**Who's in Control of Your Control System?**

**Device Fingerprinting for Cyber-Physical Systems**

**הקדמה**

מערכות בקרה תעשייתיות (ICS) משמשות לשליטה ובקרה על תשתיות קריטיות כגון חשמל, גז, ביוב ומים. אופיין המורכב של מערכות אלה מהווה אתגר אבטחתי גדול. מחד, בשל גודלן הפיזי של מערכות אלו, אשר יכול להגיע לגדלים עצומים כמו רשת החשמל או המים, ומאידך, בפן התוכנה, כאשר אלו מכילות רכיבים ישנים אשר לא ניתן לעדכנם או להוסיף אליהם רכיבי הצפנה ואבטחה מתקדמת. מצב זה, נובע כיוון שהסיכון בהוספת רכיבים או עדכונים אל המערכות הקיימות יוצר חשש כי מערכות אלו יקרסו כתוצאה של תקלה או מחסור בכוח עיבוד וישפיעו על חייהם של עשרות או מאות אלפי בני אדם. מצד שני, חוסר היכולת להגן על מערכות אלו, יכול להסתיים גם הוא במצב קטסטרופאלי עד כדי סיכון חיי אדם.

השיטה המוצגת במאמר מגיעה בכדי לפתור את המצב הקיים, ולהציע פתרון אבטחתי מותאם אישית למערכות השליטה התעשייתיות. השיטה מתבססת על יצירת "טביעות האצבע" (Device Fingerprinting) הן של מידע המועבר ברשת והן של מידע מרכיבים פיזיים. יש לציין כי בעוד שנושא זה נחקר לעומק, אף פתרון אינו מציע שיטת "טביעות אצבע" המותאמת למערכות מסוג זה.

רוב השיטות הנפוצות כיום לעבודה עם "טביעות אצבע" או "חתימות מכשיר" משתמשות בשיטת גילוי אקטיבית ,אשר, אמנם מגיעה לאחוזי זיהוי גבוהים אך עם זאת זקוקה לשליחה של כמות תעבורה גדולה יחסית ברשת, דבר שעלול לסכן את התשתית הקיימת ולהסתיים בתרחישי אסון. שיטות גילוי פאסיביות, מתוקף היותן "שקטות" יחסית ולא דורשות התייחסות מצד המכשיר הנבדק, אמנם יותר מתאימות למערכות שליטה תעשייתיות אך עם זאת מספקות כמות מוגבלת של מידע או שדורשות הוספת ציוד מיוחד.

קיימים מספר סוגים שונים של מערכות בקרה תעשייתיות, העיקרית שבהם היא SCADA – Supervisory Control and Data Acquisition. הפונקציות העיקריות במערכות בקרה תעשייתיות הן השגת מידע (Data acquisition) ושליטה (Control) אשר מבוצעות ע"י פרוטוקולי SCADA כגון DNP3 ו- Mudbus.

השיטה הראשונה המוצגת, יוצרת את "טביעות האצבע" ע"י ניצול אופיין הייחודי של פונקציות השגת המידע על מנת לייצר חתימה המבוססת על מעבר של חבילות (Packets) משכבת התעבורה אל שכבת האפליקציה.

השיטה השנייה המוצגת, מנצלת את פונקציות השליטה, על מנת לייצר חתימה המבוססת על אופיים של ההתקנים במערכות SCADA ועל מאפיינים פיזיקליים ייחודיים בהתקנים אלו.

שילוב שיטות אלה משיג "טביעת אצבע" המבוססת על החומרה, התוכנה והיבטים פיזיקליים שונים, ומספק תוסף נאה לאבטחה על מערכות בקרה תעשייתיות.

**היכרות עם Device Fingerprinting ו- ICS**

מונחים שיש להבהיר להמשך:

ICS – מערכות שליטה ובקרה תעשייתיות

מונח כללי המתייחס לסוגים שונים של מערכות המשמשות בייצור תעשייתי, שתי המערכות הנפוצות כיום הן SCADA – Supervisory Control And Data Acquisition ו- DCS – Distributed Control System. כמו כן, קיימות גם מערכות בקרה קטנות יותר כגון בקרים עם לוגיקה שניתן לתכנת PLCs – Programmable Logic Controllers (המופיעים כרכיבים במערכות SCADA הגדולות יותר).

מערכות בקרה תעשייתיות משמשות לרוב בתעשיות החשמל, מים, נפז, גז ועוד. באמצעות המידע המתקבל ונשלח מתחנות רחוקות, בין אם נשלח בצורה אוטומטית או ע"י חיווי משגיח, ניתן לפקוד על ביצוע פעולות ברכיבים מרוחקים כמו פתיחת שסתומים או מעצורים. בנוסף ניתן לנתר מרחוק את המערכת כולה על ידי חיווים המתקבלים מתחנות מרוחקות.

Device Fingerprinting

שיטה בה נאסף מידע אודות מחשב/מכשיר מרוחק, למטרות זיהוי. שיטה זו ניתנת לשימוש על מנת לזהות בצורה מלאה או חלקית משתמשים והתקנים ברשת, אפילו אם אלו אינם משתמשים ב"עוגיות" (Cookies) או לא הזדהו באופן רשמי. המוטיבציה לשימוש בשיטה זו נובע מערכו המשפטי של טביעת אצבע אנושית. במקרה "האידיאלי", לכל המשתמשים ברשת תהייה טביעת אצבע שונה, וזו לא תשתנה אף פעם. תחת ההנחות הללו, מגוון (Diversity) ויציבות (Stability) המידע, ניתן היה לסווג בצורה ייחודית את כל המכונות ברשת ללא הסכמה מפורשת של המשתמש עצמו.

