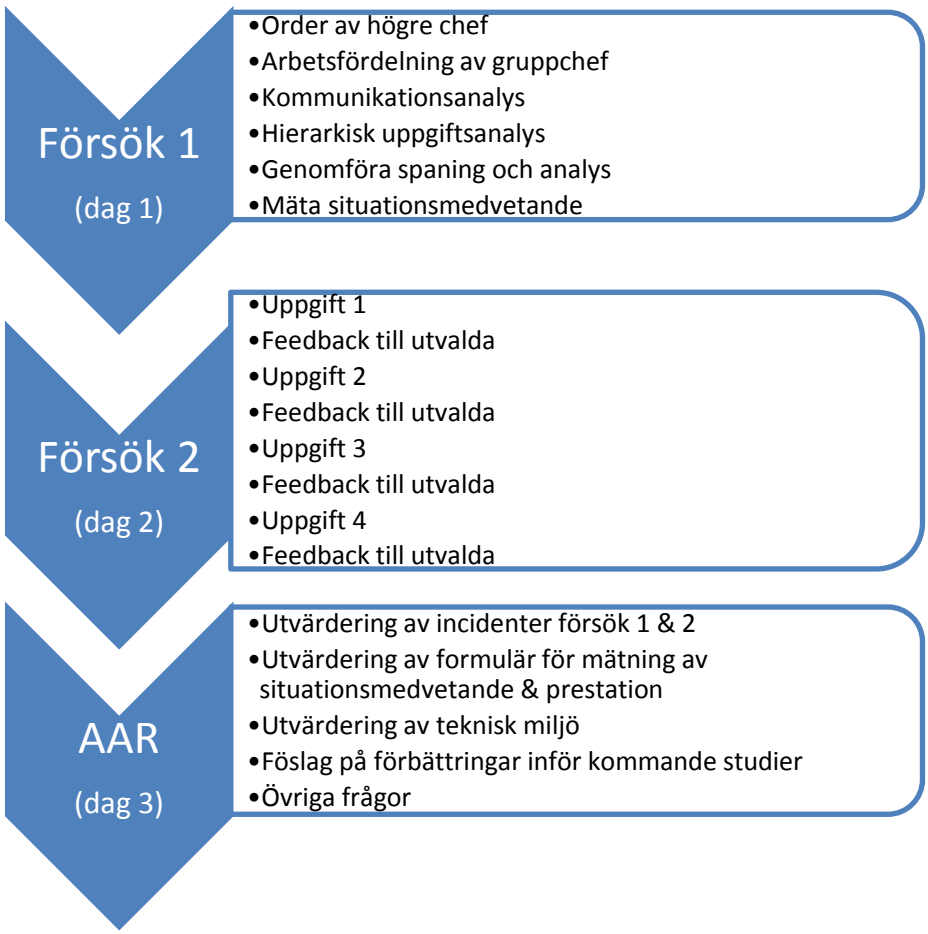
Fortsättningsvis benämns situationsmedvetande för SA och cybersituationsmedvetande för CSA.

1.2 Utveckling av logganalysförmåga

Projektet ÖvExCy försöker identifiera lämpliga utbildningsscenarier och försök för att utveckla logganalysförmåga samt identifiera hur övnings- och experimentanläggningar bör konstrueras för att kunna utveckla logganalysförmåga. Även om det finns goda erfarenheter både inom Sverige och internationellt av att öva logganalysförmåga i övnings- och experimentanläggningar finns endast begränsad forskning om bästa sättet att öva. Att öva i övnings- och experimentanläggning har vissa nackdelar jämfört med att öva i driftsatta system och lärande i det dagliga arbetet. Till exempel krävs det att resurser läggs på en övnings- och experimentanläggning samt att den simulering som sker i anläggningen inte stämmer med den verklighet som logganalytikern ska verka i. Men det finns också fördelar. Till exempel försvinner risken att misstag eller angrepp orsakar faktisk skada. Det gör det även möjligt att öva på analys av potentiella framtida hot och det är rättframt att avgöra ifall en analys eller situationsbedömning är korrekt. Det övergripande syftet med studien som presenteras i denna rapport var att öka deltagarnas förmåga till logganalys och att bedriva generell forskning kopplat till logganalysförmåga. Studien ska betraktas som ett första försök att undersöka nyttan att träna logganalytiker på analys av potentiella framtida hot. De två genomförda försöken utvärderade framförallt framtagen metodik för att mäta situationsmedvetande och effekten av att de tränade får veta det rätta svaret efter att uppgiften genomförts (återkoppling).

2 Övergripande beskrivning av genomförd studie

I denna studie genomförde logganalytiker träning i FOI:s övnings- och experimentanläggning kallad CRATE. Studien pågick i tre dagar och innehöll två försök (dag 1-2) och en utvärdering (eng. after-action review - AAR) (dag 3). AAR beskrivs mer detaljerat i Bilaga 11. Utvärderingen ska betraktas som en del av de försök som genomfördes de första två dagarna och beskrivs därför fortsättningsvis som en del de båda försöken. För övergripande struktur av studien se Figur 1.



Figur 1. Övergripande metodupplägg av studie.

Studien genomfördes med fem experter inom logganalys. I det första försöket hade deltagarna någon av rollerna gruppchef, spanare eller analytiker. Gruppchefen ledde arbetet och förde krigsdagbok; spanarna övervakade systemet och försökte identifiera avvikelser, oregelbundna mönster eller andra tecken på hot; analytikerna omhändertog de avvikelser som spanarna upptäckt och bedömt som prioriterade för att analysera dem ytterligare och avgöra om det var nödvändigt att vidta någon form av åtgärd. I det andra försöket arbetade deltagarna individuellt och agerade både som spanare och analytiker.

2.1 Övergripande metod

I detta avsnitt beskrivs den övergripande metod som användes i studien (försök 1 och 2). De delar i metoden som är unika för respektive försök beskrivs i kapitel 3 och 4.

2.1.1 Cybermiljön

Försöken genomfördes i en virtuell miljö skapad i CRATE. Denna miljö innehöll ett antal webbservrar, e-postservrar, filservrar, nätverksutrustning och drygt 200 kontorsdatorer med operativsystemet Windows 7 och typiska kontorsapplikationer. I dessa kontorsdatorer utförde ett program implementerat med programmeringsspråket AutoIt knapptryckningar för att simulera typiska kontorsanvändare. De simulerade användarna skickade e-post, besökte webbsidor och öppnade filer med samma frekvens som användare av riktiga kontorsdatorer på FOI och två andra arbetsplatser gjort samma veckodag 2012. Utanför den övervakade miljön fanns en extern miljö som var att betrakta som internet, med en typisk hotbild.

2.1.2 Övervakningsutrustning

Var och en av deltagarna utrustades med två laptops med 15-tums skärm. Dessa användes för att på olika sätt interagera med den övervakningsutrustning som fanns i cybermiljön. Nätverkstrafiken övervakades på ett tiotal knutpunkter och analyserades med logganalysverktyg som Wireshark, ELK och Snort. Systemhändelser i form av Windows Event Log samlades till ett centralt system och överfördes till analysverktyget ArcSight. Deltagarna hade på förhand försetts med en beskrivning miljön.

2.2 Försök 1 – arbetsmetoder

Syftet med försök 1 var att undersöka arbetsmetoder och utvärdera framtagen metodik för att mäta CSA. Deltagarna arbetade i grupp för att försöka upptäcka

och analysera hot. De vetenskapliga frågeställningar som undersöktes var följande:

- Hur upplever logganalytiker framtagen metodik för att mäta CSA?
- Vilka arbetsuppgifter utförs av logganalytiker i rollerna som chef, spanare och analytiker (utvärderas med hierarkisk uppgiftsanalys - HTA)?
- Vem kommunicerar med vem (högre chef, gruppchef, spanare, analytiker och driftstekniker)? Kommunikation syftar på muntlig kommunikation.
- Vilka lärdomar kan dras från detta försök?

2.2.1 Vetenskaplig bakgrund

För att kartlägga och förstå logganalytikers arbete användes hierarkisk uppgiftsanalys och kommunikationsanalys, och för att mäta deras förmåga utvecklades och utvärderades en metod för att mäta CSA.

SA innehåller tre nivåer: upptäcka, förstå och predicera framtid (Endsley, 1995b). Inom domäner som flygledning och för piloter finns gott om litteratur som beskriver begreppet SA och hur mätningar av det kan genomföras (Endsley, 1995a; Endsley et al., 2000; Endsley, Selcon, Hardiman, & Croft, 1998). Metoder för att mäta SA är t.ex. Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988), Situation Awareness Control Room Inventory (SACRI) (Hogg, Follesø, Torralba, & Volden, 1994), Situation Awareness Rating Teqchnique (SART) (Taylor, 1990) och modifierad Bedfordskala (VINTHEC, 1997). En av de mer etablerade metoderna är SAGAT (Endsley, 1988; Endsley & Garland, 2000) som bygger på att avbryta personer i deras arbete och be dem svara på frågor som har att göra med den nuvarande situationen. För att svaren på frågorna ska hänga ihop med personernas prestation behöver de handla om sådant i situationen som är relevant att vara medveten om. Det är inte enkelt att göra en distinktion mellan en persons SA och personens prestation i uppgiften som ska utföras. Till exempel kan SA mätas genom att fråga vilka händelser som innebär potentiella hot, och i logganalysuppgiften består det typiskt av att bekräfta att en händelse utgör ett hot samt ta reda på mer detaljer om det. Mätningar av prestation kan således överlappa mätningar av SA.

