بسم الله الرّحمن الرّحیم

جلسه ارائه هفتگی آزمایشگاه امنیت داده و شبکه

یک روش تولید شاخص برای داده‌های رشته‌ای وابسته به توزیع احتمالاتی کاراکترها

هادی حلواچی

شنبه ٤ تیر؛ ساعت ۹ صبح

دانشکده کامپیوتر، آزمایشگاه امنیت داده و شبکه

یکی از روش‌های حفظ محرمانگی رمز کردن داده‌هاست. داده‌ها پس از رمز شدن خواص خود را از دست داده و اجرای پرسجو روی آن‌ها با چالش مواجه است. یکی از راه‌های ارائه شده برای تقابل با این چالش، استفاده از شاخص در کنار داده رمزشده است. روش‌های متعددی برای ایجاد شاخص روی داده‌های عددی و رشته‌ای ارائه شده است که هر کدام از انواع پرسجوهای خاصی پشتیبانی می‌کنند. در مورد داده‌های رشته‌ای، روش‌هایی که تاکنون ارائه شده‌اند، از پرسجوهای شامل تطبیق دقیق و یا الگویی پشتیبانی می‌کنند. بیشتر روش‌ها در حالت تطبیق الگویی، از جستجو روی هر الگوی دلخواهی پشتیبانی نمی‌کنند. هیچ کدام از این روش‌ها علاوه بر مزایا و معایبی که دارند، توزیع کاراکترها در رشته‌ها را مورد توجه قرار نداده‌اند در حالیکه احتمال رخدادن کاراکترهای مختلف یکسان نیست. با در نظر گرفتن تکرار کاراکترهای مختلف در رشته‌ها و احتمال رخدادن هر کاراکتر، می‌توان میزان برخوردهای اشتباه در شاخص را کاهش داد. در این مقاله روشی برای ایجاد شاخص روی داده‌های رشته‌ای ارائه می‌دهیم که در آن شاخص، با در نظر گرفتن احتمال رخداد هر کاراکتر در هر مکان از رشته، ایجاد می‌شود. در این صورت از طرفی چون مکان قرارگیری کارکترها در ایجاد شاخص تاثیرگذار هستند هر الگوی دلخواهی در شاخص قابل جستجو بوده و از طرف دیگر به علت تاثیر احتمال رخداد کاراکترها در ایجاد شاخص برخوردهای اشتباه کاهش یافته و در نتیجه کارایی بهبود می‌یابد.

شرکت در این جلسه برای تمامی دانشجویان علاقه‌مند آزاد است.