אך בפועל, מגוון ויציבות המידע אינן לגמרי ברות השגה, ניתן אף לומר כי קיים יחס הפוך בין שני התכונות הללו.

מגוון (*Diversity*)- תכונה זו דורשת כי לא יתכנו שתי מכונות עם חתימה זהה, אך ברור כי ישנן מספר רב של מכונות אשר חתימתן זהה. על מנת להתגבר על כך, ניתן להשתמש במספר גדול יותר של מאפיינים המתארים את המכונה בתהליך ייצור החתימה ובכך להגיע לדיוק מקסימלי. יחד עם זאת מספר רב של תכונות אינו מבטיח יציבות יתרה כיוון שעם עליית מספר המאפיינים המרכיבים את החתימה, יורדים הסיכויים שכל המאפיינים יישארו קבועים ובכך היציבות פוחתת.

יציבות (*stability*)- תכונה זו דורשת כי החתימה תישאר זהה במשך הזמן, אך מובן שמצב זה לא יכול להתקיים כיוון שרוב התכונות (הגדרות תוכנה ו/או חומרה) ניתנות לשינוי. על מנת להתגבר על כך, ניתן להפחית בכמות המאפיינים המעורבים בתהליך יצירת החתימה ובכך להבטיח משתנים יציבים לאורך זמן, אך הפחתה זו תגרום לירידה במגוון.

היבט נוסף שיש להתייחס אליו הוא דרך איסוף המידע. קיימות שתי דרכים מרכזיות לאיסוף המידע לטובת יצירת "טביעת אצבע":

השיטה האקטיבית - שימוש בשיטה זו מתבסס על ההנחה כי המכונה תאפשר שליחת שאילתות בסיסיות אליה, וכך למשוך מידע רלוונטי ל"טביעת האצבע". דוגמא לשימוש בשיטה זו היא ע"י Nmap.

השיטה הפסיבית – שיטה זו אינה מתבססת על שליחת שאילתות למכונה בצורה אקטיבית אלא על סיווג עפ"י פקטורים כמו קונפיגורציה של TCP/IP או של מערכת ההפעלה או רכיבי חומרה במכשיר, ע"י ניתוח חבילות שהמכונה שולחת ממילא.

**חשיבות אבטחת סביבת ICS**

יש להדגיש את חשיבות המערכת המוצגת במאמר זה. נושא זה עולה יותר למודעות בשל עלייה בכמות ואיכות מתקפות הסייבר על תשתיות כמו רשת החשמל מערכות בנקים ועוד. בהיעדר הגנה מספקת, יכול האקר להתחבר אל מערכות תשתית קריטיות ולגרום לאסון כבד. ארה"ב הקימה צוות תגובה מיוחד עבור מקרים של התקפות סייבר על מערכות אלו ICS-CSERT – Industrial Control System – Cyber Emergency Response Team אשר פועל בכדי לצמצם את הסיכונים הקיימים במערכות אלו. ניתן למצוא מגוון דוגמאות להתקפות על מערכות מסוג זה, ביניהם הפסקות חשמל ענקיות, זיהום מי השתייה ועוד, אך הגדולה מבין התקפות אלו היא ללא ספק Stuxnet. Stuxnet היא תולעת שנועדה לפגוע במערכות שליטה ובקרה תעשייתיות, בפרט כאלו שיוצרו ע"י חברתSiemens **.** לתולעת יכולת לתכנת מחדש בקר לוגי מתכנת PLC)). זו התולעת הראשונה שנתגלתה עם RootKit וככל הנראה גם הראשונה המיועדת למערכות בקרה תעשייתיות. מטרת תולעת זו היא לפרוץ אל מערכות הבקרה התעשייתיות ודרכן לחבל במתקנים המבוקרים ע"י מערכת זו.

**מאפייני מערכות ICS  
המכשיר**למכשירי מערכות שליטה תעשייתיות יש חומרה וארכיטקטורת תוכנה פשוטים יחסית למחשבים עבור צרכים רגילים מכיוון שהם בנויים לבצע משימות מאוד מיוחדות וקריטיות ומעט מאוד משימות אחרות. מחשב יומיומי טיפוסי בימינו, בעל מעבד מרובה ליבות בטווח של 2-3 GHz עם מטמון (caching) משמעותי, ג'יגות של RAM ויכולת להחלפת הקשרים (Context switching) בין מגוון רחב של תהליכים שרצים על המכונה.

מהצד השני, עולם מערכות השליטה התעשייתיות נשלט ע"י יחידות לוגיות הניתנות לתכנות (PLCs) שרצות על מעבדים בעלי הספק נמוך בעשרות עד מאוד MHz, כמעט בלי מטמון או בכלל בלי, עשרות עד מאוד מגה בתים של RAM ומעט מאוד תהליכים.

עם כח עיבוד כל כך מוגבל, שינויים קלים בתכנות גורמים להפרשי זמן ניכרים הניתנים לאבחנה. בהתבסס על המשימה הנדרשת, מכשירים שונים בנויים עם מאפייני חומרה שונים, כמו גם עם תוכנה שונה (מערכת הפעלה, מימוש פרוטוקולי מחסנית, מספר מדידות שנלקחות, מורכבות יחידת השליטה ועוד). עובדה זו גורמת לכך שכל מכשיר מסוגל לעבד בקשות במהירויות שונות.

