

# روش جدید برای واترمارکینگ متون فارسی

## با استفاده از "ها" و "می"

وحید یزدانی<sup>۱</sup>، محمد علی دوستاری<sup>۲</sup>، حمید یزدانی<sup>۳</sup>

<sup>۱</sup> دانشجوی کارشناسی ارشد رشته فن آوری اطلاعات دانشگاه شاهد

v.yazdani@shahed.ac.ir

<sup>۲</sup> استادیار دانشگاه شاهد

doostari@shahed.ac.ir

<sup>۳</sup> مربی دانشگاه آزاد اسلامی واحد نور

eng.hamid.yazdani@gmail.com

### چکیده

با گسترش ارتباطات، در بعضی از مواقع نیاز به ارتباطات مخفی بیش از پیش احساس می‌شود. واترمارکینگ یکی از مهمترین روش‌های مخفی‌سازی اطلاعات است که برای محافظت از حق مالکیت از محتوی چند رسانه‌ها استفاده می‌شود. رسانه متنی درمیان انواع رسانه‌های مختلف کاربرد زیادی داشته و این در حالی است که، در زمینه واترمارکینگ متون فارسی یا عربی با استفاده از خصوصیات حروف و کلمات به کار رفته در متن، به نسبت متون انگلیسی و چینی تحقیقات کمتری انجام شده است. در این مقاله روش جدیدی برای واترمارکینگ متون فارسی پیشنهاد می‌شود که از علامت جمع "ها" برای اسامی و پیشوند "می" برای افعال جهت مخفی‌سازی اطلاعات استفاده کرده، که دارای استحکام و امنیت بهتری نسبت به روش‌های مشابه است.

### کلمات کلیدی

متون فارسی، مخفی‌سازی اطلاعات، پنهان‌نگاری، واترمارکینگ.

داشته، و کلیه افرادی که دارای کلید باشند قادر به رمزگشایی اطلاعات محرمانه می‌باشند. بر خلاف واترمارکینگ در رمزنگاری هیچ محافظتی از محتوی پیام پس از رمزگشایی وجود ندارد.

واترمارکینگ، اطلاعات محرمانه را در بیت‌های زائد و غیراستفاده در هر نوع رسانه پوششی مانند فایل‌های صوتی و تصویری قرار می‌دهد. مخفی کردن چنین بیت‌هایی در متن به دلیل وجود بیت‌های زائد کم، بسیار سخت خواهد بود، اما واترمارکینگ متنی به علت اندازه کوچکتر، انتقال سریع‌تر و آسان‌تر فایل‌های متنی بر روی شبکه، اهمیت زیادی دارد [۲].

برخلاف واترمارکینگ متنی، روش‌های بسیار زیاد و متنوعی برای واترمارکینگ دیگر انواع رسانه وجود دارد. این در حالی است که در زمینه واترمارکینگ متون فارسی با استفاده از خصوصیات حروف و کلمات به کار رفته در متن، همانند متون چینی و انگلیسی تحقیقات چندانی صورت نگرفته است [۱]. به همین دلیل در این مقاله روش جدیدی برای واترمارکینگ متون فارسی پیشنهاد می‌شود که از پسوند علامت جمع "ها" برای اسامی و پیشوند "می" برای برخی از افعال، جهت مخفی‌سازی اطلاعات استفاده کرده، که دارای استحکام و امنیت

### ۱- مقدمه

در قرن ارتباطات، به دلیل گسترش روزافزون ارتباطات جهانی و ابداع کانال‌های ارتباطی گوناگون نظیر شبکه‌های اینترنت، ارتباطات ماهواره‌ای و مخابراتی دیگر، اطلاعات به راحتی در اختیار طیف گسترده‌ای از مردم در سرتاسر دنیا قرار می‌گیرد. در نتیجه نیاز به ارتباطات مخفی بیش از پیش احساس می‌شود [۱].

