نهان‌نگاری در رایانامه با ظرفیت نامحدود از طریق لغت‌نامه

محسن رضوانی[1*] و منصور فاتح[2]

[1 و 2] دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، ایران

چکیده

هدف اصلی در نهان‌نگاری پنهان‌سازی یک پیام مخفی با قراردادن آن پیام در یک رسانه پوشانه است؛ به‌نحوی که کمینه تغییرات در رسانه ایجاد شده و آن تغییرات به‌راحتی قابل درک نباشد. رسانه پوشانه می‌تواند یک بستر قابل دسترس توسط عموم نظیر متن، رایانامه، صوت، تصویر یا ویدئو باشد. با گسترش استفاده از رایانامه در بین کاربران اینترنتی، ارائه روش‌های نهان‌گاری در بستر رایانامه مورد توجه قرار گرفته است؛ ولی روش‌های موجود دارای محدودیت در ظرفیت نهان‌نگاری بوده و به‌طورعمومی مصالحه‌ای بین امنیت و ظرفیت نهان‌نگاری در نظر می‌گیرند. در این مقاله یک روش نوین برای نهان‌نگاری رایانامه ارائه شده است که مبتنی بر لغت‌نامه بوده و هم‌زمان ظرفیت نامحدود و امنیت بالایی را ارائه می‌کند. در گام نخست روش پیشنهادی، پیام به‌وسیله یک لغت‌نامه فشرده و رمزشده و سپس به یک رشته‌بیتی تبدیل می‌شود. در هر مرحله با توجه به تعداد نویسه‌های محتوای رایانامه، قسمتی از رشته انتخاب‌شده، معادل ده‌دهی آن محاسبه شده و سپس با توجه به کلیدهای موجود، با آن‌ها نشانی‌های رایانامه ساخته می‌شود. ظرفیت نهان‌نگاری نامحدود در روش پیشنهادی منجر به امکان مخفی‌سازی هر میزان پیام در متن پوشانه شده است. همچنین نتایج آزمایش‌ها نشان می‌دهد که استفاده از لغت‌نامه منجر به کاهش حجم پیام و همچنین کاهش تعداد نشانی‌های گیرنده به میزان حدودی ۴۴ درصد در مقایسه با روش‌های موجود شده است. این مهم به‌طور مستقیم به افزایش سطح امنیت روش پیشنهادی کمک می‌کند.

واژگان کلیدی: نهان‌نگاری رایانامه، ظرفیت، امنیت، لغت‌نامه

A High Capacity Email Steganography Scheme using Dictionary

Mohsen Rezvani[1*] & Mansoor Fateh[2]

[1]Faculty of Computer Engineering, Shahrood University of Technology, Shahroud, Iran

[2]Faculty of Computer Engineering, Shahrood University of Technology, Shahroud, Iran

Abstract

The expansion of the use of information exchange space and public access to communication networks such as the Internet has led to the growing dependence of social institutions on the use of these networks. However, maintaining the security of information exchanged on networks is one of the most important challenges for users of these networks. One way to protect this data is to use private networks. But building these networks is not cost-effective in terms of time and cost. In contrast, the use of encryption techniques, access control mechanisms and data concealment are among the effective solutions for security in the information exchange space.

Existing methods for hiding information can be divided into three categories: cryptography, watermarking and steganography. In cryptography, a simple text is converted into encrypted text, which, of course, requires a decryption operation as well as an encryption key. In general, cryptographic techniques suffer from two major problems. The first problem is the ban on the transmission of encrypted data in dictatorial regimes, and the second problem is that cryptographers pay attention to

* نویسندهٔ عهده‌دار مکاتبات

* Corresponding author

• تاریخ ارسال مقاله: ۱۳۹۸/۴/۲۴ • تاریخ پذیرش: ۱۳۹۸/۱۱/۱۴ • تاریخ انتشار: ۱۴۰۱/۰۳/۳۱ • نوع مطالعه: پژوهشی

[DOI: 10.52547/jsdp.19.1.59]

encrypted data and stop any secret communication. The second category of information hiding methods is watermarking. Watermarking techniques are commonly used to protect the copyright of a digital content and to deal with issues such as fraud, fraud and copyright infringement in the data transfer space. In steganography methods, the transfer of information takes place in a cover through public communication channels, and only the sender and receiver are aware of the existence of a secret message. Two aspects of steganography must be observed. The first aspect is that the cover and secret content look the same in the face of statistical attacks. The second aspect is that the process of hiding the secret message in the cover is such that there is no difference between the cover and the secret in terms of the human perceptual system. In fact, the accuracy of the transmission media is maintained.