**שכבת הרשת**  
ברשתות תאגידיות מסורתיות, טלפונים ניידים ומחשבים ניידים נעים כל הזמן ומתחברים לנקודות הגישה השונות. התעבורה שהם מייצרים 'מטיילת' מרחקים עצומים, פוגשת נתבים בדרך שחווים רמות עומס בלתי צפויות ולעולם לא מובטח שחבילות עוקבות יעברו את אותו המסלול דרך הרשת. לעומת זאת, מכשירים במערכות שליטה תעשייתיות מוקדשים לביצוע משימה קריטית אחת וקבועים במיקום פיזי אחד. התעבורה שהם מייצרים 'מטיילת' מרחק קצר מאוד יחסית (גיאוגרפית) דרך רשתות בעלות ארכיטקטורה כזו שמציעה שינויי מסלול מעטים מאוד עד ללא כאלו.   
כמעט תמיד במערכות אלה רוחב הפס מוקצה מראש כך שייעשה שימוש בחלק קטן ממנו. למשל בתחנת הכוח בה נערך הניסוי, קיים רוחב פס של 100MBPs ואילו בפועל המערכת עושה שימוש רק ב-11KBPS, 0.01% מתוך רוחב הפס הקיים. עוצמות רוחב הפס הנמוכות הללו מבטיחות שמתגים ונתבים ברשת לעולם לא יהיו עמוסים מדי.

**מתודולוגיה**

הגישה המוצעת במאמר זה, מתבססת על נתונים ודגימות שנלקחו מרשת חשמל, אך ניתן בקלות להשליכן על כל מערכת שליטה תעשייתית אחרת.

אחד מאתגרי האבטחה המרכזיים במערכות שליטה תעשייתיות הוא השטח העצום שיש לאבטח, זאת עקב אופיין המבוזר של מערכות אלו. לדוגמא, שטח רשת החשמל, שעליה התבסס הניסוי, התפרש על פני 7500 קמ"ר, שליש משטחה של מדינת ישראל. מכיוון ששטח מערכות אלו מגיע לגדלים עצומים, אנו מסווגים את סוגי התוקפים לשתיים:

1. תוקף חיצוני שאין לו גישה פיזית למתקן אך השיג שליטה מבחוץ ע"י נוזקה, על אחד מן הרכיבים ברשת או במערכת עצמה.
2. תוקף אשר יש בידיו גישה פיזית לרשת והוא מחובר אליה באמצעות מחשב נייד סטנדרטי.

כשמדובר ברשת חשמל, קיימים שתי אזורי תקיפה עיקריים. תוקף יכול לנצל את (1) רשת התקשורת לטובת ההתקפה, או (2) רכיב פיזי המוטמע במערכת. בהתאם למיקום הפריצה במערכת, יכול התוקף להזריק מידע, תשובה שגויה לרשת או פקודות שגויות.

מטרת המחקר היא לפתח שיטת "טביעת אצבע" מדויקת, אשר יודעת לזהות את סוג המכשיר ממנו נשלחה חבילה ברשת. כלומר, היכולת להבחין האם החבילה נשלחה ע"י מחשב נייד של תוקף או ע"י אחד מרכיבי המערכת (IED- Intelligent Electronic Device).

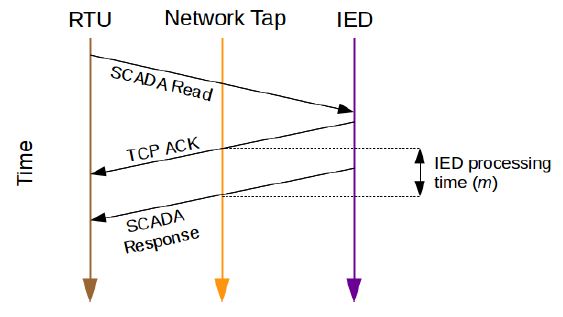
הגדרה רשמית:

נניח כי G הוא מערך המכיל את כל מערכות הבקרה התעשייתיות, ומורכב מ כך ש j מזהה את מותג המכשיר ו- k מייצג את הדגם הספציפי של המכשיר. בהינתן רצף של תצפיות, של כל מכשיר i ברשת. מטרת שיטת "טביעת האצבע" היא לזהות לאיזו תת קבוצה של G, או ליתר דיוק, לאיזה שייכות התצפיות.

כפי שצוין קיימים שתי מאפיינים עיקריים ברשתות של מערכות שליטה תעשייתיות: פונקציות השליטה (Control) והשגת המידע (Data Acquisition), פונקציות אלו מוגדרות בכל מערכת שליטה תעשייתית, ללא קשר לרכיבים הפיזיים המפוקחים (והמופעלים) ע"י המערכת ולכן שיטות "טביעת האצבע" המוצגות, אל אף שמתייחסות לרשת חשמל, נכונות עבר כל מערכת אחרת. השיטות המוצגות מנצלות את אופיין הייחודי של פונקציות אלו. השיטה הראשונה המנצלת את פונקציית משיכת המידע, נחקרה עפ"י מידע אותנטי מרשת חשמל ואושררה מול מידע ניסויי. השיטה השנייה, המנצלת את פונקציית השליטה נבחנה אך ורק עם מידע ניסויי, זאת עקב שימוש מזערי בפונקציה זו ברשת החשמל.

