Myndigheten för samhällsskydd och beredskap

Resilient Information and Control Systems (RICS)

Populärvetenskaplig sammanfattning av projektets resultat

Resilient Information and Control Systems (RICS) -
Populärvetenskaplig sammanfattning av projektets resultat

Tidsperiod: 2015-2020
Utförare: LiU, KTH, Chalmers
Ansvarig: forskare Simin Nadjm-Tehrani

Kort sammanfattning Kritisk infrastruktur är idag beroende av en motståndskraftig och säker IT-infrastruktur för att leverera sina tjänster. Inom ramen för projektet RICS har forskning bedrivits med syfte att ta fram ökad förmåga inom prevention och detektion av cyberattacker mot industriella informations- och styrsystem. Denna rapport sammanfattar resultat från detta arbete.

MSB:s Kontaktpersoner: Erik Sundström, 010-240 5371, MSB Forskningsenhet, forskning@msb.se

Foto omslag: Thomas Henrikson
Text: Simin Nadjm-Tehrani, Magnus Almgren, Mathias Ekstedt
Tryck: DanagårdLiTHO

Publ. nr: MSB1701, februari 2021
ISBN: 978-91-7927-109-1

MSB har beställt och finansierat genomförandet av denna forskningsrapport (alt. studierapport). Författarna är ensamma ansvariga för rapportens innehåll.

Förord

Forskningsrojektet RICS (Resilient Control and Information Systems) är ett femårigt projekt som MSB finansierat under området informationssäkerhet, särskilt industriella styrsystem (2015-2020). I denna rapport sammanfattas delar av projektetarbetet som har drivits i relativt nära samarbete med avnämare inom sektorn kritisk infrastruktur, t.ex. eldistribution, vatten och avlopp, byggnadsventilation, pappersproduktion. Det är författarnas förhoppning att denna skrift ska vara tillgänglig för en större krets läsare än läsare av vetenskapliga artiklar.

Projektet kommer i ny tappning pågå 2021-2023. Hela projektets vetenskapliga produktion och arrangerade evenemang finns tillgänglig via www.rics.se.

Projektgruppen vill tacka alla avnämare som har hjälpt oss med kunskaper, data, eller genom att lyssna och ifrågasätta delresultat i vår forskning.

Linköping, 2021-01-26

Simin Nadjm-Tehrani

Professor, Institutionen för Datavetenskap, Linköpings universitet

Innehåll

Inledning

Digitaliseringen påverkar alla sektorer i samhället och för med sig många fördelar. Vissa system blir mer kostnadseffektiva, i andra kan kapaciteten utnyttjas bättre. Flexibiliteten ökar i och med att man kan koppla upp sig utan att vara fysiskt nära systemet.

Även fastän fördelarna är många finns det också nackdelar. Digitaliseringen innebär att vi förlitar oss på att datorer och programvara är korrekt byggda och dessutom robusta mot cyberattacker. Det sistnämnda är speciellt viktigt eftersom många system är fjärrstyrda och kan då attackeras av utomstående, kanske till och med av personer som fysiskt befinner sig i andra länder. Även fastän detta är generella problem för alla typer av system, är de av extrem betydelse för samhällskritiska system eftersom konsekvenserna kan bli stora om de fallerar. Samhällskritiska system, t.ex. styrsystem som används inom energiförsörjning, vatten och avlopp, byggnadsventilation, trafikstyrning, och en del produktionsmiljöer, är ofta heterogena stora system med många komponenter. De utgörs av både gammal och ny teknologi och allmän teknologi (COTS) blandad med specialbyggd sådan. En del komponenter har lång levnadstid och deras styrning av den fysiska miljön kräver specifika kunskaper från domänen. Dessa processnära IT-system benämns populärt operationell teknologi (OT), i kontrast till generell IT. Informations- och kommunikationsinfrastrukturen (IKT med IT och OT sammantaget) är avsedd att leverera en tilltänkt tjänst inom givna kvalitetsramar. I kritisk infrastruktur innebär detta höga krav vad gäller korrekthet, determinism (teknisk förutsägbarhet) och tillgänglighet. Ett informationssäkerhetshot har därmed högst relevans om dess konsekvenser kan påverka leveransen av de tilltänkta tjänsterna. Utöver tillgänglighet är skydd av information och otillbörlig tillgång till data centrala inom det traditionella (IT-drivna) informationssäkerhetsområdet.

Trots att dagens samhälle är fullständigt beroende av fungerande IKT så är riskhanteringen av densamma för kritiska samhällsfunktioner inte alltid högst på agendan. Historien har visat att det är ofta först efter ett större haveri som resurser mobiliseras för att sätta fokus på hur avbrott och andra incidenter skulle förutsetts och undvikits. Samtidigt finns det en tendens bland aktörer inom säkerhetsbranschen att måla drastiska bilder av alla möjliga hot och sårbarheter innan haverier äger rum.