در این میان، روش‌های پنهان‌نگاری و واترمارکینگ به عنوان مهمترین روش‌های مخفی‌سازی اطلاعات محسوب می‌شوند. واترمارکینگ با پنهان‌نگاری و رمزنگاری متفاوت است. واترمارکینگ می‌تواند به عنوان یکی از روش‌های پنهان‌نگاری دیده شود که تمرکز بیشتر بر روی استحکام و کمتری بر روی امنیت دارد. امنیتی که پنهان‌نگاری و واترمارکینگ تامین می‌کنند، با امنیت تامین شده توسط رمزنگاری متفاوت است. امنیت واترمارکینگ توسط عدم آگاهی دیگران از وجود اطلاعات مخفی شده در رسانه، تامین می‌شود ولی در رمزنگاری اطلاع از این که در یک رسانه، اطلاعاتی مخفی است وجود

بهتری نسبت به روش‌های مشابه است. مابقی مقاله به صورت زیر سازماندهی شده است:

در بخش دوم، دسته‌بندی الگوریتم‌های واترمارکینگ و پارامترهای مؤثر در ارزیابی آن معرفی خواهد شد. در بخش سوم کارهای انجام شده در زمینه واترمارکینگ فارسی و عربی بررسی می‌شود. بخش چهارم مربوط به معرفی روش جدید برای واترمارکینگ متون فارسی است. و بخش آخر شامل جمع‌بندی روش پیشنهادی خواهد بود.

## ۲- دسته‌بندی الگوریتم‌های واترمارکینگ و پارامترهای مؤثر در ارزیابی آن

روش‌های واترمارکینگ را به صورت‌های زیر می‌توان دسته‌بندی کرد:

۱- بر اساس قابل مشاهده بودن بیت‌های واترمارک: الف) قابل مشاهده: بیت‌های واترمارک‌های تعبیه شده در هنگام رویت محتوی رسانه با یک روشی قابل دیدن هستند. ب) نامرئی: بیت‌های واترمارک‌های تعبیه شده با چشم قابل دیدن نبوده ولی با یک الگوریتم رمزگشایی مناسب می‌توان دید. ۲- بر اساس محتوی یا نوع رسانه میزبان:

الف) واترمارکینگ تصاویر  
ب) واترمارکینگ ویدئو  
پ) واترمارکینگ صدا  
ج) واترمارکینگ متن  
۳- بر اساس نیاز به سند اصلی هنگام استخراج: الف) کور: در هنگام فرایند استخراج، نیاز به سند اصلی نیست. ب) غیرکور: در هنگام فرایند استخراج، نیاز به سند اصلی است. ۴- بر اساس میزان استحکام:

الف) واترمارکینگ مقاوم: تغییر یا اصلاح محتوی رسانه تاثیری روی بیت‌های واترمارک ندارد. ب) واترمارکینگ ضعیف: تغییر یا اصلاح محتوی رسانه بر روی بیت‌های واترمارک اثر می‌گذارد [۳].

در ارزیابی الگوریتم‌هایی که برای ایجاد و کشف واترمارکینگ استفاده می‌شوند، پارامترهای زیر متداول است:

۱- استحکام: مربوط به نحوه جاسازی اطلاعات بوده، به طوری که در برابر حملات پایدار باشد. این پارامتر هرچه بالاتر باشد نشانه آن است که میزان تخریب و اثرپذیری بیت‌های واترمارک در اثر تغییرات در رسانه میزبان کمتر است.

۲- ظرفیت: این پارامتر مشخص می‌کند که حداکثر حجم ممکن که رسانه برای درج بیت‌های واترمارک قادر است فراهم کند، چقدر می‌باشد.

۳- امنیت: تعیین می‌کند که رسانه میزبان تا چه میزان وجود بیت‌های واترمارک را می‌تواند مخفی نگه داشته و باعث جلوگیری از دستکاری، جعل و حذف پیام شود [۵].

