معرفی سی و هشتمین وبینار تخصصی انجمن رمز ایران

معرفی سخنران:

آقای حمید بستانی دانشجوی سال آخر دوره دکتری علوم کامپیوتر در دانشکده محاسبات و علوم اطلاعات دانشگاه Radboud هلند هستند. پایان‌نامه کارشناسی ارشد وی به سال ۱۳۹۴ در دانشگاه آزاد اسلامی واحد تهران جنوب با موضوع بکارگیری یادگیری ماشین برای تشخیص نفوذ در اینترنت اشیاء بوده است که مقالات مستخرج شده از آن، جزء پژوهش‌های پیشگام در این حوزه محسوب شده و در مجلات معتبر علمی چاپ شده‌اند. ایشان در حال حاضر به عنوان پژوهشگر مهمان در دانشگاه‌های کینگز کالج لندن و کالج دانشگاهی لندن مشغول پژوهش بر روی امنیت یادگیری ماشین فعالیت می‌کنند. علاقه‌مندیِ تحقیقاتی ایشان در تلاقی بین هوش مصنوعی و امنیت سیستم‌ها به ویژه سیستم‌های تشخیص نفوذ و سیستم‌های تشخیص بدافزار قرار دارد. تمرکز فعلی پژوهشی آقای بستانی بر روی یادگیری ماشین خصمانه (Adversarial Machine Learning) در حوزه تشخیص بدافزار است. در واقع ایشان سعی دارند استحکام مدل‌های یادگیری ماشین مورد استفاده در تشخیص بدافزارها را در مقابل حملات خصمانه (Adversarial Attacks) افزایش دهد.

عنوان سخنرانی:

یادگیری ماشین خصمانه: چالش‌ها و راه‌حل‌ها در حوزه سیستم‌های تشخیص بدافزار

چکیده سخنرانی:

با وجود رشد سریع استفاده از یادگیری ماشین (ML: Machine Learning) برای شناسایی بدافزارها (Malwareها)، سیستم‌های تشخیص بدافزار مبتنی بر یادگیری ماشین در مقابل برنامه‌های بدافزار، که با هدف فریب دادن ML بطور دقیق ساخته شده‌اند، آسیب‌پذیر هستند. به عبارت دیگر، عوامل مخرب با تبدیل برنامه‌های بدافزار به نمونه‌های خصمانه (Adversarial Examples) اقدام به فریب سیستم‌های تشخیص بدافزار مبتنی بر یادگیری ماشین می‌کنند.

در دهه گذشته، بسیاری از مطالعات تلاش کرده‌اند تا حملات خصمانه را برای شناسایی آسیب‌پذیری سیستم‌های تشخیص بدافزار در برابر نمونه‌های خصمانه پیش‌بینی کنند.

با این حال، ارائه یک حمله خصمانه قابل تحقق (Realistic Adversarial Attack) در حوزه تشخیص بدافزار، به دلیل ماهیت متفاوت دامنه نرم‌افزار نسبت به سایر حوزه‌ها نظیر بینایی ماشین چالش های متعددی پیش رو دارد. از طرفی به‌رغم تلاش‌های تحسین‌برانگیز مطالعات قبلی در افزایش استحکام (Robustness) مدل‌های ML مورد استفاده در سیستم‌های تشخیص بدافزار نظیر آموزش خصمانه (Adversarial Training)، عملی بودن آن‌ها در ارائه استحکام خصمانه در برابر مدل‌های تهدید واقعی یک موضوع مورد مناقشه است. این سمینار ابتدا به معرفی مفهوم یادگیری ماشین خصمانه (AML: Adversarial Machine Learning) و چالش‌های پیش روی AML به ویژه در حوزه بدافزارها می‌پردازد. سپس مطالعات اخیر سخنران پیرامون حملات خصمانه قابل تحقق و همچنین مستحکم‌سازی سیستم‌های تشخیص بدافزار را مورد بررسی قرار خواهد گرفت. در انتها به معرفی مختصر زمینه‌های بالقوه پژوهشی AML در حوزه بدافزارها پرداخته می‌شود.