

# استفاده از تنوع قواعد دستوری و نوشتاری برای پنهان نگاری در زبان فارسی

سید علی اصغر شهیدیان<sup>۱</sup>، محمد صادق رستمی<sup>۲</sup> و خشایار یغمایی<sup>۳</sup>

<sup>۱</sup> دانشگاه سمنان- دانشکده برق AliShahidian2020@yahoo.com

<sup>۲</sup> دانشگاه سمنان- دانشکده برق MSR\_Anjir@yahoo.com

<sup>۳</sup> دانشگاه سمنان- دانشکده برق Khashayar.Yaghmaie@gmail.com

## چکیده

در این مقاله، روشی برای پنهان نگاری در متون فارسی بر اساس تنوع قواعد دستوری و نوشتاری، ارائه شده است. در این روش، گیرنده پیام و فرستنده ی آن بر اساس قواعد از پیش تعیین شده دستوری و نوشتاری، عملیات پنهان نگاری و بازیابی را انجام می دهند. متن ورودی، پردازش شده و افزونگی های آن مشخص میشود و با توجه به قواعدی، داده های پیام در آن متن قرار می گیرند. این روش با روش های معمول پنهان نگاری در متن، مانند تغییر پیکسل های تصویر متن، تفاوت اساسی دارد و در مقابل حملات مقاوم تر است هرچند که از نظر ظرفیت نسبت به آن ها محدودیت بیشتری دارد. روش پیشنهادی، جزو روش های پنهان نگاری زبان طبیعی محسوب می شود. این روش پنهان نگاری بلایند بوده و برای بازیابی اطلاعات نیازی به متن اصلی نیست و تنها با دانستن قواعد، می توان اطلاعات را بازیابی نمود. این روش نسبت به نویزهای معمول در عملیات اسکن و چاپ و همچنین پردازش های سیستم های تشخیص کاراکتر نوری، کاملاً مقاوم است. با توجه به آزمایشات انجام شده، این روش از عدم مشاهده پذیری بالایی برخوردار است.

## کلمات کلیدی

پنهان نگاری متن- قواعد دستوری و نوشتاری- زبان فارسی- افزونگی- زبان طبیعی- نشانه گذاری متن

می شود که در این صورت، متن حامل، اهمیت اطلاعاتی نداشته و فقط پیامی که در آن مخفی شده، مد نظر است.

## ۱- مقدمه

روش های موجود در استگانوگرافی<sup>۱</sup> یا پنهان نگاری و واترمارکینگ<sup>۲</sup> را می توان به دو گروه عمده تقسیم کرد. روش هایی که بر روی تصویر متن اعمال میشوند و روشهایی که از افزونگی های موجود در زبان طبیعی استفاده میکنند. این دو روش از نظر ظرفیت، مشاهده پذیری و مقاومت با یکدیگر متفاوت هستند.

در پنهان نگاری و واترمارکینگ، همواره مصالحه ای بین پارامتر های ظرفیت، مشاهده پذیری و مقاومت وجود دارد و افزایش یک پارامتر، کاهش دو پارامتر دیگر را منجر می شود.

در دهه ی گذشته، روش های پنهان نگاری، به طور گسترده در اسناد چند رسانه ای مورد بررسی قرار گرفته است [۱]. در مقابل، مطالعات بر روی پنهان نگاری در حوزه زبان طبیعی، تنها در چند سال اخیر مورد توجه قرار گرفته است [۲-۷].

با وجود افزایش داده های تصویری و فیلم در سال های اخیر، تعداد نسبی داده های نوشتاری منتقل شده، زیاد است. همچنین در بحث انتقال اطلاعات، داده های نوشتاری از اهمیت زیادی برخوردارند. اکثر مجله ها، روزنامه ها، مقاله های علمی و کنفرانس ها، در قالب دیجیتال نیز ارائه می شوند. این مسئله راه های جستجو و دسترسی را برای خوانندگان مهیا ساخته، اما نحوه ی توزیع و استفاده از آن ها، نویسندگان را نگران کرده است.

بحث حقوقی در اسناد متنی، بسیار مهمتر از اسناد تصویری و فیلم است، چون دسترسی و دستکاری متن، توسط کاربران بسیار راحت تر انجام می پذیرد. یکی از روش هایی که برای تشخیص اصالت متن مورد استفاده گسترده قرار گرفته است پنهان سازی يك کد عددی در داخل متن است. در برخی کاربردها نیز به کمک یک متن حامل، پیامی برای گیرنده ارسال

<sup>1</sup> Steganography

<sup>2</sup> Watermarking