**Cross Layer Fingerprinting**

השיטה הראשונה פונה אל פונקציית רכישת המידע במערכות SCADA, השיטה מבססת את חתימתה על האינטראקציה בין שכבת האפליקציה לשכבת התעבורה מרגע שליחת בקשה למשיכת מידע בשכבת האפליקציה ועד להודעה (Acknowledge) משכבת התעבורה. השיטה מודדת את הזמנים הללו על מנת לחשב את זמן העיבוד של המכשיר (IED) וליצור טביעת אצבע ייחודית לכל מכשיר במערכת בהתבסס על הפיזור של זמנים אלו. הזמן הנמדד מתגובת שכבות הרשת יקרא CLRT – Cross Layer Response Time. ומחושב עפ"י התרשים הבא:



הגדרה רשמית של החתימה:

בהינתן היסטוגרמה של מדידות הזמן (CLRT), תהי M קבוצה של מדידות זמן (CLRT) ממכשיר ספציפי. B מגדיר את מספר העמודות בהיסטוגרמת מדידות הזמן, ושווה למספר המאפיינים בווקטור "טביעת האצבע". H הוא הסף היוריסטי המייצג את ערך המקסימום הגלובלי המוערך עבור מדידת זמן CLRT. אנו מחלקים את טווח הערכים בהיסטוגרמה עפ"י המשוואה הבאה: , ונגדיר כל אלמנט בווקטור טביעת האצבע עפ"י המשוואה הבאה:

השיטה המוצעת מתבססת על ההנחה כי מדידות הזמן (CLRT) הינן קבועות למכשיר ספציפי וכי הן תשמורנה על פיזור ייחודי במרוצת הזמן. הנחה זאת נובעת מהשימוש הנמוך בכוח חישובי וקצבי עיבוד נמוכים וכן רוחב פס עצום ביחס לשימוש המזערי בהעברת נתונים. בנוסף, זמן המדידה עצמאי ואינו תלוי ב-RTT מכיוון שהמדידה מתייחסת לשליחת שתי חבילות עוקבות מאותו המקור לאותו היעד.

כל המידע הניסויי חולק ליחידות זמן (יום, שעה וכו') וכל יחידת זמן חולקה להיסטוגרמה של 200 עמודות וחושבו השונות והממוצע של המדידות בטווח הזמן. מהחישובים שנעשו ניתן היה לראות כי אפילו באמצעות שיטות מדידה בסיסיות הגיעו החוקרים להפרדה גבוהה בין הדגימות של כל מכשיר שנבדק. יתר על כן, ניתן היה לראות כי ניתן להגיע להפרדה בין שתי מכשירים זהים על קונפיגורציית התוכנה שלהם. כיוון שההפרדה בין המכשירים כל כך איכותית, כל שימוש באלגוריתם למידה סטנדרטי צריך להניב תוצאות מרשימות, כלומר סיווג באחוזי דיוק גבוהים.

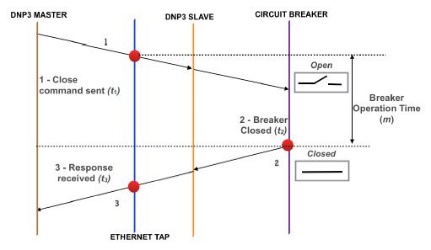
מכיוון שהמטרה והחדשנות במחקר זה אינם טמונים בשימוש בלמידת מכונה בתחום, האלגוריתמים הפופולריים ביותר בתחום נבחרו כדוגמאות. במחקר זה השתמשו החוקרים בשתי אלגוריתמי למידה, רשתות נוירונים מלאכותיות ומסווג בייס. מאחר ושתי השיטות הניבו אחוזי דיוק גבוהים, נפרט כאן רק על רשתות נוירונים. תחילה נבחר ווקטור מאפיינים (Feature vector) עפ"י משוואה 1, הדוגמאות חולקו ל 75% לאימון ו- 25% לאימות והתוצאות נמדדו עפ"י פונקציות Accuracy, Precision ו- Recall.  
עפ"י הממצאים ניתן היה לראות כי אפילו דגימות שנלקחו מפרק זמן של 5 דקות, עמדו על 93% דיוק. מכיוון שחלק מהתחנות ייצרו מידע אחת ל 2 דקות, ניתן לומר כי האלגוריתם יוכל להגיע להכרעה תוך 2 דגימות של מידע. יתרה מכך, ברוב מערכות השליטה התעשייתיות לא ייגרם נזק קטסטרופאלי לאחר הזרקת פקודה אחת, לרובן מערכות הגנה מובנות, לכן רוב ההתקפות המתוחכמות ייצרו נזק רק לאחר תקופת זמן ממושכת, למשל על ידי שיטוי במערכת הבקרה וחימום כור לטמפרטורות גבוהות עד לפיצוץ.

בהמשך, בוצע סיווג גם עפ"י מסווג בייס שגם הראה אחוזי דיוק גבוהים במיוחד, אך יחד עם זאת נשאלת השאלה: מה יקרה אם אין בידינו מספיק נתונים בכדי לבנות מכונת למידה מונחית? האם נוכל לסווג את רכיבי המערכת ע"י מכונת למידה לא מונחית? נראה שכן, מן הדגימות שנלקחו, ניתן היה לראות כי ישנה התפלגות גאוסיינית של הנתונים, ועל כן החוקרים החליטו להשתמש באלגוריתם אשכול , ניתן היה לראות, כי האשכול המקורי של הנתונים שנאספו, אשר הראו הפרדה מוחלטת, היה כמעט זהה לאשכול הנוצר ע"י האלגוריתם. כשנמדדו הביצועים של המודל, נמצא Accuracy של 92.86%,Precision של 0.891 ו- Recall של 0.956.