Några av de mer väldokumenterade historiska attackerna inkluderar intrången mot vattenreningsinfrastrukturen i Maroochy Shire i Australien 2000 med en så-kallad ”insider”. Ett decennium senare såg världen ett fall av cyberkrigsföring med Stuxnet där attackerarens avsikt verkade vara att stoppa anrikningen av uran i Iran. Stuxnet är ett av de första exemplen på kod som specifikt riktar sig mot avancerade styrsystem och som dessutom involverade icketekniska element. Sedan attackerades elnätet i Ukraina, inte bara vid ett tillfälle utan faktiskt två år i rad, med stora strömavbrott som följd. Dessa attacker tydliggör att också samhällskritiska system som används av alla (dvs. elnätet) numera kan attackeras.

Samtidigt som cyberskydd utvecklas kontinuerlig, är de flesta av sådana skyddssystem inriktade mot vanlig IT infrastruktur med större resurser i processorkraft och minne. Samhällskritiska system är oftast mer begränsade och körs dessutom i annorlunda miljöer där protokoll och programvara skiljer sig från IT system. Det innebär att en del av de cybersäkerhetslösningar vi har tillgängliga helt enkelt inte passar dessa miljöer. Dessutom får några av egenskaper som vi eftersträvar inom IT system (tillgänglighet, konfidentialitet) andra dimensioner i system där kärnan är att styra fysiska processer.

Projektet ”Säkra IT-system för drift och övervakning av samhällskritisk infrastruktur” (eng. ”Resilient Information and Control Systems” – RICS) har arbetat mot ökad förståelse av informationssäkerhetsproblem inom just samhällskritiska system. Huvudsakliga arbetsmetoder för att öka informationssäkerhet inom samhällskritiska styrsystem har följts i två huvudspår.

- Prevention – med en god förståelse för de risker som systemet utsätts för arbetar man med att minska sannolikheten för att kända säkerhetsbrister leder till ett totalhaveri varvid systemets huvudsakliga funktion uteblir. Genom att lära känna systemets brister genom analyser av sårbarheter och risker, kan man prioritera de delar av systemet eller de aktörer som kommer i kontakt med systemet som är viktigast för att undvika haverirer.
- Detektion – genom att övervaka förloppet av ett system kan man känna igen avvikande beteenden (så kallade anomalier) och därmed motverka oönskade förlopp tidigt och motverka ett systemhaveri. Även om en del oönskade scenarier går att förutse finns det andra antagonistiska hot som inte är kända i förväg.

Målet med denna skrift är att beskriva hur projektets arbete har skett, vilka nya metoder har utvecklats inom dessa två områden och hur arbetet kan komma till användning efter projektslut av: 1) företag som äger systemen och har ansvaret för driftsäkerhet gentemot sina kunder och myndigheter, 2) företag som levererar konsulttjänster eller specialiserade produkter riktade mot samhällskritiska system, 3) myndigheter som har ansvaret att samordna säkerheten inom en sektor (t.ex. elleverans), eller har intresse att vidareutbilda personal inom dessa sektorer till en högre kompetens.

Prevention

Inledning

De systemegenskaper för samhällskritiska system som beskrevs i inledningen ovan gör det svårt att bedöma en övergripande säkerhetsnivå. Det är inte säkert att man har en fullständig medvetenhet om alla tillgångar i systemet i alla konfigurationer delvis beroende på att det inte är möjligt att använda traditionella datainsamlingsverktyg överallt eftersom miljöerna är känsliga (i och med att de styr

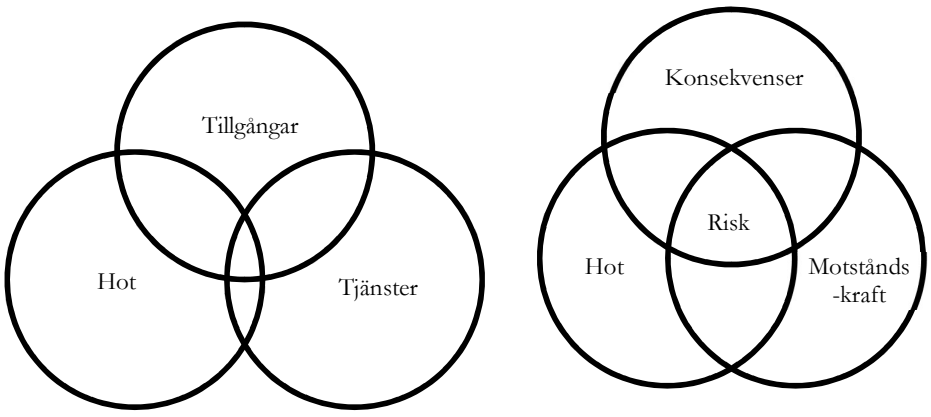
kritiska processer). Samtidigt är det självklart så att utan en helhetsbild av vad som ingår i sammansatta cyberfysiska system och en hel del kunskaper om sårbarheter i varje komponent och varje operationell process så är det svårt att bedöma vari de största riskerna till leveranskontinuitet ligger.