Steganography methods use image, video, protocol, audio, and text platforms to hide information. Steganography in the text is difficult due to very little local variation. Humans are very sensitive to textual changes. Hence it is difficult to spell in the text. However, due to the high use of text in digital media, the insensitivity of text to compression, the need for less memory to store and communicate more easily and faster, many methods for steganography have been introduced in it. In addition, text is still one of the major forms of communication available to the general public around the world.

In this paper, we propose a new email steganography scheme using a dictionary-based compression. In the proposed scheme, a number of email addresses containing a hidden message will be generated using the submitted text. The submitted text is sent to the generated and recipient addresses at the same time. This does not reveal the identity of the recipient of the message, and only the recipient can extract secret message using other email addresses. In the proposed method, two steganography keys are used. Using these two keys increases the security level of the proposed method. Also, the capacity of the proposed method is unlimited, which of course is a great advantage in a steganography method. This unlimited capacity provides high security for the proposed method. Another advantage is that the proposed method is not limited to the type of the cover-text. Initially, the secret message is converted to a bit string by a dictionary. Then the operation of embedding the secret message in the recipient's addresses is done by the steganography keys.

The efficiency of steganography algorithms depends on various factors such as lack of detection by the human eye, lack of detection by statistical methods, and capacity. The proposed method does not change the cover-text. Hence, this method is not detectable by humans or statistical methods. The capacity of the proposed method in this research is based on built-in email addresses. As the text of the message increases, the number of emails created increases too. Of course, this increase in the address of the emails created can lead to suspicion of the emails sent. Therefore, the parameter of the number of emails created is also important in the evaluation. In this paper, the efficiency of the proposed method is evaluated based on the two parameters and compared with existing methods. The results of this evaluation show that the proposed method, in addition to providing unlimited capacity in steganography, produces fewer email addresses generated as well as fewer message bits after compression.

Keywords: Email Steganography, Dictionary, Capacity, Security.

۱- مقدمه

گسترش استفاده از فضای تبادل اطلاعات و دسترسی همگانی شبکه‌های ارتباطی نظیر شبکه اینترنت، منجر به وابستگی روزافزون نهادهای اجتماعی به استفاده از این شبکه‌ها شده است. با این حال، حفظ امنیت اطلاعات تبادل شده در شبکه‌ها یکی از مهم‌ترین چالش‌های موجود برای کاربران این شبکه‌ها است. یکی از راه‌های حفظ امنیت این داده‌ها، استفاده از شبکه‌های خصوصی است؛ اما ساخت این شبکه‌ها از لحاظ زمانی و هزینه‌ای بهینه نیست. در مقابل استفاده از روش‌های رمزنگاری، سازوکارهای کنترل دسترسی و مخفی‌سازی داده‌ها از جمله راه‌کارهای مؤثر برای امنیت در فضای تبادل اطلاعات است [8].

روش‌های موجود برای مخفی‌سازی اطلاعات را می‌توان به سه دستهٔ رمزنگاری[1]، ته‌نقش‌نگاری[2] و نهان‌نگاری[3] تقسیم کرد. در رمزنگاری یک متن ساده به یک متن رمزشده تبدیل می‌شود که البته برای درک آن انجام عملیات رمزگشایی و همچنین کلید رمزنگاری ضروری است. به‌طورکلی روش‌های رمزنگاری از دو مشکل اساسی رنج می‌برند. نخستین مشکل، ممنوعیت انتقال داده‌های رمزشده در حکومت‌های دیکتاتوری و دومین مشکل، جلب توجه رمزگشاها[4] به داده‌های رمزشده و متوقف‌کردن هرگونه ارتباط مخفی است [9]. دستهٔ دوم روش‌های مخفی‌سازی اطلاعات، ته‌نقش‌نگاری‌ها هستند. روش‌های ته‌نقش‌نگاری به‌طورعمومی با هدف حفظ حق

[1] Cryptography
[2] Watermarking
[3] Steganography
[4] Interceptors

[DOI: 10.52547/jsdp.19.1.59]

نشر یک محتوای دیجیتالی و مقابله با مسائلی مانند تقلب، کلاه‌برداری و نقض قانون انتشار در فضای انتقال داده استفاده می‌شود [10]. در روش‌های نهان‌نگاری، انتقال اطلاعات در یک پوشانه[1] و از طریق کانال‌های ارتباط عمومی انجام می‌شود و تنها فرستنده و گیرنده از وجود پیام مخفی اطلاع دارند. دو جنبه در نهان‌نگاری باید رعایت شوند. جنبهٔ نخست این‌که محتوای پوشانه و نهانه در مقابل حملات آماری یکسان به نظر برسند. جنبهٔ دوم اینکه فرآیند پنهان‌سازی پیغام مخفی در پوشانه به‌گونه‌ای باشد که از نظر دستگاه ادراکی انسان تفاوتی بین پوشانه و نهانه وجود نداشته باشد. درواقع صحت رسانهٔ انتقال حفظ شود [11].

