پژوهشکده علوم کامپیوتر

پروپوزال همکاری پژوهشی تک پروژه غیر مقیم

عنوان

تشخیص ترکیبی ناهنجاری‌ها و تخمین عیوب برای ایمن‌سازی سیستم‌های سایبری فیزیکی

Hybrid Anomaly Detection and Fault Estimation for Secure Cyber-Physical Systems

نام و نام خانوادگی متقاضی: دکتر فائزه فریور	نشانی: تهران، انتهای بزرگراه شهید ستاری شمال، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، ساختمان ابن سینا، دانشکده برق و کامپیوتر	نشانی پست الکترونیکی: farivar@ieee.org
شمارهٔ تلفن ثابت: ۸۸۸۸۷۶۸۲-۰۲۱	شمارهٔ تلفن همراه: ۰۹۱۲۱۴۶۳۸۴۵	همکاران طرح:

چکیده: سیستم‌های سایبری فیزیکی[1] بخش جدایی‌ناپذیر زیرساخت‌های حیاتی هستند، اما در برابر تهدیدات سایبری و عیوب فیزیکی که می‌توانند عملکرد آنها را مختل کنند، آسیب‌پذیر هستند. این تحقیق به معرفی یک رویکرد ترکیبی خواهد پرداخت که با ترکیب تکنیک‌های پیشرفته تشخیص ناهنجاری و تخمین عیوب، امنیت و تاب‌آوری سیستم‌های سایبری فیزیکی را افزایش می‌دهد. در این رویکرد، با بهره‌گیری از روشهای تشخیص نفوذ مبتنی بر یادگیری ماشین و بکارگیری رویتگرهای خطی و غیرخطی قابل تنظیم به شناسایی ناهنجاری‌ها، بازسازی حملات سایبری و تخمین خطاها به صورت بلادرنگ پرداخته می‌شود. هدف از این پژوهش، پوشش‌دهی شکاف بین تشخیص ناهنجاری مبتنی بر محور و تخمین خطا بر اساس نظریه کنترل می‌باشد که منجر به توسعه چارچوب‌های امنیتی مقاوم و انطباق‌پذیر برای سیستم‌های سایبری فیزیکی خواهد شد.

واژگان کلیدی: تشخیص ناهنجاری، سیستم سایبری فیزیکی، یادگیری ماشین، کنترل غیرخطی، هوش مصنوعی

توضیحات در خصوص دستاوردهای طرح پژوهشی سال ۱۴۰۳

این طرح پیشنهادی در راستای تکمیل و توسعه طرح پژوهشی سال قبل است که به طراحی کنترل کننده‌های مبتنی بر هوش مصنوعی برای تشخیص و جبرانسازی حملات سایبری در سامانه‌های سایبری فیزیکی پرداخته شد و دستاوردهای آن به شرح زیر است:

1. S. Barchinezhad, MS. Haghighi, **F. Farivar**, A. Khonsari, “Enhancing the Resilience of Cyber-Physical Syster to DoS Attacks by Adopting Latency-Reducing TCP Variants”, **IEEE Transactions on Dependable and Secu Computing**, submitted, May 2024. **(IF=7.0)**

   ارسال شده در اردیبهشت ۱۴۰۳ و اکنون در مرحله ریوازیر دوم

2. Z. Li, Y. Ding, J. Li, **F. Farivar,** “Crowdsourcing Framework for Security Testing and Verification of Industr Cyber-Physical Systems”, **IEEE Network**, submitted, Sep. 2024. **(IF=6.8)**

   ارسال شده در شهریور۱۴۰۳ و اکنون در مرحله ریوازیر اول

[1] Cyber-Physical Systems (CPS)

Abstract:
Cyber-Physical Systems (CPS) are integral to critical infrastructures, yet they remain vulnerable to cyber threats and physical faults that can disrupt operations. This proposal explores a hybrid approach combining advanced anomaly detection and fault estimation techniques to enhance CPS security and resilience. Leveraging machine learning-based intrusion detection alongside an adjustable proportional-integral observer framework, this approach aims to detect anomalies, reconstruct cyber-attacks, and estimate faults in real time. This research seeks to bridge the gap between data-driven anomaly detection and control-theoretic fault estimation, contributing to the development of robust and adaptive CPS security frameworks.

Keywords:
Anomaly Detection, Cyber-Physical Systems, Machine Learning, Nonlinear Control, Artificial Intelligence