Många ansatser för att bedöma risker och dokumentera det underliggande resonemanget har tre ingredienser som belyses i Figur 1 (vänstre). Vilka tjänster som är fokus för analysen? Vilka systemtillgångar eller säkerhetsrelaterade tillgångar är nödvändiga för att kunna säkerställa leveransen av tjänsten och bör därmed vara en del av analysen? Vilka hotbilder ska beaktas?

Metoder och angreppssätt för analys av cybersäkerhetsrisker

Riskanalys av tekniska system är ett väletablerat område och avsikten här är inte att täcka all teoribildning inom riskanalys, utan endast att presentera de synsätt som drivit arbetet inom RICS.

Generellt sett brukar riskanalys ta sin utgångspunkt i att risk definieras som ett värde som bestäms av sannolikhet gånger konsekvens för att något oönskat händer i det studerade systemet. Inom ramen för OT-miljöer för kritisk infrastruktur kan vi inledningsvis tänka på konsekvenser som har negativa effekter på den fysiska processen och övrig verksamhet som bedrivs av infrastrukturägaren. Den oönskvärda konsekvensen skulle för ett elnätsbolag kunna handla om oplanerade och långvariga strömavbrott, skador på transformatorer, eller läckta kunduppgifter. Storleken på konsekvensernas upplevda skada kan variera mellan olika roller och aktörer relaterade till verksamheten. Sannolikheten å andra sidan att sådana negativa verksamhetseffekter skulle realiseras kan generellt sett påverkas av många källor, alltifrån stormar till materialutmattning och mänskliga misstag.



Figur 1: Centrala element som bidrar till förståelse av risker inom verksamheten visas på vänstersidan och komponenter som krävs för kvantifiering av risker visas på högersidan.

Vad gäller riskanalys inom cybersäkerhetsområdet avgränsas orsaksfloran till att på ett eller annat sätt involvera hot som utövas genom IKT. Med denna avgränsning kan man konstatera att sannolikheten för att konsekvenserna skall realiseras då beror dels på IKT-infrastrukturen och dess komponenters inneboende motståndskraft mot attacker samt attackerarnas kompetens, drivkrafter och resurser. Vi illustrerar detta i Figur 1 (högre).

Trots att riskanalys alltså tydligt involverar analyser inom tre områden som förs samman för att tillhandahålla övergripande riskbedömningar så har i princip alla riskanalysmetoder sitt fokus eller sin utgångspunkt i någon av dem. Dessa tre områden täcker dessutom tre väldigt separata kunskapsdomäner; någon viss process med tillhörande verksamhet (exempelvis elnät), IT och OT samt tillhörande säkerhet, samt hotaktörers egenskaper och förmågor. Vi kan vidare konstatera att dessa tre i sig är komplexa och att denna inbördes komplexitet tydligt påverkar den övergripande riskanalysen. För att exempelvis förstå konsekvenserna av en cyberattack där en viss brytare öppnas i ett elnät måste man förstå elnätets utformning och driftläge. Kanske utlöser brytarfrånslaget en kaskadeffekt som orsakar ett stort strömavbrott, eller kanske händer ingenting.

Hur svårt det är för angriparen att ta sig tillräckligt långt in i IKT-infrastrukturen för att skicka signalen att öppna brytaren beror på hur en stor mängd styrsystem och andra datorer är sammankopplade och vilka typer av IT-skyddsmekanismer som finns implementerade. Hur troligt det är att olika attackerare faktiskt försöker att genomföra angreppet beror exempelvis på deras politiska och ekonomiska motiv samt på deras kunskap om såväl den angripna IKT-infrastrukturen som elnätet. Alla dessa ämnen måste riskanalytikern skaffa sig information och kunskap om. Om inte alla dessa delar analyseras med samma nogsamhet kommer riskanalysen i slutänden att vara obalanserad. Att göra en välbalanserad riskanalys är alltså mycket svårt, men är dock kanske inte nödvändigt i alla situationer. Om syftet med riskanalysen exempelvis främst är att ta fram en beredskapsplan för händelser av lyckade angrepp blir det naturligt att fokusera på verksamheten och dess konsekvenser, är målet att öka verksamhetens motståndskraft mot cyberangrepp blir fokus de komponenter som skyddar IKT-infrastrukturen, och om frågan är hur stor IT-säkerhetsbudget organisationen bör ha kommer fokus oundvikligen hamna på hotbilden.