Efter en genomgång av publicerad forskning konstaterade Franke och Brynielsson (2014) att det finns lite empirisk forskning kopplad till SA i cybermiljö och att mer sådan forskning behövs. Det skiljer cybersäkerhetsområdet från andra domäner där det finns mer forskning om SA. I litteratur kopplat till cybersäkerhet och cyberförsvar är begrepp mer otydligt definierade (Barford et al., 2010; Franke & Brynielsson, 2014) och den litteratur som beskriver hur situationsmedvetande bör mätas är ytterst begränsad

(Brynielsson, Franke, & Varga, 2016). I denna rapport görs ett försök att identifiera frågor som kan användas för att mäta logganalytikers situationsmedvetande. Litteraturstudier (Lif, Thorstensson, & Sommestad, 2015) och en målinriktad uppgiftsanalys (Endsley & Garland, 2000) gjord tillsammans med professionella logganalytiker inom cybersäkerhet låg till grund för de frågor som utvecklats och används i detta försök. I försök 1 undersöktes frågornas relevans och sattes också i relation till mätningar av prestation, som utgjordes av de incidentrapporter som producerades under försök 1.

Kommunikationsanalys (Have, 2007) med förvalda kategorier användes för att undersöka hur deltagarna interagerade med varandra (Bilaga 3) och hierarkisk uppgiftsanalys (Annett, 2003; Stanton, 2006) genomfördes av en expert inom incidenthantering utifrån en mall (Bilaga 10) för att kartlägga vilka aktiviteter som deltagarna utförde i de tre rollerna gruppchef, spanare och analytiker.

2.3 Försök 2 – återkoppling

Syftet med försök 2 var att undersöka om logganalytikers förmåga förbättrades med återkoppling avseende hur respektive attack genomförts jämfört med då återkoppling inte gavs. Detta var en individuell uppgift där deltagarna skulle upptäcka och analysera hot på motsvarande sätt som i första försöket. De vetenskapliga frågeställningar som undersöktes var följande:

- Förbättras logganalytikers prestation om de får återkoppling under träning jämfört med när återkoppling saknas?
- Hur upplever logganalytikerna formuläret för mätning av prestation?
- Vilka lärdomar kan dras från detta försök?

2.3.1 Vetenskaplig bakgrund

Återkoppling kan delas in i två huvudsakliga delar: resultat och process (Hoffman et al., 2014). Resultat syftar på om deltagarna har gjort rätt eller fel medan process syftar på hur de arbetar. Båda delarna anses viktiga för att uppnå god skicklighet (Hoffman et al., 2014). Kontinuerlig återkoppling kan vara positivt för vissa uppgifter (Frederiksen & White, 1989; Pellegrino, Chudowsky, & Glaser, 2001) och kan ske genom att en person eller ett användargränssnitt i en dator ger återkoppling. Forskning har också visat att prestationen även kan försämras på grund av återkoppling under vissa betingelser (Kluger & DeNisi, 1996). Bland annat tycks tidpunkten för när återkoppling ges påverka inlärningen (Hoffman et al., 2014). Direkt återkoppling i samband med utförande av en uppgift kan till och med vara negativt för långsiktig inlärning. Kanske för att mycket återkoppling i direkt anslutning till uppgiften förhindrar deltagarnas

reflektion och möjlighet att organisera informationen på ett strukturerat sätt (Hoffman et al., 2014).

Den enda forskning som identifierats kopplat till logganalysförmåga och återkoppling har undersökt hur detaljerad återkoppling kontra översiktlig återkoppling påverkar prestation i en syntetisk uppgift (Ben-Asher & Gonzalez, 2015). Studien visar att detaljerad återkoppling på analyser av tidigare angrepp ökade förmågan att upptäcka kommande angrepp av liknande typ betydligt bättre än aggregerad översiktlig återkoppling. Det tycks alltså finnas en poäng med att träna i övnings- och experimentanläggningar där detaljerad återkoppling kan ges till de tränade. Försök 2 beskriver ytterligare test av återkopplingens påverkan på prestation i logganalysarbete.

3 Försök 1: undersökning av arbetsmetoder

Syftet med detta försök var att undersöka arbetsmetoder och utvärdera framtagen metodik för att mäta CSA.

3.1 Metod

Logganalytikernas huvuduppgift var att övervaka, analysera och fatta beslut om lämpliga åtgärder för upptäckt av hot/attacker (Cichonski, Millar, Grance, & Scarfone, 2012; ENISA, 2010; Sommestad & Hunstad, 2013). Försöket genomfördes med deltagare som arbetade i rollerna chef, spanare och analytiker. Rollerna alternerades från för- till eftermiddag. Det fanns även en roll som drifttekniker som deltagarna kunde vända sig till för att fråga om systeminformation eller delge information för att åtgärda uppkomna problem i systemet. Rollen som drifttekniker spelades av en person i spelledningen. Data samlades in genom användande av framtagna formulär för CSA (Bilaga 1 och Bilaga 2), registrering av kommunikation (Bilaga 3 och Bilaga 8) och hierarkisk uppgiftsanalys (Bilaga 10) genomfördes med en expert inom logganalys som innan försöket startade hade fått instruktioner och under försöket fick stöd av forskare när det behövdes.

3.1.1 Attacker/incidenter

Angreppen utfördes med verktyget SVED (eng. Scanning, Vulnerabilities, Exploits and Detection) (Holm & Sommestad, 2016) och bestod bland annat av: nätverksavsökningar inifrån och utifrån nätet, lösenordsgissningar på nätverkstjänster, infekterade USB-stickor, överbelastningsattacker och e-mail med skadliga bilagor. Dessa utfördes så att belastningen för logganalytikerna skulle vara jämn och hög under hela dagen.

3.1.2 Formulär för mätning av CSA

Framtagandet av formulär för mätning av CSA gjordes utifrån kunskap från litteratur (Lif et al., 2015), en målinriktad uppgiftsanalys (Endsley & Garland, 2000) gjord tillsammans med professionella logganalytiker inom cybersäkerhet och utifrån kunskaper och erfarenheter från deltagarna i projektet *Övning och experiment för operativ förmåga i cybermiljön*. Ett formulär framtogs för spanare (Bilaga 1) och ett formulär framtogs för analytiker (Bilaga 2) eftersom arbetsuppgifterna i de både rollerna är olika.

3.1.3 Kommunikationsanalys

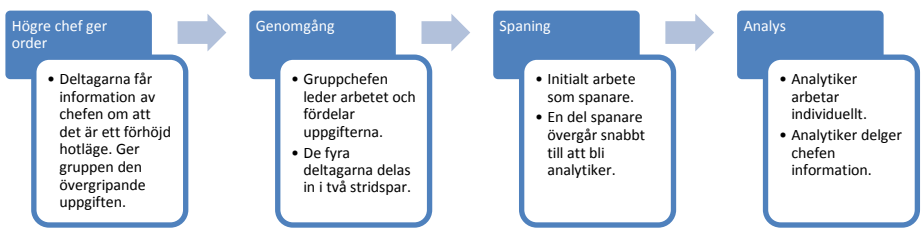
En forskare följde kommunikationen i gruppen genom att notera vem som samtalade med vem, enligt principen sändare och mottagare av information. Dessutom så klassificerades all information enligt: ämnesdiskussion, metoddiskussion, överlämning, beslut, tekniska problem, oenighet, utbildningsbrist, inspel/information och annat. Dessa kategorier bedömdes inkludera det innehåll som gruppen diskuterade. För att underlätta datainsamling och analys användes ett excelblad med macro för tidstämpling, sändare, mottagare, kategorier och ett fält för att notera kommunikationens innehåll.

3.1.4 Enkät

En enkät användes efter genförts försök där deltagarna gavs möjlighet att framföra sina subjektiva åsikter avseende bland annat hur mycket de upplevde att de lärt sig, vilken nytta de hade av det de lärt sig, upplevd svårighetsgrad och om de upplevde att försöket var verklighetstroget. Även en avslutande enkät användes (Bilaga 4). I den avslutande enkäten fanns även öppna frågor där deltagarna kunde beskriva positiva och negativa erfarenheter från försöket.

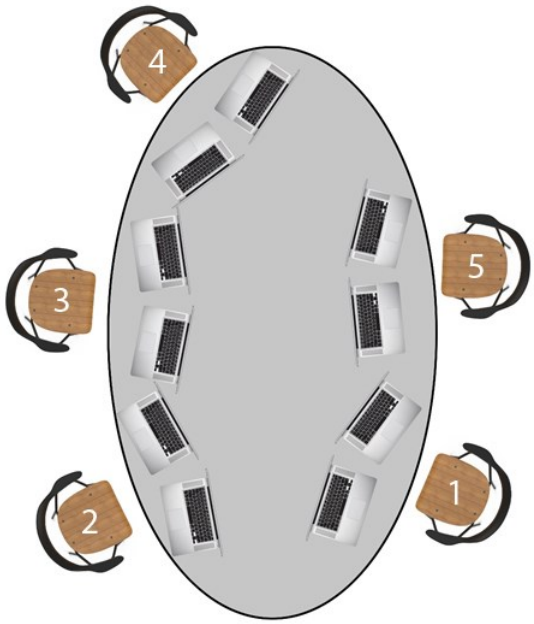
3.1.5 Genomförande

Studien startade med att fem deltagare samlades för genomgång av högre chef (ordinarie högre chef från deras ordinarie verksamhet), som förklarade att det fanns ett förhöjt hotläge avseende cyberattacker. Högre chef förklarade vidare att den ordinarie gruppchefen för incidenthanteringsgruppen inte hade möjlighet att delta i arbetet och utsåg därför en av deltagarna till ny gruppchef för att leda verksamheten och fördela arbetet mellan resterande fyra deltagare som delades in i två stridspar. Deltagarna gavs arbetsuppgifter som spanare eller analytiker. När spanare upptäckte anomalier eller hot som ansågs prioriterade så kunde extra resurser läggas på analysuppgifter. Hur rollfördelningen skiftade över tid var beroende av antalet attacker, hur allvarliga hoten bedömdes vara och gruppchefens beslut. Proceduren beskrivs i Figur 2.



Figur 2. Övergripande procedur för arbetet med information från högre chef, gemensam genomgång som leds av gruppchefen, spaning och fördjupad analys.