روش‌های نهان‌نگاری از بسترهای تصویر، ویدئو، پروتکل، صوت و متن برای مخفی‌سازی اطلاعات استفاده می‌کنند. نهان‌نگاری در متن، به‌دلیل تغییرات محلی بسیار کم، مشکل است. انسان به تغییرات متنی بسیار حساس است؛ ازاین‌رو نهان‌نگاری در متن دشوار است؛ اما به‌دلیل استفاده زیاد متن در رسانه‌های دیجیتال، حساس نبودن متن نسبت به فشرده‌سازی، نیاز به حافظهٔ کمتر برای ذخیره‌سازی و برقراری ارتباط آسان‌تر و سریع‌تر، روش‌های زیادی برای نهان‌نگاری در آن معرفی شده‌اند [11]؛ علاوه‌براین موارد، متن هنوز هم یکی از عمده‌ترین اشکال برقراری ارتباط در سراسر جهان است که در دسترس عموم کاربران اینترنت است.

در این مقاله، یک روش نوین برای نهان‌نگاری اطلاعات در بستر رایانامه و مبتنی بر لغت‌نامه ارائه می‌شود. در روش پیشنهادی، با استفاده از متن ارسالی، تعدادی نشانی[2] رایانامه که حاوی پیام مخفی هستند، تولید خواهد شد. متن ارسالی هم‌زمان به نشانی‌های تولیدشده و نشانی گیرنده ارسال می‌شود. با این کار هویت فرد گیرندهٔ پیام آشکار نشده و تنها گیرنده با استفاده از سایر نشانی‌های رایانامه می‌تواند اطلاعات مخفی را استخراج کند. در روش پیشنهادی، از دو کلید نهان‌نگاری استفاده می‌شود. استفاده از این دو کلید منجر به افزایش سطح امنیت روش پیشنهادی می‌شود. همچنین ظرفیت روش پیشنهادی نامحدود است که البته مزیت بزرگی در یک روش نهان‌نگاری محسوب می‌شود. این ظرفیت نامحدود، امنیت بالایی برای روش پیشنهادی ایجاد می‌کند. مزیت دیگر، محدودنبودن روش پیشنهادی به نوع متن پوشانه است. در ابتدا پیام مخفی به‌وسیلهٔ یک لغت‌نامه به یک رشته‌بیت تبدیل، سپس عملیات جاسازی پیام مخفی در نشانی‌های گیرنده، به‌وسیلهٔ کلیدهای نهان‌نگاری انجام می‌شود.

کارایی الگوریتم‌های نهان‌نگاری به عوامل مختلفی مانند عدم تشخیص توسط چشم انسان، عدم تشخیص توسط روش‌های آماری و ظرفیت وابسته است. روش پیشنهادی تغییری در متن پوشانه ایجاد نمی‌کند؛ ازاین‌رو، این روش توسط انسان یا روش‌های آماری قابل تشخیص نیست. ظرفیت روش پیشنهادی در این پژوهش، مبتنی بر نشانی‌های رایانامه‌های ساخته‌شده است. با افزایش متن پیام، تعداد نشانی‌های رایانامه ساخته‌شده افزون می‌شوند. که البته این افزایش نشانی رایانامه‌های ساخته‌شده می‌تواند منجر به مشکوک‌شدن رایانامه ارسالی شود؛ ازاین‌رو، پارامتر تعداد نشانی رایانامه‌های ساخته‌شده نیز در ارزیابی حائز اهمیت است. در این مقاله، کارایی روش پیشنهادی مبتنی بر دو پارامتر یادشده و در مقایسه با روش‌های موجود مورد ارزیابی قرار گرفته است. نتایج این ارزیابی نشان می‌دهد که روش پیشنهادی علاوه‌بر ارائه ظرفیت نامحدود در نهان‌نگاری، تعداد نشانی‌های رایانامه تولیدشده کمتر و همچنین تعداد بیت‌های پیام کمتری بعد از فشرده‌سازی تولید می‌کند.

در ادامه این مقاله، ابتدا در بخش دوم مروری کلی بر روش‌های نهان‌نگاری در متن انجام می‌شود؛ سپس در بخش سوم، جزییات روش پیشنهادی برای نهان‌نگاری در رایانامه ارائه می‌شود. برای درک بهتر روند نهان‌نگاری پیشنهادی، جزییات محاسبات در این روش با یاد مثالی در بخش چهارم ارائه خواهد شد. در بخش پنجم، نتایج آزمایش‌ها برای ارزیابی روش پیشنهادی شرح داده می‌شوند. در بخش ششم، نتیجه‌گیری مقاله ارائه می‌شود.