Metod och angreppssätt varierar också kraftigt inom riskanalysområdet. I sin enklaste form bedöms de tre områdena på enkla nominalskalor (exempelvis 1-5 eller hög-medel-låg) som sedan vägs samman till ett totalt riskvärde eller som illustreras i riskmatriser (med sannolikhet och konsekvens på axlarna). I andra änden av metodspannet finns de som använder statistiska beräkningar med fördelningar över tid, kostnader och andra storheter som också tar hänsyn till osäkerheten i bedömningarna. En allmän diskussion inom riskanalysområdet är hur kvantitativt och statistiskt det är möjligt och meningsfullt att genomföra riskanalyserna. Denna diskussion hänger i sin tur också samman med en annan

diskussion om datadriven respektive "antagandedriven" analys. I princip skulle man vilja basera sina analyser på statistiska data från observerade fenomen (hur lång tid tar det för olika attackerare att lyckas med olika typer av attacker och vad blir kostnaderna av strömavbrott på olika platser och tider, etc.). Denna typ av data är dock förstås en bristvara, i alla fall utanför organisationen, inte minst eftersom riskanalys naturligt omgärdas av mycket sekretess, vilket gör att man istället ofta är hänvisad till experters bästa gissningar som grund för sina bedömningar. Beroende på tillgänglig information och syfte är det vanligt att riskanalyser också antingen följer ett angreppssätt som primärt är uppifrån-och-ned eller nedifrån-och-upp.

Ytterligare en inneboende utmaning inom riskanalysen är att hantera den strukturella komplexiteten i det analyserade systemet, som nämnts ovan. Det är förstås så att det finns en stor mängd attackerare som utgör det totala hotet, det finns en stor mängd potentiella attackytor på IKT-infrastrukturen, det finns en stor mängd attackvägar från dessa attackytor som leder fram till en stor mängd värdefulla tjänster och information i verksamheten, som i sin tur alla potentiellt har en stor mängd olika typer av konsekvenser. Även för starkt avgränsade analysobjekt så har kombinationerna av alla dessa varianser en tendens att explodera och överskrida vad som är hanterbart för analytikern. Att skaffa sig en övergripande bedömning av en "total risk" är således svårt, både konceptuellt och praktiskt. Lösningen blir förstås att förenkla problemet på olika sätt. Metodmässigt kan man skönja angreppssätt som är checklistbaserade respektive beroendebaserade. Den föregående kategorin kan illustreras med många typer av standarder som exempelvis kan stipulera en uppsättning goda IT-säkerhetsskydd som kan prickas av för att uppnå lägre övergripande risk. Den senare kategorin baserar istället ofta på någon form av graf- eller trädstruktur i vilken orsak- och verkanssamband beskrivs. Exempel på detta är klassiska metoder baserade på felträd samt attackgrafer. Det är utvidgningar av denna senare paradigm som RICS arbetat med.

Forskning inom RICS

Forskarna på Kungliga Tekniska högskolan (KTH) arbetar med metoder och formalismer för att automatiskt generera attackgrafer utifrån modeller av IKT-infrastruktur. Tanken är att de ingenjörer som förvaltar och utvecklar IKT-infrastruktur skall kunna få automatiskt stöd att genomföra riskanalyser om de kan tillhandahålla en beskrivning av den existerande eller tilltänkta systemdesignen. Detta analysstöd tillhandahålls i form av vad som skulle kunna ses som virtuella penetrationstester av systemmiljön eftersom attackgraferna visar vilka möjliga sätt miljön kan angripas.

Inom ramen för RICS har förfinade analyser utvecklats specifikt inom smarta elnät med ett scenario för införande av distribuerad och förnybar elproduktion. I samarbete med ett annat forskningsprojekt[1] utvecklades en omfattande referensarkitektur. Referensmodellen beskriver IKT-infrastruktur hos en

elnätsoperatör med centralt styrsystem (SCADA), transformatorstationsautomation och elmätare hos kunder, en elproducent med både handelssystem och driftsystem, en styrsystemsleverantör samt en stamnätsoperatör. Alla systemmiljöerna är förenklade men ändå realistiska i sin arkitekturella uppbyggnad. Totalt sett avbildas 24 olika datornätverk i en modell som innehåller 560 olika systemkomponenter. Studien jämför sedan fyra olika försvarsstrategier, tex användandet av DMZ-nätverk, olika grader av uppdaterad programvara och härdning av operativsystem. För varje scenario beräknas en uppskattad fördelning för hur lång tid det tar för en hypotetisk attackerare att nå olika systemkomponenter i infrastrukturen samt tillhörande attackvektor. På så sätt kan olika scenarier jämföras med varandra utifrån ett attackmotståndskraftsperspektiv. Analyserna finns redovisade i en artikel i tidskriften Energy Informatics.[2]