Deltagarna placerades runt ett bord på bestämda platser med två bärbara datorer på varje plats enligt Figur 3. På förmiddagen agerade deltagaren som satt på plats tre som gruppchef och på eftermiddagen agerade deltagaren som satt på plats ett som gruppchef.



Figur 3. Deltagarnas placering runt bordet med två bärbara datorer på varje plats.

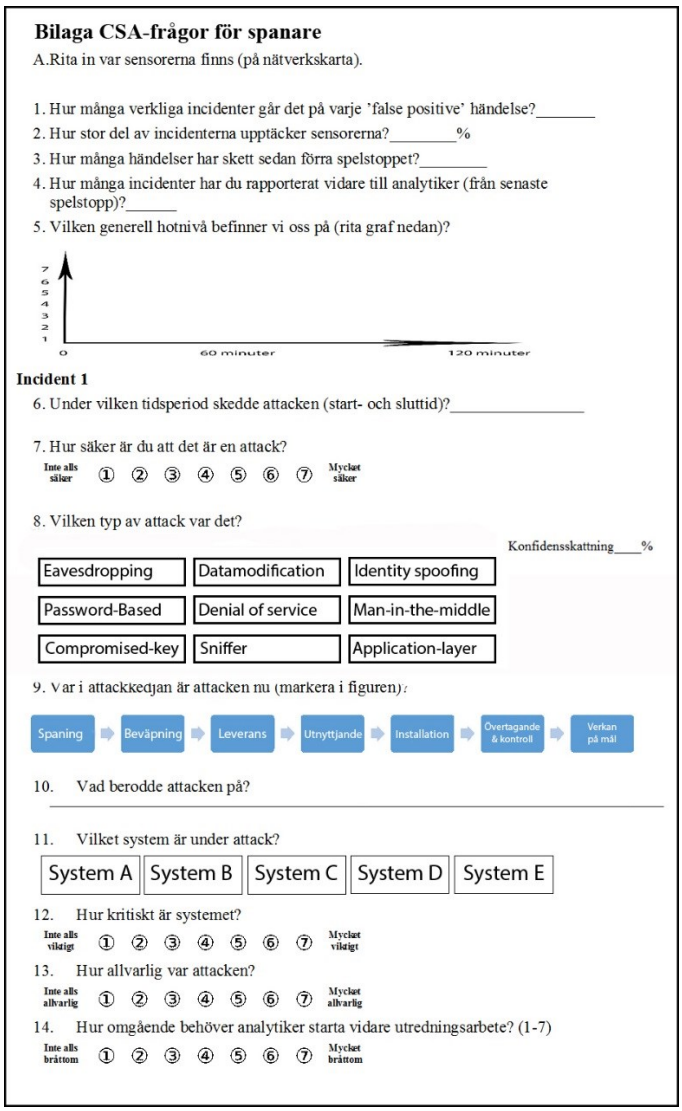
3.2 Resultat

Nedan redovisas resultaten kopplade till mätning av CSA, den hierarkiska uppgiftsanalysen och kommunikationsanalysen.

3.2.1 Mätning av CSA

De två formulären för att mäta logganalytikernas CSA utvärderades genom AAR (dag 3) där innehållet diskuterades i grupp. Formulären (Bilaga 1) mottogs mycket positivt och ansågs mäta CSA på ett tillfredställande sätt. Deltagarna förslog förändringar för att mäta CSA för spanare respektive analytiker. Deltagarna fick skriva ner förbättringsförslag på hur CSA bättre kunde mätas i rollen som spanare och analytiker. Dessutom diskuterades CSA utförligt under tredje dagens AAR där muntliga förslag framfördes. Deltagarnas skriftliga och muntliga kommentarer analyserades, sammanställdes och användes för att utveckla formulären. Formulär omarbetade enligt dessa förslag finns i Figur 4

och Figur 5; originalformulären finns i Bilaga 1och Bilaga 2. En viktig del är att figuren i formuläret ska ritas efter ett standardiserat system. Hur detta ska faktiskt bör genomföras är del av framtida arbete. Typ av attack (fråga 8 för spanare och fråga 2 för analytiker) är ett exempel på förvalda attacker som kan inkluderas, men ska inte betraktas som ett färdigt förslag på attacktyper.



Bilaga CSA-frågor för spanare

A.Rita in var sensorerna finns (på nätverkskarta).

1. Hur många verkliga incidenter går det på varje 'false positive' händelse?_____
2. Hur stor del av incidenterna upptäcker sensorerna?_____%
3. Hur många händelser har skett sedan förra spelstoppet?_____
4. Hur många incidenter har du rapporterat vidare till analytiker (från senaste spelstopp)?_____
5. Vilken generell hotnivå befinner vi oss på (rita graf nedan)?

Incident 1

6. Under vilken tidsperiod skedde attacken (start- och sluttid)?_____

7. Hur säker är du att det är en attack?

Inte alls säker ① ② ③ ④ ⑤ ⑥ ⑦ Mycket säker

8. Vilken typ av attack var det?

Konfidensskattning____%

Eavesdropping	Datamodification	Identity spoofing
Password-Based	Denial of service	Man-in-the-middle
Compromised-key	Sniffer	Application-layer

9. Var i attackkedjan är attacken nu (markera i figuren):

Spaning → Beväpning → Leverans → Utnyttjande → Installation → Övertagande & kontroll → Verkan på mål

10. Vad berodde attacken på?

11. Vilket system är under attack?

System A | System B | System C | System D | System E

12. Hur kritiskt är systemet?

Inte alls viktigt ① ② ③ ④ ⑤ ⑥ ⑦ Mycket viktigt

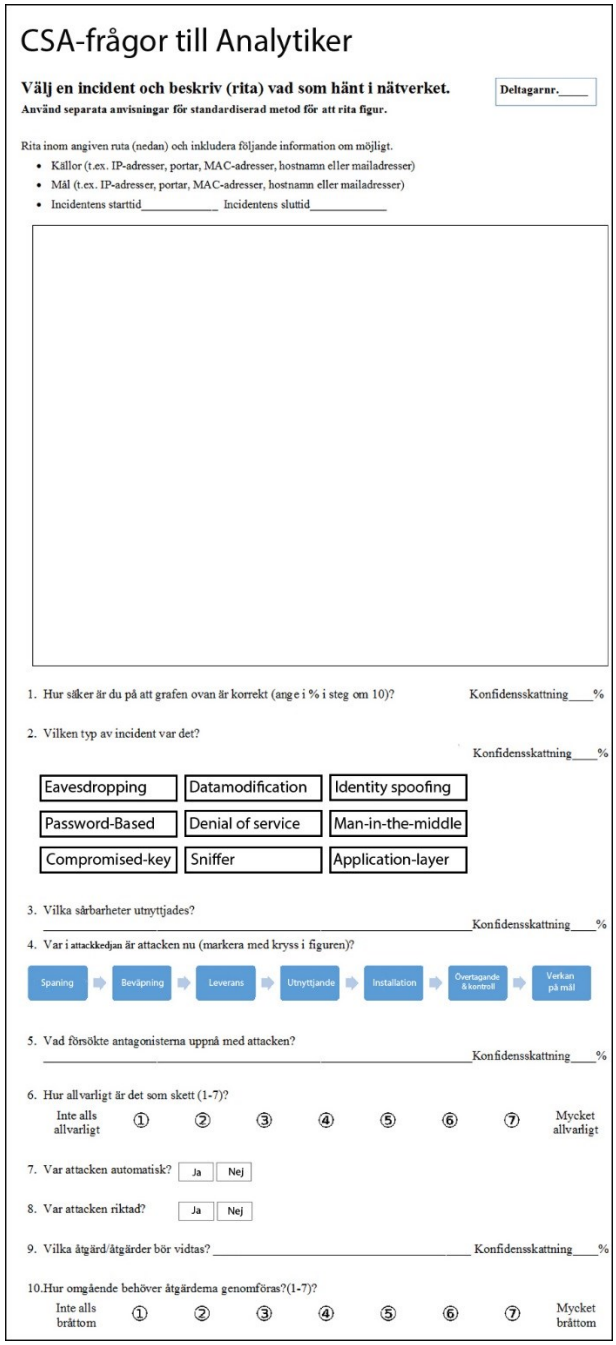
13. Hur allvarlig var attacken?

Inte alls allvarlig ① ② ③ ④ ⑤ ⑥ ⑦ Mycket allvarlig

14. Hur omgående behöver analytiker starta vidare utredningsarbete? (1-7)

Inte alls bråttom ① ② ③ ④ ⑤ ⑥ ⑦ Mycket bråttom

Figur 4. Formulär för mätning av spanares CSA efter att deltagarnas förbättringsförslag införts.



CSA-frågor till Analytiker

Välj en incident och beskriv (rita) vad som hänt i nätverket. Deltagarnr.____

Använd separata anvisningar för standardiserad metod för att rita figur.

Rita inom angiven ruta (nedan) och inkludera följande information om möjligt.