## ۳- بررسی کارهای انجام شده در زمینه واترمارکینگ متون فارسی و عربی

با توجه به خصوصیات حروف فارسی و عربی و همچنین پیوستگی آن‌ها در کلمات، کارهای کمی در زمینه واترمارکینگ متون فارسی انجام شده است. از جمله کارهای انجام شده در زمینه واترمارکینگ متون فارسی و عربی، می‌توان به موارد زیر اشاره کرد:

### ۳-۱- روش شبه فاصله

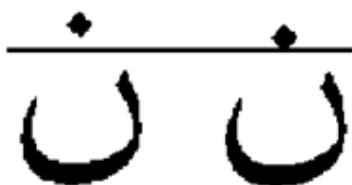
از آنجایی که در زبان فارسی از نیم فاصله در برخی از کلمات استفاده می‌شود (مانند فاصله بین "می" و "باشد" در کلمه "می-باشد") می‌توان از آن برای درج بیت‌های واترمارکینگ استفاده کرد به طوری که در هنگام برخورد به نیم فاصله بدین صورت عمل می‌شود که برای درج بیت صفر به نیم فاصله چیزی اضافه نکرده و برای درج بیت یک به نیم فاصله یک اضافه خواهد شد.

ظرفیت این روش پایین بوده و فقط در متونی که نیم فاصله در آن‌ها به کار رفته، قابل استفاده است. این روش دارای امنیت پایینی بوده ولی با تغییر در نوع و اندازه قلم متن پوششی اطلاعات محرمانه پایدار خواهد ماند [۶].

### ۳-۲- روش نقطه

زبان فارسی شامل ۱۸ حرف دارای نقطه می‌باشد. از آنجایی که تعداد حروف دارای نقطه در متون فارسی و عربی زیاد است، این روش از نقاط حروف برای درج اطلاعات محرمانه بهره می‌گیرد. بدین صورت که برای درج بیت صفر مکان نقطه در حروف دارای نقطه تغییری نکرده و برای درج بیت یک، مکان نقطه در حروف دارای نقطه کمی به سمت بالا انتقال می‌یابد. شکل (۱) نحوه تغییر مکان نقطه در حرف "ن" را نشان می‌دهد.

از ویژگی‌های این روش می‌توان به ظرفیت بالا و استحکام پایین اشاره کرد. همچنین در این روش با تغییر اندازه و نوع قلم و نوشتن مجدد متن پوششی اطلاعات محرمانه از بین خواهد رفت [۷].



شکل (۱): نمونه ای از واترمارکینگ با جابجایی نقاط [۷]

### ۳-۳- روش استفاده از علائم موجود بر حروف

در این روش جهت مخفی سازی داده بر روی متن پوشش ابتدا باید مطمئن شد که کلیه علائم در پوشش متنی حاضر هستند. سپس به ترتیب کلیه بیت‌هایی که قرار است مخفی شوند، متناظر با یک علامت قرار می‌گیرند. حال اگر مقدار بیتی که قرار است مخفی شود یک بود، علامت متناظر با آن بیت، بدون هیچ تغییری در متن حاضر خواهد بود. ولی اگر مقدار بیتی که قرار است مخفی شود صفر بود، علامت متناظر با آن بیت، از متن حذف خواهد شد [۴]. شکل (۲) نمونه‌ای از این نوع واترمارکینگ را نشان می‌دهد.

متن پوششی: مُسْتَفْعِلْ

بیت‌های محرمانه: 0 1 1 0 0 1

متن استگانوگرافی: مُسْتَفْعِلْ

شکل (۲): نمونه ای از واترمارکینگ با استفاده از علائم [۱]

در رویکرد دیگری از این روش، از علامت فتحه برای درج بیت یک، و از سایر علائم برای درج بیت صفر استفاده می‌شود. هرچند ممکن است به علت درج بیت‌های یک و یا صفر متوالی و عدم تناسب علائم موجود بر حروف با مقدار بیت ذخیره شده، از همه علائم استفاده بهینه نشود [۱].