För att automatiskt kunna generera attackgrafer från systemmodeller på det sätt som beskrivs ovan behövs att modellerna följer en fördefinierad struktur. Detta görs genom att modellerna beskrivs i speciellt utformade domänspecifika språk (DSL, Domain Specific Languages). Dessa språk definierar vilka typer av attacker och försvar som *potentiellt kan finnas* i en viss domän och dess olika systemkomponenter. Man kan i ett språk exempelvis stipulera att systemmodeller skall innehålla *nätverk*, *datorer*, *data* och *inloggningsuppgifter* och att om *datorer* är kopplade till samma *nätverk* kan dessa kommunicera med varandra, men också att är möjligt att försöka exekvera skadlig kod från den ena *datorn* på den andra bara på grund av att de är kopplade till samma *nätverk*. För att kunna programmera sådan attacklogik i ett domänspecifikt språk behövs ytterligare en nivå av formalism för hur detta skall göras, nämligen ett metaspråk. I RICS har vi bidragit till utvecklingen av Meta Attack Language (MAL). MAL beskriver alltså de grundläggande primitiver som används för att bygga domänspecifika språk (attacksteg, försvarsmekanismer, systemkomponenter) samt hur deras beroenden används för att generera probabilistiska attackgrafer. Utöver den formella beskrivningen av MAL finns även en språkkompilator och ett antal språkutvecklingsinitiativ samlade på GitHub.[3] MAL har inom ramen för andra projekt sedermera använts till att bygga domänspecifika språk för säkerhetsanalyser inom exempelvis transformatorstationsautomation, fordonsautomation och molnmiljöer.

Sett från perspektivet i Figur 1 (vänstre), har det arbete som forskare vid Linköpings universitet (LiU) drivit följt en tillgångsbaserad ansats. Ansatsen bygger på följande iakttagelse: om risker mot ett system som är under utformning ska kvantifieras, så krävs det metoder för att identifiera vilka tillgångar som behöver skyddas. Metoden går ut på att redan vid anskaffningen identifiera de tillgångar som bör vara i fokus för säkerhetsanalysen. De flesta skyddsvärda tillgångar kan betraktas som data som skall lagras, eller data som skall skickas inom nätverk (kunddata, mätvärden viktiga för att kunna bevara stabiliteten i den fysiska processen). Men även kunskapen om en viss algoritm, lagrad kod för att realisera den eller nycklar för att komma åt den är att betrakta som tillgångar. Metoden går

ut på att genom detta fokus på tillgångar och de systemkomponenter som tillgången huseras i eller passeras igenom identifiera vilka attacker som är sannolika och vilka säkerhetsmekanismer som kan motverka dem. Detta resonemang förs i flera iterationer. Allteftersom nya säkerhetsmekanismer, som i sin tur lägger komplexitet till systemet, tillförs, identifieras nya tillgångar (t.ex. kryptonycklar, certifikat och andra tillgångar) som blir potentiellt lika viktiga att skydda som det ursprungliga systemets skyddsvärda objekt.

Ett senare arbete vid LiU tar avstamp i det tredje perspektivet i Figur 1 (vänstre): tjänster med fokus på systemägarens affärsfokus. Innan man modellerar systemet för att resonera kring riskerna ställer man frågan: Om systemet ska tillhandahålla sina tjänster vad är det som krävs för en säker operation av systemet? Vilka element bidrar till detta och hur kritiskt är varje element? Vilka andra (element eller tjänster) krävs för en säker operation och vilka beroenden skapar dessa? I detta arbete som initierades inom RICS genom att etablera samarbete med universitetet i Cardiff och Airbus group (security and innovation) har vi fokuserat explicit på SCADA system. Genom att samla input från 36 domänexperter på olika nivåer i verksamheter (management, driftingenjör, säkerhetsansvarig osv.) börjar vi från ett avnämareperspektiv. 17 av dessa avnämare kom från kontakter inom RICS projektet. De olika experterna angav svar till likartade frågor och efter vår bearbetning har vi lyckats skapa en generisk beroendemodell för SCADA system (modellen skapades från 1521 element angivna av experter av vilka 640 element var unika). Modellen kan ses som ett målorienterat träd som på den högsta nivån har sex högnivå-element som bidrar till SCADA systemets mål. Dessa element täcker vitt skilda delar av beroenden som påverkar riskanalysen (Management, Employees, Data, System life cycle, System architecture, External dependencies). Arbetet är realiserat som en anpassningsbar modell (en blueprint) med hjälp av Open Groups ”dependency modelling standard” inom verktyget iDepend.[4] Modellen har tillämpats på ett antal exempel och kan utvärderas vidare inom verksamheter som visar intresse.

Sammanfattning av arbetet inom prevention

Riskanalys och dokumentation av vilka antaganden som ligger bakom en viss riskbedömning är fortfarande ett område under utveckling. Arbetet i RICS och vår exponering av resultaten till olika avnämare har visat att modellering hjälper att skapa de dialoger som krävs för att bestämma ”rätt” finkornighet på systemets delar och det fokus som passar just i det sammanhang som arbetet bedrivs. Olika aktörer, alltifrån systemutvecklare till driftpersonal och beslutsfattare kan ha olika ingångar och olika behov av förutsägelser, vilket gör att en flora av ansatser för riskanalyser inte nödvändigtvis är en nackdel, utan kanske till och med en fördel för att ta fram olika synvinklar.