- Källor (t.ex. IP-adresser, portar, MAC-adresser, hostnamn eller mailadresser)
- Mål (t.ex. IP-adresser, portar, MAC-adresser, hostnamn eller mailadresser)
- Incidentens starttid_____ Incidentens sluttid_____

1. Hur säker är du på att grafen ovan är korrekt (ange i % i steg om 10)? Konfidensskattning____%

2. Vilken typ av incident var det?

Konfidensskattning____%

Eavesdropping	Datamodification	Identity spoofing
Password-Based	Denial of service	Man-in-the-middle
Compromised-key	Sniffer	Application-layer

3. Vilka sårbarheter utnyttjades?
_____Konfidensskattning____%

4. Var i attackkedjan är attacken nu (markera med kryss i figuren)?

Spaning → Beväpning → Leverans → Utnyttjande → Installation → Övertagande & kontroll → Verkan på mål

5. Vad försökte antagonisterna uppnå med attacken?
_____Konfidensskattning____%

6. Hur allvarligt är det som skett (1-7)?

Inte alls allvarligt ① ② ③ ④ ⑤ ⑥ ⑦ Mycket allvarligt

7. Var attacken automatisk? Ja Nej

8. Var attacken riktad? Ja Nej

9. Vilka åtgärd/åtgärder bör vidtas? _____ Konfidensskattning____%

10. Hur omgående behöver åtgärdema genomföras?(1-7)?

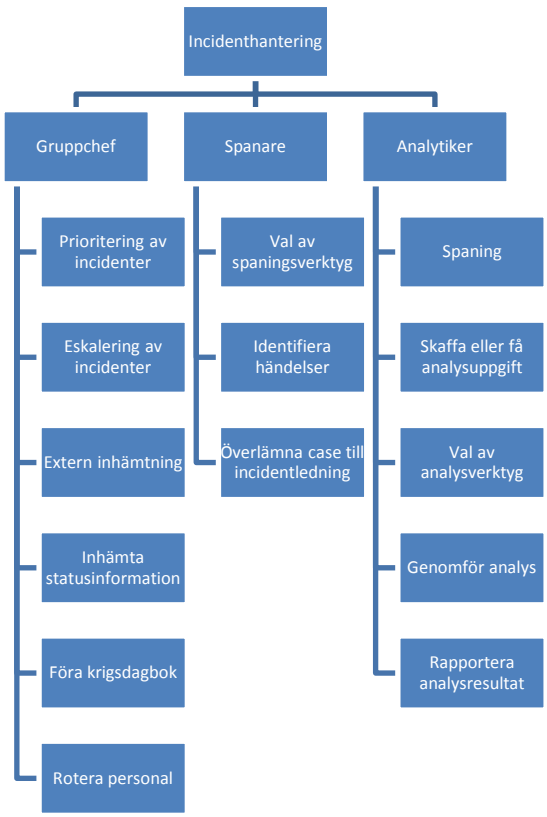
Inte alls bråttom ① ② ③ ④ ⑤ ⑥ ⑦ Mycket bråttom

Figur 5. Formulär för mätning av analytikers CSA efter att deltagarnas förbättringsförslag införts.

De omarbetade versionerna av formulären för mätning av CSA ansågs så värdefulla att de bedömdes kunna användas som rapporteringsverktyg i det dagliga arbetet för logganalytikerna. Nästa steg i arbetet med formulären är att digitalisera formulären för att de ska kunna användas effektivt vid datainsamling och i den operativa verksamheten.

3.2.2 Hierarkisk uppgiftsanalys

Den hierarkiska uppgiftsanalysen resulterade i en övergripande struktur av de aktiviteter som gruppchef, spanare och analytiker utförde under försöket, vilket illustreras i Figur 6. Utifrån detta resultat bör en vidare analys genomföras för att se vilka aktiviteter som bör inkluderas för respektive roll i en optimal situation. Detta beskriver det övergripande innehållet i deltagarnas arbetsuppgifter, vilket innebär att de mer detaljerade uppgifterna här exkluderats. Dessa resultat beskriver de aktiviteter som *faktiskt* genomfördes.



Figur 6. Resultat från hierarkisk uppgiftsanalys genomfört under försöket.

Försöket genomfördes med bra resultat och logganalytikerna fann sig väl i de roller de tilldelades. Logganalytikerna kommunicerade väl med varandra och periodvis nästan i överkant, eftersom kommunikationen ibland uppfattades som ett störningsmoment om operatören själv inte arbetade med incidenten. Vidare identifierades behov av checklistor och rutiner för en del arbetsmoment, bland annat:

- Checklista vid validering av misstänkt incident innan eskalering.
- Rutin för överlämning mellan roller. Till exempel ett system där ärenden som ska undersökas vidare kan bokföras tillsammans med tillhörande loggposter.
- Metod för att få överblick av vem som har vilken roll för ögonblicket. Förslagsvis integrerat med systemet för ”tickets”.
- Metod för att bibehålla överblick.

Miljön skilde sig från operatörernas ordinarie arbetsmiljö avseende logganalysverktyg, insamlade systemhändelser och arbetsplatser. Därför behövde betydligt mer tid läggas på aktiviteter som annars sker med automatik, som t.ex. val av logganalysverktyg för en viss analys.

3.2.3 Kommunikation

En kvantitativ analys av kommunikationen genomfördes av förmiddagens försök utifrån sändare, mottagare och vilken kategori av information som diskuterades. Kommunikationen under eftermiddagen exkluderades på grund av tekniska problem. Sändare är den person som initierar kommunikationen och mottagare är den person som sändaren adresserar informationen till. Som framgår av Tabell 1 så var deltagare tre (gruppchefen) den mest aktiva personen och som initierade kommunikation med övriga deltagare. De övriga fyra deltagarna sände ungefär lika mycket kommunikation (10,0 -15,4 procent). En högre chef hade under försöket en mindre roll och sände 3,0 procent av kommunikationen. De högra kolumnerna i tabellen redovisar vem som mottog informationen från respektive sändare. Försöket var mycket intensiv och sammanlagt identifierades 201 tillfällen när deltagarna sände information under förmiddagens försök. Troligen fanns det ännu fler tillfällen eftersom uppgiften att observera all kommunikation i realtid var mycket krävande. Deltagare 1 och 4 kommunicerade inte alls med varandra under försöket. Det är av intresse att notera att de satt längst ifrån varandra rent fysiskt. Deltagare 2 och 4 kommunicerade 70 procent respektive 80,8 procent med gruppchefen medan deltagare 1 och 5 kommunicerade mycket mer med varandra än med gruppchefen (32,3 respektive 16 procent med gruppchefen).

Tabell 1. Frekvenstabell för kommunikation med procentuell fördelning av vem som är sändare av information och vem som informationen är riktad till (mottagare).

Sändare	**Sänd information (%)**	**Mottagare**	**Mottagen information av respektive deltagare(%)**
Deltagare 1	15,4	2	3,2
		3	32,3
		4	0
		5	54,8
		Driftledning	6,5
		Högre chef	0
		Hela gruppen	3,2
Deltagare 2	10,0	1	0
		3	70
		4	15
		5	10
		Driftledning	10
		Högre chef	0
		Hela gruppen	5
Deltagare 3 (gruppchef)	43,3	1	18,4
		2	16,1
		4	25,3
		5	6,9
		Driftledning	10,3
		Högre chef	4,6
		Hela gruppen	18,4
Deltagare 4	12,9	1	0
		2	11,5
		3	80,8
		5	0
		Driftledning	0
		Högre chef	0
		Hela gruppen	7,7
Deltagare 5	12,4	1	60
		2	8
		3	16
		4	4
		Driftledning	4
		Högre chef	0
		Hela gruppen	8
Driftledning	3,0	3	83,3
		Ej kategoriserat	16,7
Högre chef	3,0	3	100

Ämnesdiskussioner var det överlägset vanligaste samtalsämnet (81,1 procent). I denna kategori omfattar alla diskussioner om spaning och analys av uppkomna incidenter. Endast 2,49 procent faller utanför de förutbestämda kategorierna (Tabell 2). Kategorin 'annat' innebär att innehållet i diskussionen inte rymts i uppsatta kategorier medan 'ej kategoriserat' innebär att den forskare som observerat kommunikationen inte hunnit med att kategorisera all information. För att få en djupare förståelse för vad som kommunicerades är en möjlighet är att spela in försöket på video och genomföra kommunikationsanalysen efteråt.

Tabell 2. Frekvenstabell för kommunikation med procentuell fördelning av vilka kategorier av information som diskuterades.

Kategori sänd information	**Procentuell fördelning**
Beslut	12,4
Ämnesdiskussion	81,1
Metoddiskussion	2,0
Överlämning	0,5
Tekniska problem	1,0
Oenighet	0,5
Annat	2,5
Ej kategoriserat	0,0

3.2.4 Enkätsvar

Den övergripande bedömningen av försökets värde gjordes utifrån den enkät deltagarna besvarade i slutet av försöket (Bilaga 4). Enkäten omfattade ett antal frågor som besvarades genom att ange ett värde mellan ett och sju, där ett innebar att deltagaren inte alls höll med och sju att deltagarna instämde helt. Medelvärdena på svaren visar att de upplevde att:

- De lärde sig mycket under försöket (M=5,2).
- Försöket var mycket verklighetstroget (M=5,8).
- De hade nytta av det de lärt sig under försöket i sitt dagliga arbete som logganalytiker (M=5,4).
- Svårighetsgraden var relativt hög (M=5,2).
- De saknade en tydlig strategi för att lösa uppgiften (M=3,8).

Dessa medelvärden visar att deltagarna ansåg att de lärde sig mycket, att innehållet var verklighetstroget och att de har nytta av det de lärt sig. De upplevde vidare att det var en hög svårighetsgrad och att strategin de använde för att lösa uppgifterna endast delvis var tydlig. I enkätens öppna frågor fick deltagarna möjlighet att beskriva vad de var speciellt positiva respektive negativ till. Deltagarnas förslag på vad som kan göras bättre inför kommande försök var framför allt att undvika initiala konfigurationsfel i nätverket och utplacering av sensorer. Valet av logganalysverktyg som används under försöket bör dessutom

utvärderas vidare för att optimera deltagarnas möjlighet att upptäcka hot. Deltagarna var speciellt positiva till att försöket gav dem möjlighet att träna i grupp där de provade att arbeta som gruppchef, spanare och analytiker. De var även positiva till att i dessa roller öva sina arbetsmetoder, processer, fördelning av arbetsuppgifter och ledningsidé.

