Att upptäcka pågående cyberintrång inom kritisk infrastruktur

Digitaliseringen påverkar alla sektorer i samhället och för med sig många fördelar. Vissa system blir mer kostnadseffektiva, medan i andra kan man utnyttja deras kapacitet mera och köra dem närmare deras maxkapacitet. 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 från andra länder. Även fastän dessa problem är viktiga för alla typer av system, är det av extrem betydelse för samhällskritiska system av främst två anledningar. Först kan man konstatera att sådana system oftast används länge (20-tals år), vilket medför att de system vi har idag byggdes utan avancerad cybersäkerhet. Sedan kan man också konstatera att konsekvenserna kan bli stora om de fallerar. Utöver dessa möjliga fjärrattacker kompliceras bilden av möjligheten för illasinnade aktörer som agerar inne i organisationen med avsikt att påverka systemets tjänster.

Vi har sett ett ökande antal attacker mot samhällskritiska system. Några av de bättre dokumenterade fallen är de följande. Ett av de första väldokumenterade intrången mot vattenrening infrastruktur skedde i Maroochy Shire i Australien 2000 (det rörde en så-kallad *insider).* Ett decennium senare såg världen ett fall av cyberkrigsföring och med den skadliga koden Stuxnet där attackens avsikt var att stoppa anrikningen av uran i Iran. Det är ett av de första exemplen på kod som specifikt riktar sig mot avancerade styrsystem, och som dessutom involverade icketekniska element (social engineering som gav tillgång till isolerade nätverk). Några år sedan attackerades elnätet i Ukraina med stora strömavbrott som följd, inte en gång utan två år i rad. Också detta har dokumenterats och analyserats i detalj, men attackerna tydliggör att också samhällskritiska system som används av alla (elnätet) numera kan attackeras.

Samtidigt som cyberskydd utvecklas kontinuerlig, är de flesta av dessa system inriktade mot vanliga IT system 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 några av egenskaper som vi eftersträvar inom IT system (tillgänglighet, konfidentialitet) får andra dimensioner i system där kärnan är att styra fysiska processer. Man kan till exempel fundera om ett system som är ostyrbar och därmed kan få den fysiska processen bete sig utan kontroll är tillgänglig eller inte.

Ett av målen med vårt projekt, Resilient Information and Control Systems (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 de kända attacker ser ut och systemet larmar när sådana mönster hittas i systemet (missbruksdetektering). I det andra definierar man hur normalbeteendet för ett system ser ut, och i det fallet larmar man när systemet verkar uppföra sig annorlunda mot det normala (avvikelsedetektering). Det säger självt att missbruksdetektering är bra för att notera när kända attacker är på gång. Medan anomalidetektering krävs om hittills okända attacker ska detekteras.

Tyvärr fungerar anomalidetektering inte så bra på vanliga IT system eftersom de används på så många olika sätt och det är svårt att skapa en generell normalmodell. Samhällskritiska system, däremot, har ofta ett mer definierat normalbeteende i och med att i dessa system består av styrenheter (så-kallade operationstekniska delen eller OT) vars mål är att styra en process som beter sig enligt fysiska lagar, även om dessa kan kommunicera med andra system inklusive IT enheter i kontorsmiljön. Detta har inneburit att avvikelsedetektering ses som mycket lovande för OT delen av dessa system.

Genomgång av alternativa möjligheter

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 styrsytem 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 övrigt programvarusystem: 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 fokuseras 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. En insider hot innebär att man måste samla data på ställen (och därmed den sorts data) som är svårt att påverka även för den som har rättigheter att komma åt systemet, t.ex. fysiska systemets olika mätvärden i relation till varandra. En attackerare som i realtid ändrar på flera mätvärden i harmoni med fysiska lagar har 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 elller 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 paketet (så-kallade "payload"). Ett specialfall av protokollspecifika delar i varje paket är den tidstämpel som anger när detta paket skapades, dvs i anslutning av vilket tid avlästes något mätvärde i processen eller skickades en styrsignal av operatören för att ändra i styrningen.

Anomalidetektering inom RICS projektet

Under RICS-projektet 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 är 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. Metoder konstruerade av andra forskare har också byggt liknande system, men de flesta av dessa använder den insamlade datan för att förutsäga framtiden genom modellen som bygger på datat. Om sedan nästa värde från systemet skiljer sig från denna prognos skapas ett alarm. Det är dock svårt att förutsäga framtiden och dessa prognoser innehåller mycket fel, som en attackerare kan gömma sina attacker i. RICS metod som är döpt till PASAD 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.

Den andra RICS metoden kan antyda andra avikande fenomen, t.ex. att någon beräkningsmodul i systemet agerar långsammare än normalt (är överbelastat p.g.a. en oavsiktligt fel eller en attack), en sekvens av paket som ser ut att vara i rätt format och tillsynes berör rätt mätvärden har matats in av en obehörig istället för riktiga paket osv. Det senare (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 det gjorde i stuxnet fallet). 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. Båda för att känna igen kommunikation mellan enheter som fungerar genom pollning (styrenheten frågar regelbundet efter värdet och sensormodulen svarar), och genom spontana utskick. Det senare avvänds t.ex. för att SCADA master ska varslas om avvikande processvärde i den fysiska processen eller OT systemet. Metoderna har testats på tre olika protokoll som är vanliga inom industriella styrsystem, nämligen Modbus, S7, och IEC-60870-5-104.

Båda metoder har testats på flera sorter data och visat sig kunna igenkänna riktiga avvikelser utan att skapa för många falska larm.

Tillämpningar i praktiken

Forskningsmodeller 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 riktigt 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. Dock behövde det syntetiska datat 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 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 i byggnadsventilationsstyrning, var våra första avnämare som delade data med oss fritt för forskningsändamål. I samma väva 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 data som samlades inom en Siemens testnätverk i detta sammanhang och utövade 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 en plus då vi kunde sprida forskningsresultaten vid relevanta konferenser av högt renommé.

Under år 3 av projektet fick vi en tredje källa av riktigt data för anomalidetektering med hjälp av tidsegenskaper var 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 filtererade data som kunde användas för modellinlärning som senare används för anomalidetektering, men det filtererade datat avslöjar inga känsliga attribut av 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 ett rikt datamängd valde vi här att köra metoden som baseras på spektralanalys direkt hos dem, installerad på en liten dator kopplad till deras nätverk. För att lösa problemet med möjliga känsligt 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, lite som att det faktiskt var ett riktigt system.

Slutligen ville vi undersöka om vi 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 hos ABB skulle direkt kunna välja om de vill övervaka en process med vår metod mot cyberhot. Försöket fungerade väl och metoden kunde implementeras och köras i ABBs programvara.

Sammanfattning

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.

Tekniskt har vi fått belägg för att intrång i OT system kräver speciella metoder. De vanliga/kommersiella verktyg inom IT är inte tillämpbara. Metoder som vi har skapat har en bra chans att kunna bäddas in i verkliga system och ge mervärde till systemägarna.

RICS har skapat både kompetens inom intrångsdetektering för cyberfysiska system genom att forskarutbilda och examinera två doktorander samt flera master- och kandidatstudenter, men också skapat samarbetsytor med näringslivet som annars kunde inte ha funnits.

Mer att läsa finns på RICS webbsida www.rics.se under publikationer, och genom att kontakta projektets forskare.