مقدمه

سیستم‌های سایبری فیزیکی به عنوان یکی از ارکان اساسی زیرساخت‌های حیاتی مدرن شناخته می‌شوند و در حوزه‌هایی مانند اتوماسیون صنعتی، شبکه‌های هوشمند، سامانه‌های حمل‌ونقل و مراقبت‌های بهداشتی کاربرد گسترده‌ای دارند. این سیستمها با ترکیب پردازش داده و کنترل فیزیکی، امکان تصمیم‌گیری خودکار و نظارت بلادرنگ را فراهم می‌کنند. با این حال، پیچیدگی و اتصال گسترده‌ی سیستم‌های سایبری فیزیکی باعث شده است که این سیستمها در معرض تهدیدات سایبری و عیوب فیزیکی قرار گیرند، که می‌تواند منجر به نقص عملکرد و حتی پیامدهای فاجعه‌بار شود. نمونه‌های زیادی از حملات سایبری دنیای واقعی گزارش شده است، مانند نفوذ آب Maroochy در مارس ۲۰۰۰ [1]، چندین خاموشی برق اخیر در برزیل [2]، کرم کامپیوتری StuxNet در ژوئن ۲۰۱۰ [3] و سایر حوادث امنیتی مختلف بر سیستم‌های صنعتی [4]. در مثال های مذکور، حملات به سیستم فیزیکی و یا شبکه‌های ارتباطی صورت می‌گیرد به عنوان مثال، کرم Stuxnet تهدیدهای جدی برای سیستم ایجاد کرد و سیستم کنترل صنعتی PLC را آلوده و مجدد برنامه‌ریزی کرد[5] و [6]. سیستم‌های سایبری فیزیکی مستعد آسیب پذیری‌های سایبری هستند که بر روی سیستم‌های کنترل اصلی آنها تأثیرگذار است. از این رو، تشخیص نفوذ برای شناسایی حملات رخ داده و تحلیل اثرات آن حمله مخرب و سپس جبران‌سازی اثر حمله و حفظ عملکرد درست و نرمال سیستم سایبری فیزیکی، امری ضروری و حیاتی است.

از منظر دیگر، می‌توان حملات سایبری را به انواع هدفمند، غیر هدفمند و حوادث شبکه دسته‌بندی نمود. حملات هدفمند به سیستم‌های سایبری فیزیکی، مهمترین حملات هستند زیرا استراتژی‌های مهاجم برای آسیب رساندن به اجزای سیستم سایبری فیزیکی می‌باشد. حملات غیرهدفمند، مشابه حوادث شبکه است که در نتیجه خرابی‌های ناخواسته شبکه ایجاد می‌شوند که دسته دیگر از حملات سایبری، حملات پنهان[1] هستند که زیرمجموعه حملات فریب و از نوع هدفمند هستند. در این حملات، مهاجم کنترل برخی از قسمت‌های سیستم سایبری فیزیکی به عنوان مثال سیگنال کنترل از طریق کانال شبکه را بدست می‌گیرد. هدف مهاجم این است که اثر حمله را به اندازه کافی کوچک نگه دارد تا توسط ناظران انسانی شناسایی نشود.

روش‌های سنتی امنیت سایبری مانند رمزنگاری و کنترل دسترسی برای مقابله با تهدیدات پیشرفته‌ای نظیر حملات تزریق داده‌ی جعلی، حملات عدم سرویس‌دهی (DoS) و حملات بازپخش کافی نیستند. علاوه بر این، خطاهای فیزیکی نظیر خرابی حسگرها یا نقص عملکرد عملگرها می‌توانند پایداری و عملکرد سیستم را تحت تأثیر قرار دهند. از این رو، رویکردی جامع که شامل شناسایی ناهنجاری و تخمین عیوب باشد، برای بهبود تاب‌آوری سیستم‌های سایبری فیزیکی ضروری است.

[1] Covert Attacks

در این پژوهش، یک چارچوب ترکیبی برای شناسایی ناهنجاری‌ها و تخمین عیوب پیشنهاد می‌شود که از ترکیب سامانه‌ی تشخیص نفوذ مبتنی بر یادگیری ماشین و رویتگرها و تخمینگرهای قابل تنظیم بهره می‌برد. این روش امکان شناسایی حملات سایبری، بازسازی تهدیدات و تخمین عیوب را به صورت بلادرنگ فراهم می‌کند. هدف اصلی این پژوهش، پوشش دهی شکاف میان روش‌های داده‌محور تشخیص ناهنجاری و روش‌های مبتنی بر کنترل برای تخمین خطا است. مهم‌ترین دستاوردهای این پژوهش عبارتند از: ۱) ادغام روش‌های یادگیری ماشین برای تشخیص ناهنجاری با روش‌های مبتنی بر کنترل برای تخمین خطا، ۲) توسعه‌ی یک چارچوب بلادرنگ برای شناسایی، موقعیت‌یابی و کاهش اثرات حملات سایبری و نقص‌های فیزیکی، ۳) ارزیابی روش پیشنهادی بر روی مدل‌های استاندارد سیستم‌های سایبری فیزیکی برای بررسی اثربخشی و پایداری آنها. لازم به ذکر است که شبیه‌سازی های اولیه انجام شده و نتایج نشان دهنده کارایی ایده‌های پیشنهادی می باشد.

پیشینه پژوهش

در سال‌های اخیر، پژوهش‌های متعددی از زوایای گوناگون، از جمله امنیت، آسیب‌پذیری، کنترل و کاربردهای مختلف، به بررسی سامانه‌های سایبر-فیزیکی پرداخته‌اند. با افزایش چشمگیر حملات سایبر-فیزیکی، نیاز به راهکارهای مقابله با این تهدیدات بیش از پیش احساس می‌شود و پس از حادثه‌یStuxnet، توجه ویژه‌ای به این موضوع معطوف شده است [7]. در این بخش، به مرور برخی از مطالعات مرتبط در این زمینه پرداخته می‌شود.