3.3 Diskussion

Resultatet från försöket visar att det var värdefullt både för vidare forskning och för logganalysgruppen att vidareutveckla sina arbetsmetoder. Det visade sig vara svårt att stanna kvar i rollen som spanare. Det hände att spaningen övergick i analys istället för att lyfta upp händelsen till gruppchefen för fördelning mellan tillgängliga analytiker. Rollen gruppchef var ganska ny för de logganalytiker som innehade den under försöket, vilket innebar att en del prioriteringar blev svårare och att osäkerhet fanns vid uppgifter kring eskalering och inhämtning av information externt. Det behövs också metodstöd för att stödja gruppchefen att upprätthålla överblicken över arbetet eftersom det är lätt att fokusera för mycket på pågående incidenter. En ytterligare reflektion är att försöket speglar en situation med hög intensitet. Organisation och processer måste anpassas för att kunna bibehålla uthållighet och för att över tid hantera ett högintensivt arbete.

Den hierarkiska uppgiftsanalysen visar att gruppchefen har en krävande arbetssituation där bland annat bedömning av vilken eller vilka incidenter som ska prioriteras och kontinuerligt förande av krigsdagbok. Samtidigt så ansvarar gruppchefen för att fördela arbetsuppgifter och säkerställa att gruppmedlemmarna tar paus i arbetet för att effektiviteten ska kunna bibehållas över tid. Spanaren har en mer fokuserad uppgift med att framför allt identifiera vilka händelser som behöver analyseras och säkerställa att gruppchefen får vetskap om dessa. Även analytikern bedriver viss spaning men genomför framförallt djupare analys av de händelser som spanaren identifierat. Dessutom ska analytikern rapportera sina resultat till gruppchefen som i sin tur rapporterar vidare till högre chef.

Analysen av kommunikation visar bland annat att gruppchefens stod för en stor del av kommunikation, vilket medförde att han var högt arbetsbelastad. Forskare som observerade försöket noterade att gruppchefen blev ordentligt trött efter cirka en timme. Utrymme för ritytor på väggar och digitala hjälpmedel kan göra arbetet enklare. En annan viktig fråga är hur länge en spanare kan sitta fokuserat och titta på sin skärm där händelser och misstänka incidenter visualiseras. Den fullständiga analysen är inte klar och måste delvis genomföras av logganalytikern för att passa deras specifika arbetsuppgifter, men det är tydligt att det behövs ett genomtänkt och systematiskt schema för att logganalytikerna ska klara sina arbetsuppgifter när hotnivån är hög och det finns många incidenter att analysera.

Slutligen mottogs formulären för att mäta CSA mycket positivt och efter justeringar så finns nu ett verktyg för att mäta logganalytikers situationsmedvetande för både spanare (Figur 4) och analytiker (Figur 5). Även om formulären och enkäterna under försöket fungerade bra ska de inte betraktas som färdiga utan bör användas och utvärderas vid fler försök innan deras kompletthet och utformning kan fastslås med säkerhet.

4 Försök 2: undersökning om effekt av återkoppling

Försökets syfte var att undersöka om logganalytikers förmåga förbättras med återkoppling jämfört med när återkoppling inte ges. Analytikerna arbetade individuellt och deras uppgift var att upptäcka och analysera incidenter.

4.1 Metod

Experimentet baserades på en mellangruppsdesign avseende återkoppling (med och utan) och uppgifter (fyra olika uppgifter fördelade på två angreppstyper). Fokus var att undersöka effekten av återkoppling där hypotesen var att återkoppling på en uppgift gav en positiv effekt (avseende prestation) på efterföljande uppgifter. Två angrepp av samma typ analyserades under förmiddagen (uppgift A och B) och under eftermiddagen (uppgift C och D). Efter varje uppgift samlades incidentbeskrivningar in enligt ett standardiserat format. Dessa ses i försöket som ett prestationstest (PT) som jämfördes med angreppet för att undersöka hur bra deltagarna lyckades med uppgiften (Bilaga 9). Även en enkät (E) användes som bland annat frågade om deltagarna upplevde att de lärt sig något (Bilaga 5). Högre chef hade på förhand delat in logganalytikerna i två likvärdiga grupper. Ena gruppen (enligt Tabell 3) fick återkoppling efter respektive uppgift (F) medan den andra gruppen (enligt Tabell 3) inte fick någon återkoppling över huvud taget. Återkopplingen bestod av en femton minuter lång muntlig genomgång av vad som gjordes under angreppet, möjlighet att ställa frågor samt möjlighet att jämföra det faktiska angreppet mot den egna analysen och tillgängliga loggar. Efter att ha fått återkoppling så ombads de tre deltagarna som fått återkoppling att åter besvara enkäten (E). Datainsamling skedde genom att logga händelser i det tekniska systemet men redovisas inte vidare i denna rapport. Tabell 3 sammanfattar designen och datainsamlingen. Deltagarna tilldelades en timme för varje analysuppgift följt av 15 minuters återkoppling.

Tabell 3. Design och datainsamling.

	Angreppstyp 1			**Angreppstyp 2**	
Fp 1	$A^{PT, E}$	$B^{PT, E}$	Lunch	$C^{PT, E}$	$D^{PT, E}$
Fp 2	$A^{PT, E,}$	$B^{PT, E}$		$C^{PT, E,}$	$D^{PT, E,}$
Fp 3	$A^{PT, E, F \& E}$	$B^{PT, E, F \& E}$		$C^{PT, E, F \& E}$	$D^{PT, E, F \& E}$
Fp 4	$A^{PT, E, F \& E}$	$B^{PT, E, F \& E}$		$C^{PT, E, F \& E}$	$D^{PT, E, F \& E}$
Fp 5	$A^{PT, E, F \& E}$	$B^{PT, E, F \& E}$		$C^{PT, E, F \& E}$	$D^{PT, E, F \& E}$

4.1.1 Attacker/incidenter

Angreppen utfördes med verktyget SVED (Holm & Sommestad, 2016) och av fyra olika angreppsekvenser (A-D). Angreppssekvenserna utfördes separat och påbörjades under den första halvan av den timme som analytikerna hade tilldelats för att lösa respektive uppgift. Nedan beskrivs angreppen (A-D) kortfattat:

A. En e-postbaserad datormask som skickades till sju användare och efter att ha öppnats av dessa försökte masken e-posta sig själv till andra användare i cybermiljön.

B. En skadlig kod som via USB-sticka infekterade maskiner i miljön. Efter infektion försökte (men misslyckades) den skadliga koden att sprida sig själv till andra datorer genom att angripa nätverkstjänster. Den spred sig också till delade mappar, vilket infekterade ett fåtal maskiner.

C. En riktad skadlig kod som först komprometterade en maskin via en elakartad länk i ett e-postmeddelande. Därefter infekterade hotaktören flera andra maskiner i olika nätverk med hjälp av lösenordshashar för ett administratörskonto som erhölls från den ursprungligt komprometterade maskinen. Från varje övertagen maskin extraherades periodvis känslig data (t.ex. filer och tangentbordstryckningar).

D. En laptop som kopplades in i ett av kontorsnäten, spred skadlig kod via en USB-sticka och sedan angrep ett logistiksystem i nätet genom lösenord som kommits över från kontorsdatorn.

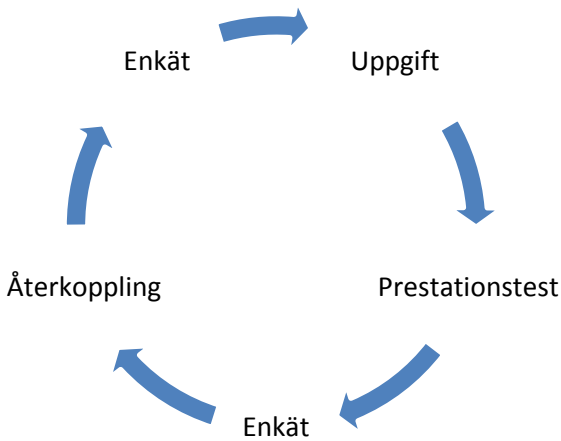
4.1.2 Formulär för mätning prestation

Framtagandet av formulär för mätning av prestation gjordes utifrån kunskaper och erfarenheter från deltagarna i projektet *Övning och experiment för operativ förmåga i cybermiljön* (Bilaga 9). Denna erfarenhet bygger både på deltagande i tidigare försök för logganalytiker och litteratutgenomgång (ENISA, 2010; NICE, 2015; Sommestad & Hunstad, 2013). Deltagarnas prestation utvärderades av två oberoende experter enligt framtagen mall (Bilaga 7) men dessa individuella data beskrivs inte vidare i denna rapport.

4.1.3 Genomförande

Deltagarna informerades av experimentledaren att deras uppgift var att spana och analysera upptäckta incidenter. De placerades runt ett bord (Figur 3) och instruerades att arbeta enskilt. Innan försöket startade fick deltagarna formuläret för prestationsmättning och uppmanades att succesivt nedteckna relevant information. Efter genomförd uppgift lämnades formuläret in och en enkät besvarades (Bilaga 5). Därefter fick gruppen som inte delgavs återkoppling lämna rummet (två personer) medan övriga tre deltagare mottog återkoppling

genom att en forskare förklarade detaljer i den attack som de utsatts för. Deltagarna fick även möjlighet att ställa frågor. Slutligen så fick deltagarna som mottagit återkoppling svara på enkäten ytterligare en gång (Bilaga 5). Därefter fick deltagarna tio minuters paus för att sedan börja arbeta med nästa uppgift. I slutet av dagen så fick deltagarna även fylla i en avslutande enkät (Bilaga 6).



Figur 7. Process där deltagarna som fick återkoppling arbetade med uppgift, prestationstest, enkät, återkoppling och enkät. Deltagarna som inte fick återkoppling genomförde inte de två sista stegen (återkoppling och enkät vid andra tillfället).

4.2 Resultat

På grund av de tekniska problem som omgärdade uppgift D kunde inte uppgiften användas för att undersöka lärandeeffekt. Prestationsmätningarna och enkätsvaren på uppgifterna A-C samt den avslutande enkäten beskrivs nedan.