Detektion

Inledning

Ett av målen med RICS har varit att förstå hur man kan bygga mekanismer för att skydda kritisk infrastruktur med hjälp av så-kallade intrångsdetekteringssystem. Dessa system började utvecklas redan under 1980-talet med två varianter. I den första definierar man hur kända attacker ser ut och systemet larmar när sådana mönster hittas i systemet (missbruksdetektering). I det andra definierar man hur systemets beteende under normala förhållanden ser ut, och i det fallet larmar man när systemet verkar uppföra sig annorlunda mot det normala (avvikelse/anomalidetektering). Det säger sig självt att missbruksdetektering är bra för att upptäcka kända attacker medan anomalidetektering krävs om hittills okända attacker ska detekteras.

Tyvärr fungerar anomalidetektering inte så bra rent praktiskt på vanliga IT system. IT system används på så många olika sätt vilket innebär att det är svårt att skapa en generell normalmodell. Samhällskritiska system, däremot, har ofta ett mer definierat normalbeteende i och med att dessa system består av styrenheter (inom OT-delen) vars mål är att styra processer som beter sig enligt fysiska lagar, även om dessa enheter kan kommunicera med andra system inklusive IT enheter. Detta har inneburit att avvikelsedetektering ses som mycket lovande för OT-delen medan IT-delen kan skyddas av mer traditionella cybersäkerhetsmetoder.

Metoder och angreppssätt inom anomalidetektion

Trots att dagens OT börjar likna IT systemen i och med att man använder sig av lösningar för nätverkande och protokoll som är gemensamma, t.ex. vid trådlös åtkomst och avläsning av sensorer, så ligger i grund och botten ett tidstyrt styrsystem i de flesta SCADA miljöer. Systemets funktion är byggt för ett visst ändamål som inte varieras inom korta tidsintervaller. Denna relativa stabilitet över relativt långa perioder bör vara en fördel vid monitorering för avvikelser. Men det finns fortfarande samma typer av sårbarheter som inom IT systemens mjukvara som kan leda till haveri, t.ex. felaktiga anrop och fel format på input (datapaket) som förflyttar sig genom nätverket. För att motverka dessa behöver man använda samma mekanismer som i övrig programvaruutveckling: genomgång av designval, välfungerande implementeringsprocesser, analys av risker och fokus på känsliga punkter. Men man kan även utnyttja en ny möjlighet som är svårare att finna inom IT system: att under den operationella fasen observera nyckeldata om systemets beteende och kontrollera ifall något avviker från normalitet. Detta kan ge en utökad nivå av skydd vid avsiktlig felanvändning utöver kvalitetssäkring vid anskaffande.

Anomalidetektering sker med hjälp av kod som skapas för det ändamålet och lagras i delar av systemet där nyckeldata kan avläsas, antigen i noder (så kallade host-based detection) eller i nätverkspunkter där data passeras och kan lagras för realtidsanalys eller senare undersökningar (så kallade network-based detection). De olika nyckeltalen som man bör lägga fokus på är delvis beroende av systemets egenskaper och delvis metoden som man använder för monitorering. En annan faktor som påverkar vilka mätvärden som registreras har att göra med vilka tänkbara hotmodeller som beaktas. Ett insider hot innebär att man måste samla data på ställen (och därmed den sorts data) som är svår att påverka även för den som har stora rättigheter i systemet, t.ex. genom att sätta ett fysiskt systems olika mätvärden i relation till varandra. En attackerare som i realtid ändrar på flera mätvärden i harmoni med fysiska lagar har en mycket svårare uppgift än en som matar nätverket med paket som har fel format. Föreställningar av hot utifrån kan leda till att man fokuserar på signaler som kommer in i systemet eller skickas mellan delsystemen och observerar deras eventuella avvikelser.

Även när man granskar dataflöden över tid, genom att tillfälligt lagra en sekvens av paket och granska dessa, kan man fokusera på olika fält, t.ex. de fält som är avsedda för ett specifikt protokoll (t.ex. instruktionsfält, flaggor) eller mätvärden/kommandovärden inuti. Ett specialfall av protokollspecifika delar i varje paket är den tidstämpel som anger *när* detta paket skapades, dvs. i anslutning till att något mätvärde i processen avlästes eller en styrsignal skickades av operatören.

Forskning inom RICS

Under RICS har vi studerat två varianter på anomalidetektering: en som avser att övervaka avvikelser som går att iaktta i den fysiska processen, och en som granskar avvikelser i det mönster över tid som kommunikation sker i nätverket. De beskrivs närmare nedan.