از مزایای این روش‌ها علاوه بر ظرفیت بالا به کارگیری آسان آن بوده و در برابر حملاتی نظیر OCR و چاپ مقاوم هستند. ولی این روش‌ها فاقد استحکام بوده، و از این رو بهتر است برای واترمارکینگ ضعیف یا تمامیت داده استفاده شود.

### ۳-۴- روش استفاده از حروف مشابه با کدهای متفاوت

حروف فارسی و عربی در چهار حرف اختلاف دارند. مابقی حروف تقریباً شبیه به هم بوده و تنها با توجه به مکان قرارگیری‌شان در کلمه تفاوت جزئی خواهند داشت، که از جمله آن می‌توان به حروف "ک" و "ی" فارسی با "ک" و "ی" عربی اشاره کرد. جدول (۱) تفاوت شکل ظاهری این دو حرف را در آغاز، وسط، انتها و جدا از کلمه در دو زبان فارسی و عربی نشان می‌دهد.

در این روش برای درج بیت صفر از شکل عربی حروف "ک" و "ی"، و برای درج بیت یک از شکل فارسی حروف "ک" و "ی" استفاده می‌شود. در حالت‌هایی که شکل دو حروف در دو زبان فارسی و عربی تفاوتی ندارد (مانند حرف "ی" در ابتدای کلمات)، اطلاعاتی درج نخواهد شد. ظرفیت این روش پایین بوده ولی دارای شفافیت اداری بالایی است.

همچنین این روش در برابر انتقال به محیط‌های دیجیتالی مانند HTML, Pdf, Word مستحکم بوده ولی برا اثر نوشتن مجدد متن پوششی اطلاعات محرمانه از بین خواهد رفت [۸].

جدول (۱): تفاوت شکل ظاهری حروف "ک" و "ی" [۸]

وضعیت کلمه	زبان	حرف "ی"	حرف "ک"
آغاز	فارسی	ی	ک
	عربی	ی	ک
وسط	فارسی	ی	ک
	عربی	ی	ک
انتها	فارسی	ی	ک
	عربی	ي	ك
جدا	فارسی	ی	ک
	عربی	ي	ك

### ۳-۵- روش استفاده از کلمه "لا"

در این روش از یک کلمه یکسان که دارای دو شکل ظاهری متفاوت بوده و در سیستم استاندارد یونیکد، دو کد متفاوت دارد، برای درج بیت‌های واترمارک استفاده می‌شود. از جمله این کلمات، کلمه "لا" می‌باشد. در کلمه "لا" برای درج بیت یک از "لا" و برای درج بیت صفر از شکل عادی، یعنی "لا" استفاده خواهد شد. کلمه "لا" دارای کد FEFB و کلمه "لا" دارای کد ۰۶۴۰ در سیستم استاندارد یونیکد است.

هرچند ظرفیت این روش پایین بوده ولی این روش محدود به اسناد الکترونیکی نبوده و بر اثر چاپ، اطلاعات محرمانه پایدار باقی خواهد ماند. در این روش، نوشتن مجدد متن پوششی باعث از بین رفتن اطلاعات محرمانه خواهد شد [۵].

### ۳-۶- روش استفاده از کاراکتر کشیده

کاراکتر کشیده عربی در نوع الکترونیکی تنها برای مرتب‌سازی و فرمت‌بندی به عنوان یک کاراکتر عادی در نظر گرفته می‌شود. نکته قابل ذکر در استفاده از کاراکتر کشیده این است که تمام حروف به دلیل موقعیت‌شان در کلمات و ماهیت نوشتاری زبان عربی قابل گسترش نیستند. همچنین کاراکتر کشیده نمی‌تواند پس از حرف پایانی کلمه یا قبل از شروع اولین حرف کلمه اضافه شود. کاراکتر کشیده را می‌توان به بعد یا قبل از حروف که قابلیت اضافه شدن کاراکتر کشیده را دارند، اضافه کرد.