4.2.1 Effekten av återkoppling på upplevt lärande

De tre deltagarna som erhöll återkoppling svarade på frågan ”Hur mycket upplevde du att du lärde dig under passet?” såväl före och efter återkoppling (fråga 2 på enkät som användes under försöket, Bilaga 5). Medelvärdena på dessa svar indikerar att återkoppling har en positiv effekt på lärande. Svaren på den sjugradiga skalan steg från 3,7 i snitt innan återkoppling till 4,8 efter återkoppling. Två av deltagarna upplevde liknande positiv lärandeeffekt (ökning med 2,7) medan den tredje personen inte upplevde en positiv lärandeeffekt (skattade lika före- och efter återkoppling). Deltagarnas skattning efter genomfört försök (Bilaga 6) visar att de upplevde att de lärde sig mycket av att få återkoppling (M=6,0).

4.2.2 Effekten av återkoppling på prestation

De incidentrapporter som producerades av deltagarna jämfördes mot de som producerats av ramverket SVED. Effekten av återkoppling kan i detta experiment ses som den skillnad i prestationsförbättring som finns mellan deltagare som fått återkoppling respektive deltagare som inte fått återkoppling. Dessvärre uppstod en golveffekt på grund av att uppgifterna var alltför svåra för att lösas inom angiven tid och med tillgängliga logganalysverktyg. Få rapporter identifierade ett enda attacksteg korrekt och endast 7 av 110 nedskrivna misstänkta händelser hade att göra med de faktiska angreppen. Detta uppmärksammades också av deltagarna själva som angav svårighetsgraden till ett medel på 5,2 på den sjugradiga skalan. Trots denna golveffekt finns antydningar till positiv effekt från återkoppling:

- Av de två personer som inte fick återkoppling upptäckte en person attacksteg i uppgift A och uppgift C. Ingen av dem upptäckte attacksteg i uppgift B.
- De tre personer som fick återkoppling upptäckte inget attacksteg i uppgift A eller B, men två personer upptäckte attacksteg i uppgift C.

Det finns alltså en indikation till att återkoppling ökar prestationen och att utebliven återkoppling har saknar påverkan på prestationen, men att dra slutsatser från dessa bristfälliga mätningar är vanskligt eftersom det endast var fem deltagare i försöket. Fler liknande experiment med uppgifter av lägre svårighetsnivå och mer kontroll för störvariabler som deltagarnas kompetens i olika specialområden behövs för att få klarhet i återkopplingens effekt på prestation.

4.2.3 Avslutande enkät

Den övergripande bedömningen gjordes utifrån den avslutande enkäten (Bilaga 6) som inkluderade ett antal frågor som besvarades genom att ange ett värde mellan ett och sju, där ett innebar att deltagaren inte alls höll med och sju att deltagarna instämde helt. Medelvärdena på svaren visar att de upplevde att:

- De lärde sig mycket under försöket (M=5,0).
- Försöket var verklighetstroget (M=5,0).
- De hade nytta av det de lärt sig under försöket i sitt dagliga arbete som logganalytiker (M=4,6).
- Svårighetsgraden var relativt hög (M=5,2).
- De saknade en tydlig strategi för att lösa uppgiften (M=4,0).

Dessa medelvärden visar att deltagarna ansåg att de lärde sig mycket, att innehållet var verklighetstroget och att de har viss nytta av det de lärt sig. De upplevde vidare att det var en relativt hög svårighetsgrad och att strategin de använde för att lösa uppgifterna endast delvis var tydlig.

4.3 Diskussion

Resultaten från enkäterna. incidentrapporter och de muntliga kommentarerna från deltagarna visar tydligt att uppgifterna upplevdes som mycket svåra, vilket till stor del bedöms bero på att de logganalysverktyg som användes inte var konfigurerade optimalt. Detta hämmade sannolikt delvis inlärningen och gjorde objektiva mätningar av återkopplingens påverkan problematiska. En lärdom från detta försök är att förberedelserna rent tekniskt behöver förändras för att undvika de initiala teknikproblemen och för att deltagarna ska ha bättre konfigurerade logganalysverktyg. En annan viktig lärdom är att svårighetsgraden behöver anpassas och om uppgifter tenderar att vara för svåra så kan försöksledarna delvis hjälpa deltagarna för att de ska kunna genomföra djupare analys av de identifierade incidenterna. Trots tekniska begränsningar framfördes mycket positiva kommentarer efter experimentet och det finns nu en bra grund att stå på inför kommande försök.

5 Slutsatser och fortsatt arbete

Denna rapport beskrev erfarenheter från träning av logganalytiker i kombination av att nya vetenskapliga metoder prövas. Två heldagsförsök följt av en halv dags utvärdering.

Första försöket (dag 1) gav deltagarna värdefull träning både avseende de faktiska uppgifterna som logganalytiker och i hur de bör arbeta som grupp. Deltagarna hade under försöket möjlighet att arbeta enligt sina vanliga rutiner och de kunde identifiera vad som fungerade bra och vad som de behöver träna mer på. Den hierarkiska uppgiftsanalysen resulterade i tre roller: gruppchef, spanare och analytiker. Gruppchefen har en krävande arbetssituation som bland annat innefattar bedömning av incidenters prioritet och förande av krigsdagbok. Därutöver ansvarar gruppchefen för att fördela arbetsuppgifter och säkerställa att gruppmedlemmarna tar paus i arbetet för att deras effektivitet ska bibehållas över tid. Spanaren har en mer fokuserad uppgift som framförallt handlar om att identifiera vilka händelser som behöver analyseras och kommunicera detta till gruppchefen. Analytikern bedriver också viss spaning men ägnar merparten av tiden åt djupare analys av de händelser som spanaren identifierat och gruppchefen prioriterat. Dessutom ska analytikern rapportera sina resultat till gruppchefen som i sin tur rapporterar vidare till högre chef.

Syftet med andra försöket (dag 2) var att undersöka om logganalytikers förmåga förbättras mer när de ges återkoppling efter en uppgift jämfört med när ingen återkoppling. Försöket var individuellt och fokuserade på att utveckla individuella färdigheter. Även om prestationsmåtten avseende återkoppling inte blev som önskat och endast fem personer deltog i försöket så bekräftades att återkoppling är viktigt för lärande i logganalys. Såväl prestationerna för de olika deltagarna, deras självskattningar av sitt lärande och diskussionerna under utvärderingen pekar på att återkoppling ger mervärde. Försök i anläggningar som CRATE har alltså påtagliga fördelar jämfört med till exempel lärande i det dagliga arbetet.

Det initiala formuläret för att mäta CSA för logganalytiker hade föregåtts av litteraturstudier och intervjuer med logganalytiker men aldrig utvärderats under försöksverksamhet. Den återkoppling som gavs i samband med försöket visar att framtagen metod fungerade bra men utifrån deltagarnas kommentarer modifierades formuläret delvis för att fungera så optimalt som möjligt. Det konstaterades även att en metod för att mäta gruppchefens CSA bör tas fram. Mätmetoden för kommunikationsanalys fungerade överlag bra, men alternativen för att kategorisera kommunikationen kan förfinas.

Inför kommande försök så behöver det finnas färdigkonfigurerade logganalysverktyg på plats. Det är anses viktigare att verktygen är intrimmade och fungerande än att de är identiska med de logganalysverktyg logganalytikerna

använder till vardags. Det kan till och med finnas ett värde i att prova andra logganalysverktyg än vad deltagarna är vana vid. Dessutom bör uppgifternas svårighetsgrad kontrolleras innan försök för att skapa ett försök som är anpassad till deltagarna. Detta innebär bland annat att verktyg för logganalys finns tillgängliga innan försöket för att säkerställa att angreppen går att se. Deltagarna påpekade att de övervakade nätverken inte behöver vara av realistisk storlek för att uppgiften ska upplevas som meningsfull och lärorik. Mindre och enklare miljöer bör testas framgent. En tänkbar framtida lösning är också att utföra uppgifter via fjärruppkopplingar mot anläggningen, till exempel i form av entimmes pass en gång i veckan.

Slutsatserna från denna studie är: 1) den framtagna metoden för mätning av CSA och prestation upplevs mycket positivt av logganalytiker men de nya versionerna måste utvärderas under kommande försök, 2) kommunikationsanalysen och den hierarkiska uppgiftsanalysen ger en bra bild av hur logganalytiker arbetar, 3) muntlig återkoppling ger en positiv inlärningseffekt, 4) försöksuppläggen fungerade bra men bland annat logganalysverktyg bör vara färdigkonfigurerade och testade innan försöket.

6 Referenser

Annett, J. (2003). Hierarchical Task Analysis *Cognitive Task Design*. London: Lawrence Erlbaum Associates, Publishers.

Barford, P., Dacier, M., Dietterich, T., Fredrikson, M., Giffin, J., Jajodia, S., . . . Yen, J. (2010). Cyber SA: Situational Awareness for Cyber Defense %U http://dx.doi.org/10.1007/978-1-4419-0140-8_1. In S. Jajodia, P. Liu, V. Swarup, & C. Wang (Eds.), *Cyber Situational Awareness* (Vol. 46, pp. 3-13 %G English): Springer US %8 2010-01-01.

Ben-Asher, N., & Gonzalez, C. (2015). Training for the unknown: The role of feedback and similarity in detecting zero-days attack. *In Proceedings of 6th Internatinal Conference on Applied Human Factors and Ergonomics (AHFE 2015)*.

Brynielsson, J., Franke, U., & Varga, S. (2016). Cyber Situational Awareness Testing. In B. Akhgar & B. Brewster (Eds.), *Combatting Cybercrime and Cyberterrorism, Advanced sciences and Technologies for Security Applications.* Switzerland: Springer International Publishing.

Cichonski, P., Millar, T., Grance, T., & Scarfone, K. (2012). *Computer Incident Handling Guide*. NIST National Institute of Standards and Technology.

Endsley, M. (1988). *Situation global assessment technique (SAGAT)*. Paper presented at the Aerospace and Electronics Conference, Hawthorne, CA.