I det första fallet observerar vi att en orsak till att samhällskritiska system attackeras är för att i slutändan påverka en fysisk process, som skedde i de dokumenterade fallen ovan med Maroochy, Stuxnet och Ukraina. Metoden modellerar normalbeteende på mätvärden och hur givaren används. Fördelarna ligger i att metoden är datadriven, dvs. man behöver inte bygga upp en separat modell av (det komplicerade) systemet utan man kan använda historiska data direkt. Andra forskare har också undersökt datadrivna metoder, men de flesta av dessa använder insamlad data för att bygga en modell för att *förutsäga framtiden,* där modellen beräknar en prognos av vad nästa värde borde vara. Om sedan nästa mätta värde från systemet skiljer sig från denna prognos skapas ett alarm. Det är dock svårt att förutsäga framtiden så dessa prognoser är inte så exakta. Det medför att en attackerare kan gömma sina attacker i bruset från systemet. RICS metod, döpt till PASAD[5], bygger på en enkel men kritisk observation: förutsäg aldrig framtiden utan arbeta hela tiden med värden från nutid. Sålunda byggs en spektralmodell upp och när ett nytt värde samlas in jämförs detta med modellen

direkt för att se om det liknar värden vi borde se. Om inte skapas ett alarm. Därmed har vi kunnat visa att RICS modell verkar vara mer robust än andra liknande lösningar.[6]

Den andra metoden utvecklad i RICS kan antyda om andra avvikande fenomen, t.ex. att någon beräkningsmodul i systemet agerar långsammare än normalt (är överbelastat p.g.a. ett oavsiktligt fel eller en attack), eller en sekvens av paket som ser ut att vara i rätt format och tillsynes berör rätt mätvärden har matats in av en obehörig för att ersätta riktiga paket osv. Det senare (den felaktiga injiceringen) kan åstadkomma olämpliga styrreaktioner och skapa kaos eller t.o.m. framkalla felaktiga reaktioner från operatören som skadar processen (t.ex. släcker delar av elnätet) eller utrustningen (får enheten att slitas eller gå sönder så som gjordes i Stuxnet). Metoden bygger på att lära sig tidsegenskaper hos dataflöden och deras variationsmönster över tid. Flera maskininlärningsmetoder har kombinerats för att känna igen olika tidsbaserade mönster hos ett normalt dataflöde. Både för att känna igen kommunikation mellan enheter som fungerar genom pollning och genom spontana utskick. Det första används i implementationer där styrenheten frågar regelbundet efter värdet och sensormodulen svarar, och det andra används då sensorn är tilltänkt att meddela endast när värdet från fysiska processen över/understiger ett gränsvärde. Metoderna har testats på tre olika protokoll som är vanliga inom industriella styrsystem, nämligen Modbus, S7, och IEC-60870-5-104.[7,8]

Sammanfattning av arbetet inom detektion

Arbetet inom RICS har gett forskarna en djupare förståelse av anomalidetekteringsproblemet för SCADA system. Alla framtagna metoder har diskuterats på konferenser, granskats för tidskrifter, och citeras av andra forskare inom fältet internationellt. Men vad vi har lärt oss utöver detta är hur arbetsflödet för att samla riktigt data, skapa syntetiska data i emulerade labb som liknar riktiga miljöer, och hur känsligheten i systemen gör att en lång period för skapande av tillit är nödvändig för att samarbeta långsiktigt med industrin. Mer om dessa erfarenheter kommer i nästa avsnitt.

Tillämpningar i praktiken

Forskningmodeller kan vara intressanta för en vidare analys och förståelse av systemen, men ett mål med RICS har varit att samverka och påverka avnämare med riktiga data och riktiga system. Dock visade det sig svårare än förväntat att få tillgång till intressant data, och än mer utmanande att faktiskt testa metoderna i riktiga miljöer trots att många aktörer visade ett stort intresse för forskningen och möjlig avknoppning med tillämpningar.

Ett problem med samhällskritisk infrastruktur är just att den är viktig och särskilt attraktiv för angripare. Därav skyddas i många fall dessa system så att själva topologin, typ av system, och data endast är tillgängliga för organisationen som

underhåller systemet. Ett universitet å andra sidan arbetar oftast på andra principer där det är viktigt att dokumentera arbetsmetodik och resultat så att andra kan ta del av dem. Samtidigt är validering av metoderna endast övertygande om data som liknar riktig data används.

Inom RICS hade vi två spår för att kunna utvärdera algoritmer på relevanta system. Eftersom vi förutsåg att det skulle ta tid att extrahera data ur riktiga system, arbetade vi tillsammans med FOI och ABB för att bygga en virtuell testmiljö som kunde användas för validering. Vi har beskrivit RICS-el, en emuleringsmiljö för att validera algoritmer i en vetenskaplig publikation[9]. Dock behövde den syntetiska skapade datamängden kompletteras med data från riktiga system.

Under de första forskningsåren inom RICS möttes många avnämare med olika typer av system men det fanns alltid en konflikt i att systemägaren ville kontrollera hur data skulle användas mot behovet i forskningen att kunna redovisa resultat öppet. Vi hade dock turen att tidigt komma i kontakt med två avnämare som förstod problemet, och speciellt att om vi inte kan prova algoritmer i riktiga miljöer kan vi inte säga hur effektiva de verkligen kommer att vara. Göteborg kretslopp och vatten samt Modio, ett företag som specialiserar sig på byggnadsventilationsstyrning, var våra första avnämare som delade data med oss fritt för forskningsändamål inom RICS. Samtidigt deltog vi i ett svenskt stormöte och mässa (4SICS som senare döptes om till CS3 STHLM) där många aktörer inom kritisk infrastruktur närvarar. I det sammanhanget sätts praktiska labbmiljöer upp där korta demonstrationer ges. Vi fick därmed data som samlades inom ett Siemens demonätverk i detta sammanhang och utövade då våra första testanalyser.