**Physical Fingerprinting**

שיטה זו פונה אל פונקציית השליטה במערכות SCADA, ע"י יצירת טביעת האצבע באמצעות מאפיינים פיזיקליים ייחודים ברכיבים עצמם. השיטה משתמשת באותה המשוואה ליצירת החתימה (1), כאשר M היא קבוצת מדידות זמנים של פעולות מכשירים ו- H הוא הסף היוריסטי המקסימלי לפעולות אלו.

ההנחה היא, שזמן הפעולה של מכשירי המערכת משתנה ממכשיר למכשיר ומיצרן ליצרן, וכך מייצר טביעת אצבע מושלמת. כאשר נשלחת פקודה אל אחד מרכיבי המערכת, ממסר הרכיב מחזיר תגובה לא מסונכרנת (asynchronous) לגבי האירוע שהתרחש, אשר ניתן להסניף דרך הרשת ולחשב את זמן הפעולה של המכשיר. התגובה יכולה גם להכיל רצף של אירועים שהוקלטו ((SER- Sequence of event recorder, לכן זמני הפעולה יכולים להימדד על ידי שתי שיטות:

1. חתימות זמן לתגובות לא מוכרות / מוזמנות- מחושבות על ידי מערכת ההפעלה בנקודת ההאזנה, תוך שימוש בהפרש בין הזמן בו הפקודה הובחנה לבין הזמן בו התגובה הובחנה: m=t3-t1.
2. **חתימות זמן SER – מחושבות מההפרש בין הזמן בו הפקודה הובחנה בנקודת ההאזנה לרשת לבין חתימת הזמן של המאורע בשכבת האפליקציה: m=t2-t1.

לצורך הניסוי, נמדדו הזמנים של פעולת סגירת ופתיחת זרם חשמלי במכשירי המערכת, בחיבור שבין ה-Slave ל-Master. כפי שניתן לראות מהתרשים הוצב מחשב שהסניף את התעבורה וביצע ניתוח מעמיק שלה. באותה נקודת ההאזנה, לחבילות ניתנות חתימות זמן על ידי מערכת ההפעלה לינוקס אשר מסונכרנת מבחינת זמנים לאותו הזמן של ה-Master וה-Slave. המכשיר ממנו נלקחו הזמנים, תומך בפרוטוקולי-Modbus, DNP3 , פרוטוקול SNTP ופרוטוקול SER שמעניק חתימות זמן ברזולוציה של מילי-שניות.

בפקודת סגירת המעגל מהמאסטר, המכשיר מייצר פלט בינארי המניע (טוען) את סליל הנעילה כדי לסגור את המעגל. ברגע בו מעגל הטעינה נטען, הקלט הבינארי חש את השינוי ונוצרת חתימת זמן.  
בפקודת'טעינה' מהמאסטר המכשיר מפעיל את הפלט הבינארי השני שיטען את סליל הטעינה מחדש (Reset) כדי לפתוח את מעגל הטעינה, דבר המתועד כאירוע בעל חתימת זמן.  
עבור ניסויים אלו, נמדדו 2500 פעולות פתיחה וסגירה על גבי פרוטוקול DNP3. הפקודות והתגובות הוקלטו בנקודת ההאזנה המתוארת בתרשים וזמני הפעולה חושבו באמצעות שתי השיטות (SER ו- unsolicited response). שיטת התגובות הבלתי מוזמנות לא ייצרה תוצאות שמישות ואילו תוצאות שיטת ה-SER היו מרשימות, לכן אלו מתוארות להלן ותועדו כחתימות פיזיות.

Hardware

DNP3 Master

DNP3 Slave

Sniffer + parser

כאשר הופעלו אלגוריתמי הלמידה על דגימות ה SER שנאספו, אלגוריתם רשתות הנוירונים הגיע לאחוזי דיוק של 86% ומסווג בייס הגיע לדיוק של 92%. הביצועים שנמדדו, היו מפעולות סגירה של הזרם, אך ישנו צורך להבדיל גם בין הפעולות השונות בכל מכשיר, לשם כך, מדדו גם את סוגי הפעולות (סגירה ופתיחה של זרם) וגם כאן, נמצא פיזור אידיאלי של הנתונים.

חשוב לציין תחילה כי אופן איסוף הנתונים לצורך יצירת טביעת האצבע יכול התבצע ב 3 צורות:

הקופסא הלבנה - מודל דינמי של המכשיר נבנה ע"פ עקרונות ראשונים, פרמטרים של המודל שזוהו מתוך שרטוטי ה-CAD (computer-aided design), קוד המקור, מדידות פיזיות וכדומה. זאת מבלי להתבונן אפילו בדגימת מידע אחת של המערכת. נתונים אלו שבאמצעותם ניתן לחזות את ההתנהגות המצופה של המערכת, משמשים ליצירת טביעת האצבע ע"י שינוי ערכי המודל תוך שימוש בהתפלגות אי ודאות.

קופסא שחורה – טביעת האצבע נבנית ישירות מתוך מידע ניסויי בלי מודל דינאמי. הגישה דורשת פרק זמן ניכר עבור מדידות ואיסוף מידע, אך מעט מאוד מידע מקדים על המערכת. עד כה, שיטה זו הייתה השיטה היחידה בה נעשה שימוש בגישות השונות לעבודה עם עקבות מכשירים.

הקופסא האפורה- ראשית נוצר מודל דינמי של המערכת ולפיו נוצרת טביעת האצבע. אותה טביעת אצבע עוברת עיבוד ושיבוח בהתבסס על מדידות ניסוייות (למעשה שילוב שתי הגישות הקודמות).