Endsley, M. (1995a). Measurement of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 65-84.

Endsley, M. (1995b). Toward a Theory of Situational Awareness in Dynamic Systems. *Human Factors, 37*(1), 32-64.

Endsley, M., & Garland, D. (2000). *Direct meusurement of situation awareness: validity and use of SAGAT*. Mahwah, NJ.: Lawrence Erlbaum Associates.

Endsley, M., Holder, L., Leibrecht, B., Garland, D. J., Wampler, R., & Mathews, M. (2000). *Modeling and measuring Siuation awareness in the infantry operational environment* (1753). U.S. Army Research Institute for the Behavioral and Social Sciences.

Endsley, M., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). *A comparative analysis of SAGAT and SART for evaluations of situation awareness.* Paper presented at the Annual meeting of the Human Factors & Ergonomics Society, Chicago, IL.

ENISA. (2010). *Good practice Guide for Incident Management*.

Franke, U., & Brynielsson, J. (2014). Cyber situational awareness – A systematic review of the literature. *Computers & Security, 46*(0), 18-31. doi: http://dx.doi.org/10.1016/j.cose.2014.06.008

Frederiksen, J. R., & White, B. Y. (1989). An approach to training based upon principled task decomposition. *Acta Psychologica, 71*(1–3), 89-146. doi: http://dx.doi.org/10.1016/0001-6918(89)90006-1

Have, P. (2007). *Doing Conversation Analysis*. London: Sage Publications Ltd.

Hoffman, R. R., Ward, P., Feltovich, P. J., DiBello, L., Fiore, S. M., & Andrews, D. H. (2014). *Accelerated Expertise*. East Sussex, UK.: Taylor & Francis.

Hogg, D., Follesø, K., Torralba, B., & Volden, F. (1994). *Measurement of the operator's situation awareness for use within process control research: Four methodological studies* (HWR-377). Halden, Norway: OECD Halden Reactor Project.

Holm, H., & Sommestad, T. (2016). *SVED: Scanning, Vulnerabilities, Exploits and Detection*. Paper presented at the MILCOM 2016, Baltimore.

Kluger, A., & DeNisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin, 119*(2), 254-284.

Lif, P., Thorstensson, M., & Sommestad, T. (2015). *Övning, träning och prövning inom logganalys* (FOI-R--4149--SE). Linköping, Sverige: Ledningssystem.

NICE. (2015). National Cybersecurity Workforce Framework. from http://csrc.nist.gov/nice/framework/

Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know*. Washington DC: National Academy Press.

Sommestad, T., & Hunstad, A. (2013). Intrusion detection and the role of the system administrator. *Information Management & Computer Security, 21*(1), 30-40. doi: doi:10.1108/09685221311314400

Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics, 37*(1), 55-79. doi: http://dx.doi.org/10.1016/j.apergo.2005.06.003

Taylor, R. M. (1990). *Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In Situational awareness in aerospace operations* (AGARD-CP-478). Neuilly-Sur-Seine, France.

VINTHEC. (1997). *Visual interaction and human effectiveness in the cockpit (Final report)* (Technical Report: VINTHEC-WP1-TR-01).

7 Bilagor

Nedan presenteras enkäter, formulär och övriga dokument som användes under de tre övningsdagarna.

Bilaga 1 - CSA-frågor Spanare

A. Rita in var sensorerna finns (på nätverkskarta).

1. Hur många verkliga incidenter går det på varje ’false positive’ händelse?_____
2. Hur stor del av incidenterna upptäcker sensorerna?_____%
3. Hur många händelser har skett sedan förra spelstoppet?_____
4. Hur många incidenter har du rapporterat vidare till analytiker (från senaste spelstopp)?_____

Incident 1

5. Under vilken tidsperiod skedde incidenten (start- och sluttid)?_____
6. Vilken typ av incident var det?
   _____
   _____
7. Vad berodde incidenten på?
   _____
   _____
8. Vad försökte antagonisterna uppnå med sin attack?
   _____
   _____
9. Hur allvarlig var incidenten?

Inte alls allvarlig	①	②	③	④	⑤	⑥	⑦	Mycket allvarlig

10. Hur omgående behöver analytiker starta vidare utredningsarbete? (1-7)

Inte alls bråttom	①	②	③	④	⑤	⑥	⑦	Mycket bråttom

Bilaga 2 - CSA-frågor Analytiker

Följande formulär användes för att mäta CSA för analytiker.

Välj en incident och beskriv (rita) vad som hänt i nätverket.

Deltagarnummer._____

Inkludera följande information om möjligt.

- Källor (t.ex. IP-adresser, portar, MAC-adresser, hostnamn eller mailadresser)
- Mål (t.ex. IP-adresser, portar, MAC-adresser, hostnamn eller mailadresser)
- Starttid_____ Sluttid_____

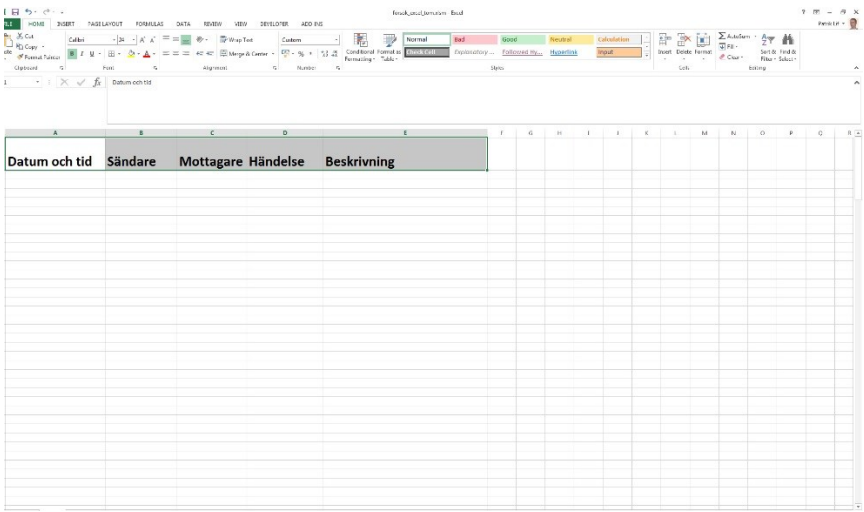
1. Hur säker är du på att grafen ovan är korrekt (ange i % i steg om 10)? Konfidensskattning____%
2. Vilken typ av incident var det (t.ex. DDos, insider, etc.)?
   _____Konfidensskattning____%
3. Vilka sårbarheter utnyttjades?
   _____Konfidensskattning____%
4. Vad försökte antagonisterna uppnå med attacken?
   _____Konfidensskattning____%
5. Hur allvarligt är det som skett (1-7)?

   Inte alls allvarligt ① ② ③ ④ ⑤ ⑥ ⑦ Mycket allvarligt
6. Var attacken automatisk? JA NEJ
7. Var attacken riktad? JA NEJ
8. Vilka åtgärd/åtgärder bör vidtas? _____ Konfidensskattning____%
9. Hur omgående behöver åtgärdema genomföras?(1-7)?

   Inte alls bråttom ① ② ③ ④ ⑤ ⑥ ⑦ Mycket bråttom

Bilaga 3 – Excelark för kommunikationsanalys

Kommunikationsanalys genomfördes under den explorativa studien (dag 1) för att kartlägga hur gruppen kommunicerar med varandra. Kommunikationsanalysen inkluderade tidpunkt, sändare, mottagare, vilken typ av händelse det var och beskrivning av händelsen. Förvald kategorisering var ämnesdiskussion, överlämning, beslut, tekniska problem, metoddiskussion, oenighet, utbildningsbrist och inspel/information från spelledning. Följande Excel mall användes:



Datum och tid	Sändare	Mottagare	Händelse	Beskrivning

Bilaga 4 – subjektiva frågor efter (dag 1)

Dessa frågor ska ställas efter dag 1, dvs. endast vid ett tillfälle.

Deltagarnummer _____

1. **Hur mycket upplevde du att du lärde dig under övningen?**

Väldigt lite	①	②	③	④	⑤	⑥	⑦	Väldigt mycket

2. **Hur verklighetstrogen upplevde du att övningen var?**

Inte alls	①	②	③	④	⑤	⑥	⑦	Mycket

3. **Hur stor nytta bedömer du att innehållet på övningen har för ditt vardagliga arbete?**

Väldigt liten	①	②	③	④	⑤	⑥	⑦	Väldigt stor

4. **Hur upplevde du svårighetsgraden under övningen?**

Mycket låg	①	②	③	④	⑤	⑥	⑦	Mycket hög

5. **Hade du någon strategi för att lösa uppgifterna?**

Ingen strategi	①	②	③	④	⑤	⑥	⑦	Tydlig strategi

6. **Beskriv kortfattat vad du upplevde som speciellt bra med övningen.**

_____

_____

_____

7. **Beskriv kortfattat vad du upplevde som speciellt dåligt med övningen.**

_____

_____

_____

Bilaga 5 – subjektiva frågor under experiment (dag 2)

Dessa frågor ska ställas efter respektive betingelse under experimentet dag två, dvs. vid fyra tillfällen. I samband med att dessa frågor ställs så ges deltagarna ett prestationstest för att mäta deras förståelse av attacken/attackernas innehåll.