مرجع [8] به مطالعه حملات فریب و حمله انکار سرویس بر روی یک سیستم کنترل شده مبتنی بر شبکه پرداخته است. حملات فریب احتمال آسیب رساندن به یکپارچگی بسته های کنترل یا اندازه گیری است که منجر به تغییر در رفتار سنسورها و عملگرها می‌شود. آنها حملات یکپارچگی[1] در مرجع [9] نامیده شده‌اند. در مقابل، حملات انکار سرویس دسترسی به منابع را از طرق مختلف مانند مسدود کردن کانال ارتباطی به خطر می‌اندازد. در مرجع [10] حملات تزریق داده‌های نادرست در شبکه ارتباطی سیستم‌های سایبری فیزیکی معرفی شده‌اند که از نوع حملات فریب هستند. به طور مشابه، حملات فریب پنهان علیه سیستم کنترل و نظارت در مرجع [11] مورد مطالعه قرار گرفته است. از مطالعات انجام شده در زمینه طراحی حملات سایبری برای سیستم‌های سایبری فیزیکی می‌توان به مرجع [12] اشاره کرد. مرجع [12]، تهدیدهای سیستم‌های سایبری فیزیکی را طبقه بندی کرده و خانواده حملات پنهان را معرفی کرده است. در این مطالعه، یک حمله فریب پنهان برای ایجاد اختلال در عملکرد سیستم کنترل پیشنهاد شده است که بر اساس اطلاعات جمع آوری شده توسط یک حمله دیگر است. این دو حمله می‌توانند رفتار یک سیستم را تغییر دهند.

همچنین، مطالعات گسترده‌ای برای کاهش یا جبران اثر حملات سایبری انجام شده است. در مرجع [13]، یک روش تحلیلی برای امنیت شبکه‌های هوشمند مبتنی بر تجزیه و تحلیل توان بر اساس تحلیل خطا ارائه شده است. روش پیشنهادی مبتنی بر یک الگوریتم تخمین حالت است که مجموع مربعات خطا را به حداقل می‌رساند. در این مطالعه، حملات سایبری به عنوان اطلاعات بد[2] مدلسازی شده است. شناسایی و جبران حملات سایبری همزمان در سیستم‌های گرید[3] مورد بررسی قرار گرفته است.

تحقیق [14] برای شناسایی و کاهش اثر سایبری مبتنی بر مدل برای سیستم‌های کنترل خودکار است. این مقاله ابتدا تأثیر حملات فریب را بر کنترل خودکار در سیستم الکتریکی مورد مطالعه قرار داده و سپس یک چارچوب کلی برای استفاده در سیستم‌های کنترلی که تحت حملات هوشمند کار می‌کنند، پیشنهاد شده است.

[1] Integrity attacks
[2] Bad information
[3] Grid systems

مطالعه [15] در زمینه تشخیص نفوذ و جبران حملات انکار سرویس و همچنین جبران تلفات بسته برای سیستم‌های کنترل‌شده با شبکه‌های حسگر بی‌سیم صنعتی است. در این مقاله، یک کنترل‌کننده پیش‌بین[1]، یک کنترل‌کننده تنظیم درجه دوم خطی[2] و ترکیبی از هر دو جبران‌کننده برای کاهش اثرات حمله از دست دادن بسته‌ها بر روی شبکه حسگر بی‌سیم صنعتی پیشنهاد شده است. سیستم کنترل در شرایط مختلف تلفات بسته مورد استفاده قرار می‌گیرد. شبیه‌سازی‌ها بر روی یک سیستم آونگ انجام شده است.

در مرجع [16]، حملات انکار سرویس در سیستم‌های کنترل مبتنی بر شبکه بررسی شده است. یک مدل صف برای شبیه‌سازی فرآیند تاخیر تصادفی برای ورود بسته‌ها و آسیب‌های تحت حملات انکار سرویس پیشنهاد شده است. تمرکز بر این است که چگونه این حملات عملکرد سیستم‌های کنترل شبکه را مختل می‌کنند. سیستم کنترل، یک کنترل‌کننده انتگرال-تناسبی است که برای یک سیستم مرتبه دوم اعمال شده است و توسط مهاجمان با تاخیر تزریق می‌شود.

در مرجع [17]، تشخیص و شناسایی خطا در سیستم‌های کنترل فرآیند با استفاده از روش‌های طبقه‌بندی داده محور مورد مطالعه قرار گرفته است. در رویکرد اول ابتدا تشخیص خطا اجرا می شود و در صورت مثبت بودن نتیجه طبقه‌بندی فعال می‌شود. در روش دوم هر دو مورد تشخیص و تعیین خروجی طبقه‌بند به طور همزمان انجام می‌شود.