در این روش برای درج بیت صفر کاراکتر کشیده به قبل یا بعد از حرف فاقد نقطه و برای درج بیت یک کاراکتر کشیده به قبل یا بعد از حرف دارای نقطه اضافه که قابلیت اضافه شدن کاراکتر کشیده را دارند، تعبیه خواهد شد. شکل (۳) یک نمونه از جزئیات فرآیند این نوع از واترمارکینگ در حالت درج کاراکتر کشیده به بعد از حروف را نشان می‌دهد.

## جدول (۲): کاربرد "ها" و "می" در اسامی و افعال زبان فارسی

مثال	نوع اسم یا فعل	نوع پیشوند یا پسوند
بیت‌ها	اسامی جمع	ها
می‌رفتم	ماضی استمراری	می
می‌رفته‌ام	ماضی نقلی مستمر	
می‌رود	مضارع اخباری	

همانطوری که در جدول (۲) مشاهده می‌شود، پسوند "ها" و پیشوند "می" با یک نیم فاصله به ترتیب به بعد و قبل از اسامی و افعال اضافه می‌شود. در روش پیشنهادی از این دو ویژگی برای درج اطلاعات محرمانه در متون فارسی استفاده خواهد شد. بدین گونه که می‌توان برای درج یک سطح از اطلاعات، این نیم فاصله را نادیده گرفته و اسامی و افعالی که دارای این پسوند و پیشوند هستند، به صورت یکسره نوشته شود. یعنی برای درج بیت یک، نیم فاصله بین پسوند "ها" با اسم مربوطه و یا نیم فاصله بین پیشوند "می" با فعل مربوطه حذف شده و کلمه به صورت پیوسته نوشته شود. برای تعبیه بیت صفر نیز، هیچگونه تغییری در بین این پسوند و پیشوند با اسم و فعل مربوطه انجام نمی‌شود.

هرچند ظرفیت این روش پایین است ولی می‌توان با استفاده از جدول نگاشت این ویژگی را بهبود بخشید. در این جدول به هر حرف به جای ۱۶ بیت (با توجه به طرح کدگذاری UTF-8 که جهت نمایش کاراکتر عربی یا فارسی از ۱۶ بیت استفاده می‌کند)، ۶ بیت اختصاص داده می‌شود. اولین حرف از جدول نگاشت دارای مقدار ۰۰۰۰۰۰ بوده و برای مابقی حروف به این مقدار یک واحد به ترتیب حروف الفبا اضافه خواهد شد. در نتیجه با صرفه‌جویی در ۱۰ بیت، ظرفیت به‌طور چشمگیری زیاد خواهد شد. نکته دیگر در استفاده از جدول نگاشت، بالا رفتن امنیت روش پیشنهادی است. شکل (۴) مثالی از این روش و جدول (۳)، جدول نگاشت برای پیام محرمانه را نشان می‌دهد.

بد	پیام محرمانه
فعل‌های ساده فعل‌هایی می‌باشند که مصدر آن‌ها صرف‌نظر از علامت‌های مصدری، بیش از یک کلمه نیست. فعل‌های جعلی را فعل‌هایی بر می‌شماریم که بن فعل ندارند. فعل‌های جعلی جزء فعل‌های ساده محسوب می‌شوند. فعل‌های پیشوندی فعل‌هایی می‌باشند که از یک پیشوند و یک فعل ساده ساخته می‌شوند.	متن پوششی
فعل‌های ساده فعل‌هایی می‌باشند که مصدر آن‌ها صرف‌نظر از علامت‌های مصدری، بیش از یک کلمه نیست. فعل‌های جعلی را فعل‌هایی بر می‌شماریم که بن فعل ندارند. فعل‌های جعلی جزء فعل‌های ساده محسوب می‌شوند. فعل‌های پیشوندی فعل‌هایی می‌باشند که از یک پیشوند و یک فعل ساده ساخته می‌شوند.	متن خروجی

شکل (۴): نمونه‌ای از واترمارک‌کینگ روش پیشنهادی

جهت افزایش و گمراهی بیشتر، می‌توان از هر دو حالت درج کاراکتر کشیده به قبل یا بعد از حروف در یک متن پوششی بهره برد. بدین‌گونه که برای بعضی از خطوط یا پارگراف مشخص در متن، از حالت درج کاراکتر کشیده به قبل از حروف، و برای مابقی خطوط یا پارگراف از حالت درج کاراکتر کشیده به بعد از حروف استفاده کرد.