1. **Deltagarnummer _____**

2. **Hur mycket upplevde du att du lärde dig under passet?**

Väldigt lite	①	②	③	④	⑤	⑥	⑦	Väldigt mycket

3. **Hur upplevde du det var att upptäcka incidenten?**

Mycket lätt	①	②	③	④	⑤	⑥	⑦	Mycket svårt

4. **Hur upplevde du det var att förstå vad som skedde under incidenten?**

Mycket lätt	①	②	③	④	⑤	⑥	⑦	Mycket svårt

5. **Hur upplevde du det var att bedöma vilka konsekvenser incidenten hade?**

Mycket lätt	①	②	③	④	⑤	⑥	⑦	Mycket svårt

6. **Hur upplevde du att du löste uppgiften?**

Mycket dåligt	①	②	③	④	⑤	⑥	⑦	Mycket bra

7. **Hade du någon strategi när du hanterade incidenten?**

Ingen strategi	①	②	③	④	⑤	⑥	⑦	Tydlig strategi

Bilaga 6 – subjektiva frågor efter experiment (dag 2)

Deltagarnummer _____

1. **Hur mycket upplevde du att du <u>lärde dig</u> under övningen?**

Väldigt lite	①	②	③	④	⑤	⑥	⑦	Väldigt mycket

2. **Hur mycket upplevde du att du <u>lärde dig av att få återkoppling</u>?**

Väldigt lite	①	②	③	④	⑤	⑥	⑦	Väldigt mycket

3. **Hur <u>verklighetstrogen</u> upplevde du att övningen var?**

Inte alls	①	②	③	④	⑤	⑥	⑦	Mycket

4. **Hur stor <u>nytta</u> bedömer du att innehållet på övningen har för ditt <u>vardagliga arbete</u>?**

Väldigt liten	①	②	③	④	⑤	⑥	⑦	Väldigt stor

5. **Hur upplevde du <u>svårighetsgraden</u> under övningen?**

Mycket låg	①	②	③	④	⑤	⑥	⑦	Mycket hög

6. **Hade du någon <u>strategi</u> för att hantera uppgifterna under övningen?**

Ingen strategi	①	②	③	④	⑤	⑥	⑦	Tydlig strategi

7. Beskriv kortfattat vad du upplevde som speciellt bra med övningen.

_____

_____

_____

_____

8. Beskriv kortfattat vad du upplevde som speciellt dåligt med övningen.

_____

_____

_____

_____

Bilaga 7 – expertbedömning av incidentbeskrivning

Bedömare _____

Avser deltagare _____ Tillfälle _____

1. **Hur bedömer du analysens <u>korrekthet</u> (figuren)?**

Mycket dålig	①	②	③	④	⑤	⑥	⑦	Mycket bra

2. **Hur bedömer du analysens <u>noggrannhet</u> (figuren)?**

Mycket dålig	①	②	③	④	⑤	⑥	⑦	Mycket bra

3. **Hur bedömer du korrektheten för typ av <u>incident</u>:**

Mycket dålig	①	②	③	④	⑤	⑥	⑦	Mycket bra

4. **Hur bedömer du korrektheten avseende <u>sårbarhet</u>.**

Mycket dålig	①	②	③	④	⑤	⑥	⑦	Mycket bra

5. **Hur bedömer du korrektheten för vad <u>antagonisten vill uppnå</u>?**

Mycket dålig	①	②	③	④	⑤	⑥	⑦	Mycket bra

6. **Hur bedömer du föreslagen <u>åtgärd</u>?**

Mycket dålig	①	②	③	④	⑤	⑥	⑦	Mycket bra

7. **Hur bedömer du incidentbeskrivningen som <u>helhet</u>?**

Mycket dålig	①	②	③	④	⑤	⑥	⑦	Mycket bra

Bilaga 8 – Frågeställningar vid kommunikationsanalys vid explorativ studie

Den explorativa studien undersökte kommunikationsmönster mellan deltagarna. De övergripande frågeställningarna som var av intresse beskrivs nedan. Dessa frågor är av övergripande intresse men avsågs inte att kunna besvaras under enbart denna studie.

1. Hur fördelar gruppchefen arbetsuppgifter?
2. Hur ställer högre chef uppgiften?
3. Vilka arbetsuppgifter utförs av respektive roll?
4. Finns det ett tydligt mönster i rollfördelningen?
5. Ges återkoppling mellan olika roller?
6. Hur prioriterar spanare när det är många larm?
7. Hur prioriterar analytiker när det är många larm?
8. Hur prioriterar gruppchefen vid hög arbetsbelastning?
9. Hur genomförs rapportering från spanare till gruppchef?
10. Hur genomförs rapportering från analytiker till gruppchef?
11. Hur genomförs överlämning av uppgift från spanare till analytiker?
12. Används någon mall för överlämning av uppgift?
13. Hur sker samarbetet mellan spanare?
14. Hur sker samarbetet mellan analytiker?

Bilaga 9 – objektiva frågor efter respektive del i experiment (dag 2)

Mall för beskrivning av attacksteg/elakartad aktivitet och sammanfattande beskrivning av incidenten.

Attacksteg/elakartad aktivitet

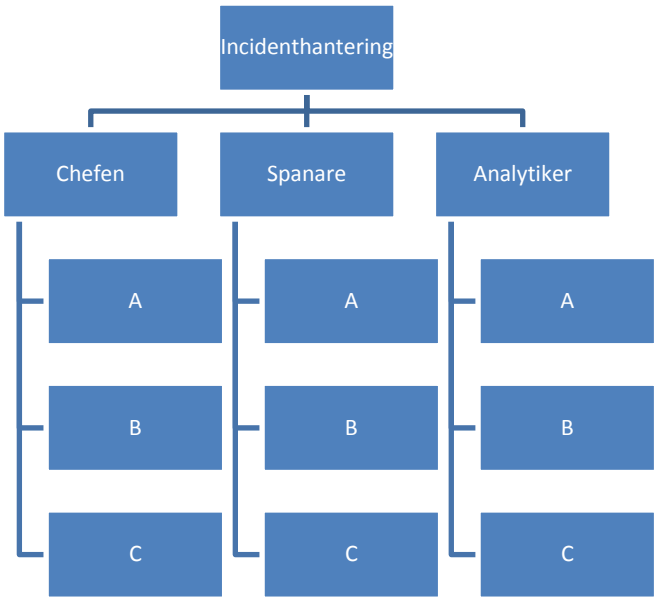
Tid	Händelse (vad som gjorts och konsekvenser/effekter)	Källa (t.ex. IP-address)	Mål (t.ex. IP-address)	Händelsetyp: Kartläggning	Händelsetyp: Angrepp	Händelsetyp: Missbruk	Del av angrepp (ange ett ID)
2016-11-07 10:02	DOS mot DNS. DNS dödl.	100 adresser av olika typer.	26.32.12.2:25		x		3
2016-11-07 12:02	20 min portskan för port 1-5000	192.128.32.1:25	126.32.12.2:25	x			2
2016-11-09 12:05	Serverattack mot webapp www.logistik.se. Användarprivilegier för www-data kompromitterade	192.128.32.1:80	126.32.12.2:80		x		2
2016-11-09 12:05	Mail med bilaga innehållande skadlig kod?	kalle@svensson.se; x-originating-ip 30.32.123.32	anders.and@ipfm.se; hvss32		x		1
2016-11-10 16:02	Nedladdning av extern fil till maskin; start av calc.exe. Användarprivilegier för sofex03 kompromitterade	152.128.32.1:25312	hvss32			x	1
2016-11-10 18:02	60 min portskan för port 1-8000	hvss32 (ursprung 152.128.32.1:25312)	hvss (allt)	x		x	2
2016-11-10 12:05	Eskalering från användare till administratör via MS14XX. Användarprivilegier för SYSTEM kompromitterade	kalle@svensson.se; x-originating-ip 30.32.123.33	anders.and@ipfm.se; hvss33		x		1

Incidentbeskrivning/angrepp

Angrepp (ID)	Starttid	Sluttid	Riktat	Autonomt	Tänkbar/trolig intention eller mål med angreppet
1	2016-11-09 12:05	pågår	X		Krysseland vill extrahera information eller få ett fotfäste i systemet.
2	2016-11-07 12:02	2016-11-10 18:02		X	Sannolikt nån datormask eller liknande.
3	2016-11-07 10:02	5 minuter senare	X		Nån random DoS-attack från nåt botnät. Oklart syfte.

Bilaga 10 – Diagram för aktiviteter och interaktion mellan gruppmedlemmar

Denna mall användes av expert inom logghantering för att genomföra hierarkisk uppgiftsanalys (Hierarctic Task Analysis - HTA) utifrån de tre rollerna gruppchef, spanare och analytiker. Syftet var att få en övergripande bild av deltagarnas genomförda aktiviteter.



Bilaga 11 – Diskussionsämnen dag 3

AAR genomfördes i helgrupp och syftade till att delge deltagarna information om försöket och att samla in erfarenheter för att utveckla framtida försök.

1. **Genomgång av incidenter (dag 1)**

Förklaring av attackerna på en övergripande nivå.

Diskussion om kommunikation i gruppen.

2. **Genomgång av CSA (dag 1)**

Genomgång av CSA-frågor och diskussion i helgrupp om det var bra frågor och huruvida frågorna behöver anpassas för spanare respektive analytiker.

3. **Genomgång av incidenterna (dag 2)**

Förklaring av attackerna på en övergripande nivå.

Diskussion om verbal återkoppling.

4. **Genomgång av protokoll för prestationsmätning (dag 2)**

Genomgång av PT-frågorna och diskutera i helgrupp om det var bra frågor och huruvida de behöver anpassas för framtida försök.

5. **Teknisk miljö**

    - Komplexiteten och funktionen på IPFM-nätet vi använt.
    - Sensorernas typ, placering, kalibrering med mera.
    - Användaragenternas beteende och funktion.
    - Arbetsplatserna under försöket.
    - Dokumentationen som finns tillgänglig.

6. **Övergripande frågor**

Diskussion om frågor som rör hela försöket och förklara eventuella oklarheter för deltagarna.

- Vad ska vi göra för att nästa försök ska bli bättre?
- Fungerade försöket utifrån ett tekniskt perspektiv?
- Lagom verklighetstroget?
- Svårighetsgrad?
- Grupp (dag 1) & individuellt. Bra eller dåligt?
- Övrigt?

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

FOI
Totalförsvarets forskningsinstitut
164 90 Stockholm

Tel: 08-55 50 30 00
Fax: 08-55 50 31 00

www.foi.se