در مقاله [18]، تزریق داده‌ها و تشخیص نفوذ در سیستم کنترل اسکادا[3] مورد بررسی قرار گرفته است. این مقاله مجموعه‌ای از تزریق داده و حملات انکار سرویس را توسعه داده است که از عدم احراز هویت در بسیاری از پروتکل‌های ارتباطی مانند MODBUS، DNP۳ و EtherNet/IP استفاده می‌کنند. این مطالعه از ویژگی‌های این پروتکل‌ها برای توسعه یک سیستم تشخیص نفوذ مبتنی بر شبکه عصبی و شناسایی حملات تزریق داده استفاده می‌کند.

■ من (مجری طرح) به همراه همکارنم در برخی پژوهشهای پیشین به امنیت سامانه‌های صنعتی از نوع سایبری فیزیکی به شرح زیر پرداخته‌ایم:

در مرجع [19]، یک چارچوب تشخیص نفوذ و جبران برای تسلط بر حملات مخفی توسعه داده شده است. یک شبکه عصبی مصنوعی به عنوان یک سیستم تشخیص نفوذ استفاده می‌شود. علاوه بر این، یک کنترل‌کننده مقاوم تطبیقی مبتنی بر روش کنترل مدلغزشی طراحی شده است تا پس از شناسایی حمله، جایگزین کنترلر کلاسیک تناسبی- انتگرالی- مشتقی شود.

در مطالعه [20]، یک رویکرد کنترل ترکیبی برای تحمل و جبران حملات سایبری رخ داده در ورودی‌ها و خروجی‌های یک سیستم سایبری فیزیکی از نوع جرثقیل دورانی پیشنهاد شده است. فرض بر این است که حملات سایبری از نوع انکار سرویس هستند و باعث از دست رفتن بسته‌ها با احتمال زیاد در کانال‌های ارتباطی مستقیم و بازخورد می‌شوند. در این مقاله، سیستم‌های کنترل مقاوم و هوشمند طراحی شده است.

در پژوهش[21]، یک رویکرد کنترلی ترکیبی هوشمند-کلاسیک برای بازسازی و جبران حملات سایبری رخ داده در کانال‌های ارتباطی مستقیم سیستم‌های فیزیکی-سایبری غیرخطی پیشنهاد شده است. یک شبکه عصبی تابع پایه‌ای شعاعی[4] به عنوان یک تخمین‌گر هوشمند برای حملات سایبری استفاده می‌شود و از روش کنترل ساختار متغیر برای جبران اثرات حملات سایبری و کنترل عملکرد سیستم سایبری فیزیکی در اهداف ردیابی استفاده می‌شود. پایداری سیستم با استفاده از تئوری پایداری سیستم‌های غیرخطی اثبات شده است. استراتژی پیشنهادی در یک سیستم کروز کنترل خودرو به عنوان بستر آزمایش مورد مطالعه قرار گرفته است.

[1] Model predictive control
[2] Linear Quadratic Regulation
[3] SCADA
[4] Radial basis function

در پژوهش [22]، حملات پنهان سایبری به وسایل نقلیه خودمختار معرفی شده است که می تواند ایمنی وسیله نقلیه و مسافران را به خطر بیندازد. یک حمله پنهانی به سیستم رهیگیر مسیر[1] در خودرو خودران طراحی و اجرا شد. در سناریو مورد مطالعه، مهاجم هوشمند، داده‌های اندازه‌گیری توسط سنسورها مانند موقعیت خط ، انحنا و غیره را دستکاری می‌کند تا کنترل‌کننده اصلی خودرو خودران را فریب دهد و وسیله نقلیه را به مرزهای جاده نزدیک کند. مهاجم تعاملی رفتارهای وسیله نقلیه میزبان از نظر انحراف جانبی و قدرت مانور را فرا می‌گیرد و سعی می‌کند خطاها را تا حدی افزایش دهد که برای راننده قابل توجه نباشد. این فرایند با استفاده از یادگیری بازیگر-منتقد انجام خواهد شد. علاوه بر این، نشان داده می‌شود که چگونه می‌توان یک سیستم تشخیص نفوذ برای چنین حملات پنهانی برای هشدار دادن به راننده طراحی کرد. از داده‌های GPS و همچنین از نقشه‌های آفلاین برای بازسازی منحنی‌های جاده و مطابقت آن با خوانش‌های سنسورها استفاده شده است.

مطالعه [23] که یکی از تلاشهای تیم مشترکی از مجری و همکاران داخلی و طرف خارجی این پروپوزال می باشد و همچنان در حال توسعه است، به مساله امنیت سیستم کنترل کروز تطبیقی خودروهای هوشمند می‌پردازد. این سامانه ها با کمک سنسورهای سرعت و رادار، سرعت خودرو را کنترل کرده و در عین حال فاصله آن را تا خودروی جلویی در حد ایمن نگاه میدارند. دو حمله مخفی جدید در این تحقیق کشف و گزارش گردید که برای پیشگیری و مقابله با آنها نیز یک روش پیشنهادی ارایه شد.