I fallet med Göteborg kretslopp och vatten hade vi två fördelar. För det första är deras process (vattendistribution) inte hemlig i motsats till många andra tillämpningsområden. Dessa system och algoritmer är välkända. För det andra planerade de ett större systemunderhåll där stora delar av deras nätverk skulle ändras. Detta betydde att de kunde samla in data i sitt system (som av sin natur inte var känslig), och sedan ge den till oss efter att de hade uppgraderat sin infrastruktur. Då kunde inte längre någon information läcka ut eftersom systemet inte längre fanns. Förutom att visa för svenska avnämare att RICS metoder kunde fungera med riktiga data från en välkänd operatör gav också denna data en fördel inom forskningen. Metoderna inom RICS var teoretiskt väl beskrivna. Med detta kunde vi också påvisa att de fungerade med riktiga data. Även användning av Modios data var ett plus då vi kunde sprida forskningsresultaten vid relevanta konferenser av hög renommé.

Under år tre av projektet fick vi en tredje källa av riktigt data för anomalidetektering med hjälp av tidsegenskaper genom en eldistributör i Sverige. I detta samarbete använde vi en ny metod. Företaget fick kod från forskarna som de fick studera och sedan använda inne i deras system för datasamling. Koden tog fram filtrerad data som kunde användas för modellinlärning som senare användes

för anomalidetektering, men denna filtrerade datamängd avslöjar inga känsliga attribut hos systemet.

Efter att vi framgångsrikt provat algoritmer med data extraherat från svenska avnämare ville vi ta nästa steg: kunde vi låta systemet köra i realtid direkt i ett system? Tillsammans med studenter kontaktade vi ett pappersbruk nära Göteborg. De har styrsystem för att kontrollera sina processer. Eftersom vi redan hade utvärderat algoritmerna med en rik datamängd valde vi här att köra metoden PASAD beskriven ovan direkt hos dem, installerad på en liten dator kopplad till deras nätverk. För att lösa problemet med möjlig känslig data exporterade vi aldrig rådata, utan bara alarm från vårt system (som i sin tur kunde inspekteras av avnämaren så att inget känsligt lämnade deras bruk). Vi kunde visa att våra algoritmer kunde köra under en längre period i en riktig miljö, och därmed att det mycket väl kan fungera som ett riktigt system.

Slutligen ville vi undersöka om våra metoder direkt kunde verka i de facto ledande SCADA system på marknaden. Genom ABB och ett studentarbete kunde vi anpassa metoden PASAD till att vara en modul i ABBs programvara, vilket skulle kunna innebära att framtida kunder som köper ett SCADA-system direkt kan välja om de vill övervaka en process med vår metod mot cyberhot. Försöket fungerade väl och metoden kunde implementeras och köras i ABBs programvara.

Slutsatser och lärdomar från RICS

Tekniskt har vi fått belägg för att intrång i OT system kräver speciella metoder. De vanliga/kommersiella verktygen inom IT är inte tillämpbara. Utvecklade RICS-metoder verkar till och med kunna bäddas in i verkliga system och ge mervärde till systemägarna. Vi har även lärt oss att riskbedömning inom kritisk infrastruktur och tillhörande verktyg kan ha flera olika nyanser och användningsområden.

RICS har skapat kompetens inom cybersäkerhet för samhällskritiska system genom att forskarutbilda och examinera två doktorander samt flera master- och kandidatstudenter, anordnat en doktorandkurs för flertalet doktorander nationellt, tillfört innehåll om dessa typ av system i våra grundutbildningskurser, verkat för att en kurs om etisk hackning skapats, och anordnat flertal seminarier och internationella konferenser. Men också skapat samarbetsytor med näringslivet som annars inte hade funnits.

Mer att läsa finns på RICS webbsida www.rics.se under publikationer, och artiklarna kan fås genom att kontakta projektets forskare om man inte kommer åt hos förlagen.

Referenser

1. https://www.sergid.eu
2. https://energyinformatics.springeropen.com/articles/10.1186/s42162-018-0010-x
3. https://www.mal-lang.org
4. https://idependeu.herokuapp.com
5. https://www.chalmers.se/sv/institutioner/cse/nyheter/Sidor/PASAD-.aspx
6. https://dl.acm.org/doi/10.1145/3243734.3243781
7. https://link.springer.com/chapter/10.1007%2F978-3-319-99843-5_5
8. https://www.usenix.org/conference/raid2019/presentation/lin
9. https://www.springer.com/gp/book/9783030058487

I samarbete med:

651 81 Karlstad Tel 0771-240 240 **www.msb.se**
Publ.nr MSB1701, februari 2021 ISBN 978-91-7927-109-1