بیت‌های محرمانه	۱۱۰۰۱۰
من حسن اسلام المرء ترکه مالا یعنیہ	متن پوششی
من حسن اسلام المرء ترکه مالا یعنیہ	متن خروجی
	۱ ۱ ۰ ۰ ۱

شکل (۳): نمونه‌ای از واترمارک‌کینگ با استفاده از حروف کشیده [۲]

در مقایسه این روش با روش نقطه این نکته قابل ذکر است که در این روش، مخفی کردن یک بیت در متن، معادل درج یک کاراکتر کشیده می‌باشد. در صورتی که روش نقطه فاقد چنین افزایش اندازه‌ای است. در نتیجه دارای ظرفیت کمتری نسبت به روش نقطه می‌باشد. در مقابل این روش از امنیت و استحکام بالاتری نسبت به روش نقطه برخوردار بوده و اطلاعات محرمانه در برابر تغییر نوع و اندازه قلم نیز، پایدار خواهد ماند [۲].

## ۴- روش پیشنهادی

متون زبان فارسی بر خلاف متون انگلیسی از راست به چپ نوشته می‌شوند. زبان فارسی دارای ۳۲ حرف بوده که ۱۸ حرف از ۳۲ حرف آن دارای ۱ تا ۳ نقطه می‌باشند. نقاط ممکن است در بالای حرف مانند "ت" و یا در پایین حرف مانند "پ" و یا در میان حرف مانند "چ" قرار بگیرند، در حالیکه در زبان انگلیسی فقط دو حرف "i" و "j" دارای نقطه می‌باشند. بر خلاف زبان انگلیسی، حروف در کلمات فارسی به هم پیوسته بوده و با توجه به مکان قرارگیری حرف در کلمه ممکن است شکل‌های متفاوتی داشته باشند. به عنوان مثال حرف "ع" با توجه به مکان قرارگیری در کلمه به صورت‌های "ع"، "عـ"، "عـ" و "ع" نوشته می‌شود [۶]. تفاوت حروف فارسی و عربی در چهار حرف "پ"، "چ"، "ز" و "گ" بوده و مابقی حروف تقریباً شبیه به هم بوده و تنها با توجه به مکان قرارگیری‌شان در کلمه تفاوت جزئی خواهند داشت.

معمولاً در زبان فارسی از "ها" برای جمع اسامی استفاده می‌شود. همچنین یکی از پیشوندهای پرکاربرد در افعال فارسی پیشوند "می" است. جدول (۲) چند مثال از کاربرد "ها" و "می" در اسامی و افعال زبان فارسی را نشان می‌دهد.

*Extension Character 'Kashida' With Pointed Letters For Arabic Text Digital Watermarking*" International Conference on Security and Cryptography-(SECRYPT), Barcelona, Spain, 28 – 31 July 2007.