درمطالعه [24]، یک حمله مخفی جدید معرفی می شود که می‌تواند توسط بدافزارهایی طریق سیستم‌های حمل و نقل هوشمند منتشر شود. به عنوان نمونه، این بدافزار یک ماژول وسیله نقلیه مانند کنترلر کروز تطبیقی[2] را آلوده می‌کند و تنظیمات آن را به گونه ای دستکاری می‌کند که برای ناظران انسانی قابل توجه نباشد، اما بطور آماری میزان تصادفات را افزایش می دهد. در این مطالعه نشان داده شده است که چگونه این کار انجام می شود و اثر آن به صورت ریاضی تجزیه و تحلیل شده است. همچنین، یک سیستم تشخیص نفوذ جدید پیشنهاد می‌شود که می‌تواند به آسانی توسط تولیدکنندگان خودرو مورد استفاده قرار گیرد. سیستم پیشنهادی با مجموعه داده های دنیای واقعی ارزیابی شده و نشان داده شده است که چگونه یک بدافزار/مهاجم می‌تواند آمار تصادف را با دستکاری مقدار فاصله ایمن در سناریوهای کروز کنترل مهندسی کند. سپس، یک سیستم تشخیص نفوذ مبتنی بر ناهنجاری برای ماژول‌های کروز کنترل مورد آزمایش قرار گرفته و نشان داده شده است که چگونه می‌تواند به طور موثر چنین حملات مخفی را شناسایی کند.

در مطالعه [25]، معرفی و ایجاد شبکه های عصبی مقاوم تحت حملات یادگیری ماشینی متخاصم با استفاده از نورون های الهام گرفته شده از بیولوژیک انجام شده است. در هسته هر شبکه عصبی مصنوعی، نورون هایی قرار دارند که مدل های تقریبی از همتایان بیولوژیکی خود هستند. با این حال، شبکه‌هایی که از نورون‌های مصنوعی ساخته شده‌اند می‌توانند در مقابل حملات یادگیری ماشینی متخاصم غیرقابل اعتماد شوند. در این مطالعه مشکل ریشه‌یابی شده و یک نورون مصنوعی قابل اعتماد ایجاد شده که می‌تواند در برابر حملات متخاصم بهتر مقاومت کند. در مدل نورون پیشنهادی، مقدار ورودی یک نورون مصنوعی با امیدریاضی یک متغیر تصادفی برنولی جایگزین می‌شود. آزمایش‌های گسترده‌ای برای ارزیابی قابلیت اطمینان این نورون جدید انجام شده و نتایج، افزایش استحکام و دقت تحت حملات خصمانه یا به‌طور معادل آن، افزایش قابل‌توجهی در میزان اعوجاجی که مهاجمان برای ایجاد نمونه‌های متخاصم موفق باید به تصاویر اضافه کنند، را تأیید می‌کند.

[1] Lane Keeping System
[2] Adaptive Cruise Control

دلایل انجام طرح

با پیشرفت فناوری و گسترش سیستم‌های سایبری فیزیکی در زیرساخت‌های حیاتی مانند نیروگاه‌ها، سیستم‌های حمل‌ونقل هوشمند و شبکه‌های صنعتی، نیاز به روش‌های کارآمد برای تأمین امنیت و اطمینان‌پذیری این سامانه‌ها بیش از پیش احساس می‌شود. تهدیدات سایبری و عیوب فیزیکی می‌توانند عملکرد سیستم‌های سایبری فیزیکی را مختل کرده و پیامدهای جبران‌ناپذیری را به دنبال داشته باشند. در این راستا، توسعه‌ی روش‌های هوشمند برای شناسایی و کاهش اثرات حملات سایبری و خطاهای فیزیکی، امری ضروری است. یکی از چالش‌های اساسی در حوزه‌ی امنیت سیستم‌های سایبری فیزیکی، عدم کفایت روش‌های سنتی امنیتی است. روش‌های رمزنگاری و کنترل دسترسی تنها در برابر حملات مبتنی بر نفوذ کارایی دارند، اما در مقابل حملاتی که در داده‌های سیستم دستکاری ایجاد می‌کنند، مانند حملات تزریق داده‌ی جعلی و بازپخش، آسیب‌پذیر هستند. از سوی دیگر، روش‌های کنترلی کلاسیک که بر اساس مدل‌سازی ریاضی سیستمها کار می‌کنند، برای تشخیص و تخمین عیوب فیزیکی مناسب هستند اما در برابر تهدیدات سایبری که منجر به تغییر رفتار سیستم می‌شوند، کارایی محدودی دارند. ضرورت این پژوهش از چند جنبه قابل بررسی است:

- افزایش پیچیدگی حملات سایبری و تأثیرات گسترده‌ی آن‌ها بر سیستم‌های سایبری فیزیکی: در سال‌های اخیر، حملات سایبری به سیستم‌های سایبری فیزیکی از نظر پیچیدگی و قابلیت پنهان‌کاری پیشرفت چشمگیری داشته‌اند. بسیاری از حملات قادرند به‌طور هم‌زمان چندین زیرسیستم را هدف قرار دهند و تشخیص آن‌ها با روش‌های معمولی دشوار است. ترکیب یادگیری ماشین و تخمین‌گرهای کنترلی می‌تواند به بهبود دقت تشخیص این حملات کمک کند.
- محدودیت روش‌های مبتنی بر یادگیری ماشین در سیستم‌های سایبری فیزیکی: روش‌های داده‌محور مانند شبکه‌های عصبی و الگوریتم‌های یادگیری عمیق، به دلیل نیاز به حجم بالای داده‌های آموزشی و وجود هشدارهای نادرست بالا، به‌تنهایی برای سیستم‌های سایبری فیزیکی قابل اعتماد نیستند. بنابراین، ترکیب آن‌ها با روش‌های کنترلی که از مدل‌های دینامیکی سیستم بهره می‌برند، می‌تواند کارایی سیستم‌های تشخیص را بهبود بخشد.
- نیاز به رویکردهای بلادرنگ برای شناسایی و کاهش اثرات حملات و خطاها: بسیاری از روش‌های موجود، به‌ویژه آن‌هایی که مبتنی بر یادگیری ماشین هستند، دارای تأخیر محاسباتی بوده و برای کاربردهای بلادرنگ مناسب نیستند. استفاده از تخمین‌گرهای قابل تنظیم همراه با الگوریتم‌های تشخیص ناهنجاری، امکان واکنش سریع‌تر و دقیق‌تر به تهدیدات را فراهم می‌آورد.
- قابلیت تلفیق در سامانه‌های واقعی: یکی از چالش‌های مهم در روش‌های تشخیص ناهنجاری و تخمین خطا، امکان پیاده‌سازی آن‌ها در محیط‌های واقعی است. چارچوب پیشنهادی با استفاده از مدل‌های کنترلی شناخته‌شده و الگوریتم‌های یادگیری ماشین، قابلیت انطباق با طیف وسیعی از سیستم‌های سایبری فیزیکی را دارد و می‌تواند به‌عنوان یک ابزار کارآمد در صنایع مختلف مورد استفاده قرار گیرد.

با توجه به این چالش‌ها، این پژوهش تلاش می‌کند تا یک روش ترکیبی برای شناسایی ناهنجاری‌ها و تخمین خطاهای سایبری-فیزیکی ارائه دهد که ضمن حفظ دقت و قابلیت اطمینان، بتواند در شرایط بلادرنگ نیز عملکرد مناسبی داشته باشد. نتایج این پژوهش می‌تواند به توسعه‌ی سامانه‌های مقاوم‌تر و تطبیق‌پذیرتر در برابر تهدیدات کمک کرده و مسیر جدیدی برای بهبود امنیت و پایداری سیستم‌های سایبری فیزیکی فراهم آورد.

اهداف و محصولات طرح

طرح از نوع پژوهشی بوده و خروجی آن منجر به انتشار مقالات علمی معتبر و یا ثبت اختراع خواهد شد. این پروژه بخشی از یک خط پژوهشی است که پژوهشگر در زمینه امنیت سامانه‌های صنعتی دنبال می‌کند. فناوری‌های تشخیص و جبران‌سازی نفوذ می‌توانند در سامانه‌های زیرساخت حیاتی کشور نظیر سیستم‌های تصفیه و توزیع آب، سیستم‌های حمل و نقل، سیستم های توزیع برق و غیره مورد استفاده قرار گیرند. هدف بعدی پس از اتمام موفقیت آمیز این پروژه، پیاده سازی فناوری بدست آمده در صنعت است. امکان‌سنجی اولیه (Feasibility Study) برای این تحقیق انجام شده است.

برنامه و نحوه اجرای طرح

این پژوهش در سه مرحله‌ی اصلی طراحی و اجرا خواهد شد که به شرح زیر است:

مرحله اول: در این مرحله، چارچوب ترکیبی برای شناسایی ناهنجاری‌ها و تخمین عیوب در پایداری سیستم‌های سایبری فیزیکی توسعه داده خواهد شد. این مرحله شامل فعالیت‌های زیر است:

- تحلیل و مدل‌سازی پایداری سیستم‌های سایبری فیزیکی: بررسی ساختار و مدل دینامیکی سیستم‌های هدف، شامل فرمول‌بندی معادلات حالت و شناسایی متغیرهای کلیدی برای شناسایی ناهنجاری‌ها و خطاها.
- طراحی سامانه‌ی تشخیص ناهنجاری مبتنی بر یادگیری ماشین: توسعه‌ی یک مدل یادگیری ماشین برای تحلیل داده‌های پایداری سیستم‌های سایبری فیزیکی و شناسایی رفتارهای غیرطبیعی. روش‌های نظارت‌شده و غیرنظارت‌شده برای تشخیص حملات سایبری ارزیابی خواهند شد.
- توسعه‌ی تخمین‌گرهای قابل تنظیم: طراحی یک مشاهده‌گر کنترلی برای تخمین خطاهای سیستم و بازسازی اثر حملات سایبری. این مشاهده‌گر باید قابلیت تنظیم تطبیقی برای سازگاری با تغییرات سیستم را داشته باشد.
- ادغام روشهای پیشنهادی و ایجاد یک چارچوب ترکیبی: توسعه‌ی یک مکانیسم ترکیبی که از خروجی‌های یادگیری ماشین و مشاهده‌گر کنترلی برای تشخیص دقیق‌تر و تخمین اثرات حملات استفاده کند.