לפיכך, נלקחו גם דגימות סינטטיות (מודל הקופסא הלבנה) על מנת להבטיח שבמצב שבו אין גישה לנתונים על מנת לייצר מודל "קופסא שחורה" נוכל להתבסס על ממודל "הקופסא הלבנה", אך גישה זו הניבה אחוזי דיוק של 80% ולכן התבססות על שיטה זו אינה רלוונטית. יחד עם זאת, במצב בו השגת דגימות מידע היא קשה (מכיוון שתהליכים מסוימים מתרחשים לעיתים נדירות מאוד בלבד, ניתן לשלב תוצאות ממודל הקופסא הלבנה והשחורה על מנת ליצור מודל קופסא אפורה ולהגיע לאחוזי דיוק גבוהים למדי.

**ביצועים**

כפי שכבר ציינו, על מנת ששיטת טביעת האצבע תהייה נכונה לכל מערכת, על הטכניקה להיות מדויקת (Accuracy) כמובן, ובנוסף עליה להיות ברת הרחבה (Scalability) . בעוד שאף אחת מן השיטות המוצעות לא הצליחה להשיג אחוזי דיוק מושלמים, שתיהן השיגו תוצאות טובות מספיק בכדי להפוך שימושיות עבור אסטרטגיית הגנה רוחבית, כתוספת וחיזוק לגישות IDS מסורתיות. גישת ה-CLRT השיגה נכונות סיווג מרשימה בגובה של 99% בחלק מהמקרים וגישת החתימות הפיזיות הצליחה לסווג בהצלחה של 92% מדידות של שתי מכשירים כמעט זהים. לצורך השוואה, הגישות הקודמות שבהן נעשה שימוש, השיגו נכונות סיווג שנעה בין 86% ל-100% כך שביצועי הגישות החדשות לא רחוקות מאוד. *ל*אלגוריתם רשת הנוירונים בו נעשה שימוש באימון שתי הטכניקות, היו רק שכבה נסתרת אחת ו-200 מאפייני קלט (Features) מה שמייצר חתימה ברת הרחבה סבירה לסיבוכיות כוח החישוב. מסווג בייס החלופי גם מאוד יעיל. יתרה מכך, התוצאות רומזות שנכונות התוצאות שהגישות מספקות גוברות גם הן ככל שמגדילים את כמות הדגימות.

כאשר משתמשים בחתימות מכשירים כדי להעצים שיטות IDS, רצוי גם שהחתימות לא יהיו קלות לזיוף. למרבה המזל, קיימות מספר סיבות לכך שהשיטות המוצעות לא כל כך קלות לשבירה. ראשית, תמיד קיימת ירושה של אקראיות במכונה של התוקף, שתגרום ליצירת כל דבר המבוסס על דיוק תזמון, להפוך ללא טריוויאלי. שנית, בשיטת החתימות הפיזיות, על שעון מכונת התוקף להיות ולהישאר מסונכרן לשעון של מכשיר המטרה, ברמת דיוק של מילי-שניה. בעוד שדבר זה לא קשה במיוחד עם מחשבים ורשתות מודרניות, לרוב המכשירים ברשתות שליטה ישנות (Legacy) יש מעבדים בעלי כוח נמוך בהרבה והם חווים הצטברויות סחיפה בשעון. לדוגמה, בסט המידע באחד הניסויים הקודמים, ה-RTU נע מהשעון של המחשב מנטר הרשת בקצב של 6 מילי-שניות בכל שעה.  
כדי לחקור ניסיונות זיוף לטכניקת החתימות הפיזיות, מאסטר העובד בפרוטוקול DNP3 הוגדר לשלוח פקודות פעולה כל שניה ואילו מכונת התוקף שינתה קוד פתוח של פרוטוקול DNP3 כך שישלח תגובות עם חתימות זמן המחושבות באמצעות הזמן הנוכחי של שעון המכונה. בניסיונות הזיוף של התוקף החלש יותר, ניכר הבדל ברור בין ההתפלגויות, בגלל השעון של המכשיר שסוטה במהירות מהשעון של מכשיר המאסטר. ניסיונות הזיוף של התוקף החזק יותר דומים למקור אך עדיין בעלות הבדל נראה לעין, הנגרם ככל הנראה מכך שהמחשב החזק של התוקף מעניק חותמות זמן מהר יותר מהמכשיר המקורי. למרות ששתי הטכניקות מציגות עמידות בפני שיטות תקיפת זיוף נאיביות, מחברי המאמר מודים כי עדיין אפשרי שהתוקף יעצב את זמני התגובה שלו בצורה מושכלת יותר כך שיתאימו לחתימת האמת של המכשיר או יממש שיטה שתשמור על סינכרוני שעון טובים יותר. יחד עם זאת, דבר זה ידרוש מהתוקף להיות בעל יכולות, ניסיון וידע *משמעותיים מאוד* כדי לתקוף בהצלחה.