- [3] Jaseena K.U., Anita John, "Text Watermarking using Combined Image and Text for Authentication and Protection", International Journal of Computer Applications (0975 – 8887) Volume 20– No.4, April 2011.
- [4] Bensaad, M.L., Yagoubi, M.B., " High capacity diacritics-based method for information hiding in Arabic text", International Conference on Innovations in Information Technology (IIT), 2011, vol.11, no.12, pp. 433 - 436, 16 June 2011.
- [5] Shirali-Shahreza, M, "A New Persian/Arabic Text Steganography Using "La" Word", in Proceedings of the International Joint Conference on Computer, Information, and Systems Sciences, and Engineering (CISSE 2007), Bridgeport, CT, USA, 2(2007), pp. 339–342, 2007.
- [6] M. Shirali-Shahreza., "Pseudo-space Persian/Arabic text steganography", Computers and Communications, 2008. ISCC 2008. IEEE Symposium on, 6-9 July 2008.
- [7] M.H. Shirali-Shahreza and M. Shirali-Shahreza, "A New Approach to Persian/Arabic Text Steganography", Proceedings of the 5th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2006), Honolulu, HI, USA., pp. 310-315, July 10-12, 2006.
- [8] M. Hassan Shirali-Shahreza, Mohammad Shirali-Shahreza, "Arabic/Persian Text Steganography Utilizing Similar Letters With Different Codes", The Arabian Journal for Science and Engineering, Volume 35, December 9, 2009.

جدول (۳) : جدول نگاشت

د	ب	حرف
۰۰۱۰۱۰	۰۰۰۰۰۱	کد

اولین حرف در شی محرمانه، حرف "ب" بوده که دارای کد ۰۰۰۰۰۱ در جدول نگاشت می‌باشد. برای تعبیه پنج صفر اول از بیت محرمانه، به ترتیب کلمات اول، سوم، چهارم، هفتم و یازدهم از متن پوششی هیچ تغییری نکرده و برای درج بیت یک، کلمه هجدهم متن به صورت پیوسته نوشته می‌شود (" فعلهای"). در شکل ۴ کلمات مشخص شده با رنگ قرمز نشان‌دهنده بیت صفر، و کلمات مشخص شده با رنگ آبی نشان‌دهنده بیت یک است.

همچنین می‌توان جهت گمراهی بیشتر بعد از درج آخرین بیت از پیام محرمانه، به طور تصادفی پسوند "ها" و پیشوند "می" را به اسامی و افعال متصل کرد. نکته قابل ذکر این است در هنگام استخراج، داشتن طول بیت محرمانه و جدول نگاشت لازم است.

## ۵- نتیجه

در این مقاله روش جدیدی برای واترمارکینگ متون فارسی ارائه شده است که با استفاده از ، پسوند "ها" و پیشوند "می" که با یک نیم فاصله به ترتیب به بعد و قبل از اسامی و افعال اضافه می‌شود، اقدام به درج اطلاعات محرمانه می‌کند.

تفاوت این روش با سایر روش مشابه را می‌توان در استحکام و امنیت بالای آن دانست. در روش پیشنهادی با استفاده جدول نگاشت طول بیت برای نمایش هر حرف کوتاه‌تر شده، و در نتیجه محیط میزبان می‌تواند اطلاعات بیشتری را در خود جای دهد. در این روش، از آنجایی که از جدول نگاشت استفاده می‌شود، و برای هر حرف به جای استفاده از ۱۶ بیت به ۶ بیت یا کمتر نیاز است، حتی با استخراج بیت-ها، تشخیص اینکه کدام بیت‌ها مربوط به یک حرف است، بسیار سخت می‌باشد. در نتیجه امنیت این روش نیز بالاست.

همچنین این روش در برابر چاپ و انتقال به محیط‌های دیجیتالی مختلف مانند Pdf و HTML مستحکم بوده، و با تغییر اندازه و نوع قلم، اطلاعات محرمانه از بین نخواهد رفت.

ظرفیت این روش پایین بوده ولی با استفاده از جدول نگاشت سعی شده است که این ویژگی بهبود یابد. همچنین در اثر نوشتن مجدد متن پوششی، اطلاعات محرمانه از بین خواهد رفت.

## مراجع

- [1] A. Gutub Y. S. Elarian, S. M. Awaideh and A. K. Alvi, "Arabic Text Steganography Using Multiple Diacritics", International Workshop on Signal Processing and its Applications, Sharjah, U.A.E, 18–20 March 2008.
- [2] Adnan Gutub, Lahouari Ghouti, Alaaeldin Amin, Talal Alkharobi, and Mohammad K. Ibrahim, "Utilizing