مرحله دوم: در این مرحله، روش پیشنهادی در بستر شبیه سازی ایجاد شده و عملکرد آن ارزیابی خواهد شد. این مرحله شامل گام‌های زیر است: انتخاب یک بستر شبیه‌سازی مناسب، تهیه‌ی مجموعه داده‌های واقعی یا شبیه‌سازی‌شده، اجرای سناریوهای مختلف حملات سایبری، تحلیل و مقایسه‌ی عملکردی روش‌های ترکیبی با روش‌های موجود.

مرحله سوم: در این مرحله، بر اساس نتایج بدست‌آمده از مرحله‌ی قبل روش پیشنهادی بهینه‌سازی شده و یافته‌ها مستند خواهند شد.

برنامه زمانی

مدت انجام این طرح، حدود ۱ سال پیش بینی شده است.

شماره مرحله / فعالیت	نام مرحله / فعالیت		درصد وزنی به کل پروژه	مدت زمان اجرا (ماه)	نمودار زمان‌بندی(ماه)											
			٪۱۰۰	۱۲	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
۱	ارائه یک چارچوب ترکیبی برای شناسایی ناهنجاری‌ها و تخمین عیوب در سیستمهای سایبری فیزیکی	تحلیل و مدل‌سازی پایداری سیستم‌های سایبری فیزیکی	٪۶۰	۸												
		طراحی سامانه‌ی تشخیص ناهنجاری مبتنی بر یادگیری ماشین														
		توسعه‌ی تخمین‌گرهای قابل تنظیم														
		ادغام روشهای پیشنهادی و ایجاد یک چارچوب ترکیبی														
۲	ایجاد بستر شبیه‌سازی و ارزیابی روش پیشنهادی	انتخاب بستر شبیه‌سازی مناسب	٪۲۰	۲												
		تهیه‌ی مجموعه داده‌های واقعی یا شبیه‌سازی‌شده														
		اجرای آزمایش‌های سناریو محور														
		تحلیل و مقایسه‌ی عملکردی روش ترکیبی														
۳	ارزیابی، بهینه‌سازی و ارائه‌ی نتایج	ارزیابی و بهینه‌سازی مدل تشخیص ناهنجاری و مشاهده‌گر کنترلی	٪۲۰	۲												
		ارزیابی تطبیق‌پذیری و مقیاس‌پذیری روش														
		مستندسازی و انتشار یافته‌های علمی														

همکاران طرح

این طرح همکار هیات علمی دیگری ندارد و توسط همکاران پژوهشی و مجری اجرا خواهد شد.

ابزار مورد نیاز

کیفیت انجام پروژه و میزان واقعی بودن بستر تست و ارزیابی روشهای توسعه داده شده، به مقدار بودجه تصویب شده برای پروژه بستگی دارد. بطور ایده‌آل این پروژه نیاز به بستر لازم برای پیاده‌سازی عملی دارد. اما برای شبیه‌سازی و در مقیاس کوچک، تجهیزات آزمایشگاهی مورد نیاز به غیر از موارد مرسوم رایانه ای، یک پلتفرم و تعدادی شبیه ساز تخصصی برای ارزیابی نتایج پروژه است. برخی از نرم افزارها و یا پایگاه داده ها در صورت نیاز باید بصورت لایسنس خریداری گردند.

بودجه تقریبی

میزان کار لازم برای این طرح در بخش پژوهشی حدود ۱۲ نفر ماه است که هزینه ای در حدود ۱/۰۰۰/۰۰۰/۰۰۰ ریال برای آن پیش‌بینی شده است. اما هر گونه حمایت پژوهشگاه دانش‌های بنیادی کمک شایانی خواهد بود. هزینه ها به نسبت درصد وزنی در مراحل انجام تحقیق توزیع می‌شوند.

مراجع

[1] J. Slay and M. Miller, “Lessons learned from the maroochy water breach,” in *IFIP International Federation for Information Processing*, 2007, vol. 253, pp. 73–82.

[2] J. P. Conti, “The day the samba stopped [power blackouts],” *Eng. Technol.*, vol. 5, no. 4, pp. 46–47, 2010.

[3] J. P. Farwell and R. Rohozinski, “Stuxnet and the future of cyber war,” *Survival (Lond).*, vol. 53, no. 1, pp. 23–40, 2011.

[4] R. Guy, “Hackers vs slackers [control security],” *Eng. Technol.*, vol. 3, no. 19, pp. 40–43, 2008.

[5] B. Bencsáth, G. Pék, L. Buttyán, and M. Félegyházi, “The cousins of Stuxnet: Duqu, Flame, and Gauss,” *Futur. Internet*, vol. 4, no. 4, pp. 971–1003, 2012.