**מגבלות**

בעוד ששתי השיטות המוצעות פועלות היטב תחת תנאים מסוימים, עדיין קיימות כמה מגבלות. שיטת ה-CLRT דורשת מפרוטוקול ה-SCADA להשתמש בהודעות 'Read' ו-'Response'. כל פרוטוקולי ה-SCADA הפופולריים ביותר מממשים הודעות אלו. בנוסף, פרוטוקול ה-SCADA חייב לשבת על מימוש TCP שמשתמש לפחות בכמות מינימלית של 'Quick ACKs' (שליחת ACK באופן מידי על חבילה במקום לעכב אותה בקפיצות). לדוגמה, מערכות לינוקס מודרניות משתמשות ב-Quick ACKs כדי להאיץ את שלב ה-Slow start של TCP בתחילת חיבור ולאחר שידורים חוזרים, אבל כל הספקים שהופיעו בתצפית בתחנת הכוח השתמשו ב-Quick ACKs עבור כל אחת מהחבילות, ככל הנראה כדי להפחית עיכובים. לכן, כמות ה-Quick ACKs שיהיו בשימוש כל מכשיר תקבע כמה מהר ניתן יהיה לייצר חתימה.  
שיטת החתימות הפיזיות דורשת רזולוציית תזמון גבוהה ולכן עליה לעשות שימוש בפרוטוקולים הכוללים חתימות זמן בתגובות עצמן. לא כל פרוטוקול SCADA תומך בפונקציונליות שכזו אבל אלו שנעשה בהם שימוש בסביבות בהן מרכיב הזמן קריטי, כמו רשת החשמל, משתמשים בחתימות זמן כאלו.  
דרישה לחתימות זמן בתעבורת רשת היא מגבלה במובן שכך יהיה לתוקף קל יותר לייצר ולזייף חתימת מכשיר אבל היא יכולה גם להיות לחוזקה הגנתית. אם תעבורת הרשת מוצפנת, על התוקף לפנות לגישת הקופסה הלבנה כדי ליצור חתימות כלשהן, דבר שאינו טריוויאלי והופך אף לקשה יותר ככל שהמכשירים ברשת הופכים למורכבים יותר.

**עבודה קשורה**

גישות לעבודה עם עקבות מכשירים בדרך כלל מסווגות לגישות אקטיביות וגישות פאסיביות,   
אחד הכלים הישנים והמוכרים ביותר לעבודה עם עקבות מכשירים ברשת הוא NMAP אשר משתמש בגישות פעילות לאיסוף מידע על מכשירים על ידי שליחת סדרת של בקשות ייחודיות. הכלי קובע את סוג מערכת ההפעלה וגרסת השרת שרצים על המכונה, בהתבסס על תגובות המכשיר. אך בעוד שכלי זה בהחלט שימושי עבור תוקפים ומגנים כאחד ברשת סטנדרטית, שימוש בכלי זה ברשת של מערכות שליטה תעשייתיות מוגבל, בעיקר מכיוון שגישות פעילות אינן רצויות בהן (מסיבות של כח חישובי נמוך וקצוב בד"כ).  
מהצד השני, קיים מגוון רחב של טכניקות המספקות עקבות לסוגי מכשירים ולמכשירים פרטניים, בצורה פאסיבית. דוגמה אחת היא כלי קוד פתוח בשם P0f אשר בוחן שדות ב- TCP & HTTP Headers על מנת לקבוע מידע על המכשיר, כגון סוג מערכת ההפעלה וגרסת הדפדפן אותו הוא מריץ.  
הניסיון הראשון ליצירת שיטות עבודה רשמיות עם פרוטוקולי רשת, פורסם בשנת 2006. המחברים השתמשו באוטומט סופי מורחב פרמטרים (PEFSMs) על מנת למדל את התנהגות המימושים השונים לפרוטוקול.  
לזיהוי גרסאות תוכנה בהחלט קיים שימוש, אך לזיהוי מכשירים פרטניים ברשת, בהתבסס על החומרה שלהם, קיים שימוש נרחב אף יותר.  
ע"י שימוש בשילוב הגישות (פאסיבית ואקטיבית), הצליח קוהנו א-האל לייצר בשנת 2005 את העבודה הראשונה שחקרה חתימות זמן של פרוטוקול ה-TCP וכך לזהות מכשיר פרטני ע"פ סטיות השעון שלו.

מחקר נוסף בגישות פאסיביות לעבודה עם עקבות התמקד בהיבטים שונים של תזמון תעבורת הרשת על מנת לייצר חתימות למכשירים וסוגי מכשירים.  
בשנת 2010, חוקרים הצליחו לנתח בצורה פאסיבית מידע שעבר דרך נקודות גישה אלחוטיות (access points) וזיהו בהצלחה כל אחת מנקודות הגישה. שנה לאחר מכן התפרסם מחקר נוסף שתיאר שיטה לייצור עקבות מכשיר בהתבסס על מודלי תזמון של מימוש פרוטוקולי שכבת האפליקציה במכשיר.  
מאמר שלישי שהשתמש בתצפיות של תזמוני תעבורת הרשת פורסם בשנת 2014 והשתמש בהתפלגות Inter-arrival times (IAT) על מנת לזהות מכשירים וסוגי מכשירים.