[6] P. K. Kerr, J. Rollins, and C. A. Theohary, “The stuxnet computer worm: Harbinger of an emerging warfare capability,” in *The Stuxnet Computer Worm and Industrial Control System Security*, 2011, pp. 1–10.

[7] N. Falliere, L. O. Murchu, and E. Chien, “W32.Stuxnet Dossier,” *Symantec-Security Response*, vol. Version 1., no. February 2011, pp. 1–69, 2011.

[8] A. Teixeira, H. Sandberg, and K. H. Johansson, “Networked control systems under cyber attacks with applications to power networks,” in *Proceedings of the 2010 American Control Conference, ACC 2010*, 2010, pp. 3690–3696.

[9] Y. Mo and B. Sinopoli, “Integrity attacks on cyber-physical systems,” in *HiCoNS’12 - Proceedings of the 1st ACM International Conference on High Confidence Networked Systems*, 2012, pp. 47–54.

[10] A. Chattopadhyay and U. Mitra, “Security Against False Data-Injection Attack in Cyber-Physical Systems,” *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 2, pp. 1015–1027, 2020.

[11] C. Kwon, W. Liu, and I. Hwang, “Security analysis for Cyber-Physical Systems against stealthy deception attacks,” in *Proceedings of the American Control Conference*, 2013, pp. 3344–3349.

[12] A. O. De Sá, L. F. R. D. C. Carmo, and R. C. S. Machado, “Covert Attacks in Cyber-Physical Control Systems,” *IEEE Trans. Ind. Informatics*, vol. 13, no. 4, pp. 1641–1651, 2017.

[13] A. S. Bretas, N. G. Bretas, B. Carvalho, E. Baeyens, and P. P. Khargonekar, “Smart grids cyber-physical security as a malicious data attack: An innovation approach,” *Electr. Power Syst. Res.*, vol. 149, pp. 210–219, 2017.

[14] S. Sridhar and M. Govindarasu, "Model-based attack detection and mitigation for automatic generation control," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 580–591, 2014.

[15] V. A. Q. Nguyen and M. Yoo, "Packet Loss Compensation for Control Systems over Industrial Wireless Sensor Networks," *Int. J. Distrib. Sens. Networks*, vol. 2015, 2015.

[16] M. Long, C. H. Wu, and J. Y. Hung, "Denial of service attacks on network-based control systems: Impact and mitigation," *IEEE Trans. Ind. Informatics*, 2005.

[17] H. A. Nozari, S. Nazeri, H. D. Banadaki, and P. Castaldi, "Model-free fault detection and isolation of a benchmark process control system based on multiple classifiers techniques—A comparative study," *Control Eng. Pract.*, vol. 73, pp. 134–148, 2018.

[18] W. Gao, T. Morris, B. Reaves, and D. Richey, "On SCADA control system command and response injection and intrusion detection," in *General Members Meeting and eCrime Researchers Summit, eCrime 2010*, 2010.

[19] F. Farivar, M. S. Haghighi, S. Barchinezhad, and A. Jolfaei, "Detection and compensation of covert service-degrading intrusions in cyber physical systems through intelligent adaptive control," in *Proceedings of the IEEE International Conference on Industrial Technology*, 2019.

[20] M. Sayad Haghighi, F. Farivar, A. Jolfaei, and M. H. Tadayon, "Intelligent robust control for cyber-physical systems of rotary gantry type under denial of service attack," *J. Supercomput.*, 2019.

[21] F. Farivar, M. S. Haghighi, A. Jolfaei, and M. Alazab, "Artificial Intelligence for Detection, Estimation, and Compensation of Malicious Attacks in Nonlinear Cyber-Physical Systems and Industrial IoT," *IEEE Trans. Ind. Informatics*, 2020.

[22] F. Farivar, M. S. Haghighi, A. Jolfaei, and S. Wen, "Covert Attacks through Adversarial Learning: Study of Lane Keeping Attacks on the Safety of Autonomous Vehicles," *IEEE/ASME Trans. Mechatronics*, 2021.

[23] F. Farivar, M. S. Haghighi, A. Jolfaei, and S. Wen, "On the Security of Networked Control Systems in Smart Vehicle and Its Adaptive Cruise Control," *IEEE Trans. Intell. Transp. Syst.*, 2021.

[24] M. S. Haghighi, F. Farivar, A. Jolfaei, A. Bayrami Asl and W. Zhou, "Cyber attacks via consumer electronics: Studying the threat of covert malware in smart and autonomous vehicles," *IEEE Trans. on Consumer Electronics*, 2023. DOI: 10.1109/TCE.2023.3297965.

[25] H. Ghiassirad, .F. Farivar, M. Aliyari, and M. S. Haghighi, ""Building Robust Neural Networks under Adversarial Machine Learning Attacks by using Biologically-inspired Neurons," *Information Sciences*, Vol.645, pp.119190, 2023.

نام و نام خانوادگی متقاضی: **دکتر فائزه فریور– دانشیار دانشگاه آزاد واحد علوم و تحقیقات تهران** تاریخ: **۱۴۰۳/۱۱/۲۵**