למרות ששלושת המאמרים הנ"ל נקטו בגישות שונות לשימוש בניתוח תזמוני תעבורת הרשת, שלושתן בלתי אפשריות למימוש ברשת של מערכת שליטה תעשייתית.  
גישת ניתוח אדוות המידע נבדקה רק על נקודות גישה אלחוטיות תחת עומסים כבדים, תרחיש שאינו מתקיים במערכות שליטה תעשייתיות, היכן שהתקשורת לרוב חוטית מסיבות של אמינות וקצבי העברת מידע.  
הגישה השנייה, מתבוננת על התנהגות שכבת האפליקציה ודורשת בסיס נתונים גדול שיכיל את כל התרחישים (sessions) האפשריים.  
הגישה השלישית והאחרונה שהוצעה, עשתה שימוש ב-IAT ודורשת מספר גדול (לפחות 2500) של דגימות מידע לקבוצת האימון על מנת להשיג תוצאות מדויקות. עבור חלק ניכר מהמכשירים בהם עושים שימוש במערכות שליטה תעשייתיות, פעולות מסוימות דורשות פרק זמן ארוך ואף זמן השהייה בין הפעלה להפעלה, מה שיגרום לאיסוף המידע להפוך למטלה קשה ובלתי סבירה לביצוע.  
גישה נוספת שפותחה משתמשת במדידות זמן של חיבורי USB על מנת לייצר חתימות עבור המכשיר המארח, אך גם זו אינה מתאימה לצרכינו, שכן מרבית המכשירים בסביבת מערכת השליטה התעשייתית כלל אינם תומכים בחיבורי USB ומכיוון שהרצון הוא לייצר חתימות בצורה פאסיבית עבור כל המכשירים ברשת במפגש אחד ולא לנדוד בין מכשירים שונים בעולם ולסווג את כולם.  
עוד גישה ייחודית, רלוונטית למאמר זה, התמקדה בשכבה הפיזית של מהלך התקשורת בין מכשירים בשכבות הנמוכות יותר. באמצעות מדידות שונות הצליחו חוקרים לזהות מכשירים ע"פ גלי הרדיו שרכיבי ה-WiFi שלהם פלטו. זו הייתה העבודה הראשונה בנושא שעשתה שימוש בשכבה הפיזית ובכל זאת בלתי אפשרית למימוש בסיבה המדוברת, בה רכיבי WiFi נמצאים בשימוש לעיתים רחוקות בלבד.

שתי הגישות המוצגות במאמר זה, מתגברות על הקשיים שהעבודות הקודמות נתקלו בהם, ע"י מתן דיוק תוצאות גבוה יותר תוך שימוש בטכניקות המותאמות למערכות שליטה תעשייתיות.  
בעוד שהשיטה הראשונה מתמקדת בשיפור השימוש במדידות לזמני התעבורה ברשת, כך שהמדידות בהן ייעשה שימוש יהיו מותאמות למערכות שליטה תעשייתיות, השיטה השנייה מרחיבה את רעיון השימוש בשכבה הפיזית לייצור חתימות ולזיהוי מכשירים ברשת, בהתבסס על זמני הביצוע המדווחים לפעולות פיזיות. בנוסף, כל העבודות הקודמות בנושא עשו שימוש בגישת ה'קופסה השחורה' שדורשת גישה למכשיר יעד כדוגמא ע"מ לייצר בסיס נתונים מקדים. מאמר זה מציג גישה שמתגברת על המגבלות הללו ע"י גישת 'קופסה לבנה' שאינה דורשת גישה מקדימה למכשיר כלשהו.

אחד השימושים העיקריים לשתי הטכניקות המוצעות לייצור חתימות מכשירים, יכול להיות להעצים את יכולותיהם של פתרונות קיימים לזיהוי חדירות ברשת (IDS) אשר בתחומם בוצעה עבודה מקדימה נרחבת. הניסיון הראשון להתאמת מערכת IDS למערכת שליטה תעשייתית הוצע ע"י מעבדות המחקר הלאומיות של מדינת איידהו (ארה"ב) והתמקד בניטור זרמי תעבורה, חיפוש דפוסים חוזרים והבנת החבילות בשכבת הרשת בניסיון למציאת חדירות.  
כמה חוקרים ניגשו לבעיה זו בכך ששינו תוכנת IDS נפוצה ומפורסמת כך שתבצע זיהוי חדירות ספציפיות הניתנות לביצוע על גבי פרוטוקולים מסוימים של מערכות שליטה תעשייתיות. אחרים ניסו למדל את המצבים אליהם מערכת שכזו יכולה להיכנס ולזהות מתי פקודה עלולה להכניס את המערכת למצב קריטי. הפתרונות האלה בהחלט מסוגלים לגלות סוגים שונים של התקפות על מערכות שליטה תעשייתיות, אך נכשלות בזיהוי מתקפות מחוכמות יותר ובהן התקפות 'הזרקת מידע שקרי'. הגישות המוצעות במאמר זה מציגות שתי גישות חדישות שהן מספיק כלליות (גנריות), כך שקל יהיה ליישם אותן במרבית מערכות השליטה התעשייתיות, והן מאפשרות זיהוי מדויק של מידע שקרי והודעות שליטה.

**סיכום**

במאמר זה הוצגו שתי שיטות חדשות, פאסיביות לייצור חתימת מכשירים ברשתות מערכות שליטה תעשייתיות. לאחר הערכת השיטות תוך שימוש בתצפיות מידע מהעולם האמתי יחד עם ניסויי מעבדה מבוקרים, הושגה נכונות סיווג בגובה של 99% ו-92% עבור השיטה הראשונה והשיטה השנייה בהתאמה. שתי השיטות הראו עמידות למתקפות פשוטות ויכולת מימוש לצד פתרונות IDS מסורתיים יותר, על מנת לחזק את אבטחת רשתות מערכות השליטה התעשיות הקריטיות.  
כעבודה להמשך, מחברי המאמר מתכננים לשפר את גישת הקופסה הלבנה ולהרחיב את השיטות כך שיתאימו למכשירי IOT. כמו כן הם מתכננים לחקור את האפשרות של פיתוח טכניקות פעילות לעבודה עם חתימות, על מנת להגביר את נכונות הדיוק.