۲- مرور کارهای پیشین

به‌طورکلی می‌توان روش‌های نهان‌نگاری را در متن به سه دسته متفاوت زبانی، ساختاری و تصادفی-آماری تقسیم کرد [1]. درهمین‌اواخر مقالات [1,2,4] مرور کلی بر روش‌های موجود در نهان‌نگاری متن ارائه کرده‌اند. در ادامه این بخشی تعدادی از مهم‌ترین روش‌های ارائه‌شده برای نهان‌نگاری متن مرور خواهند شد.

نهان‌نگاری زبانی تغییر ویژگی‌های نحوی یا معنایی متن موجود است. تغییر ویژگی‌های نحوی می‌تواند قراردادن علایم نگارشی در مکان‌های مناسب باشد [3]. در

[1] Cover
[2] Address

[DOI: 10.52547/jsdp.19.1.59]

تغییر ویژگی‌های معنایی از جانشین‌کردن کلمات هم‌معنی برای پنهان‌کردن اطلاعات مخفی استفاده می‌شود [4]. گاهی جایگزینی یک کلمه هم معنی، نتایجی غیرقابل انتظار در سطح متن به‌وجود می‌آورد؛ ازاین‌رو، یافتن تغییردهندهٔ مناسب با قابلیت اعتماد بالا، مسأله‌ای چالش‌برانگیز در نهان‌نگاری زبانی است. چَنگ و کلارک در [5,29]، با استفاده از تجزیه‌گر گرامر ترکیبی دسته‌ای[1] یک روش نهان‌نگاری زبانی در متن ارائه کرده است. بانیک و همکاران در [3] یک روش نهان‌نگاری متن مبتنی بر پردازش زبان طبیعی ارائه کرده‌اند.

نهان‌نگاری ساختاری به معنای تغییر ساختار فیزیکی متن است. برای مثال، پنهان‌کردن اطلاعات مخفی در متن با تغییر فونت فضای خالی بین کلمات، در بستر واژه‌پرداز [12]، تغییر فونت و رنگ نویسه‌های موجود در سلول رسانه صفحه گسترده [13] و از جابه‌جایی کلمات و خطوط [14] می‌توان استفاده کرد. در [15]، هر نویسه انگلیسی و نقطه به یک رشتهٔ هفت نویسه‌ای نسبت داده شده است. هر نویسه پیام مخفی می‌تواند در یکی از هفت نویسه رشتهٔ نگاشت‌شده، با تغییر فضای بین نویسه، مخفی شود. در این شکل از نهان‌نگاری در متن، با انتقال متن از یک بستر به بستر دیگر متنی، اطلاعات مخفی‌شده از بین می‌رود.

هدف از نهان‌نگاری تصادفی-آماری، تولید متن پوشانه به‌صورت خودکار با استفاده از ویژگی‌های آماری پیام مخفی مورد نظر است [11]. حَمدان و حمارشه [16] از ساختار شبکه اومگا و کلمات فرهنگ لغت برای پنهان و استخراج‌کردن پیام مخفی استفاده شده است. از مشکل‌های این روش، عدم تولید جملات با معنی و همچنین پایین‌آمدن کارایی کل سامانه با استفاده مکرر از فایل‌های فرهنگ لغت برای جاسازی هر نویسه است.

علاوه‌بر دسته‌بندی‌های یاد شده، روش‌های دیگری نیز ارائه شده‌اند. ناگاراهالی در [20]، از پیام‌های کوتاه برای متن پوشانه استفاده و اطلاعات را در شکلک‌ها مخفی کرده است. نقطه‌ضعف این روش، ظرفیت پایین آن است. گرَی در [21] از سندهای اینترنتی[2] به‌عنوان متن پوشانه استفاده کرده و اطلاعات را در صفت‌ها[3] مخفی کرده است. از ویژگی‌های بارز این روش، یک‌پارچگی بین رمزنگاری و نهان‌نگاری است. این یک‌پارچگی موجب افزایش امنیت شده است. ماجومدر و چانگدر در [22] برای نهان‌نگاری اطلاعات از خلاصه‌سازی متن استفاده شده است. به‌دلیل عدم انسجام خلاصه تولیدشده، این روش امنیت پایینی دارد. پُر و همکاران در [23] یک روش نهان‌نگاری اطلاعات و یک راه حل برای حمله به این روش ارائه کرده است. این مرجع بر روی بستر واژه‌پرداز[4] و براساس انواع فضاهای خالی موجود در متن، نهان‌نگاری را انجام می‌دهد. طالبی و همکاران در [5] یک روش نهان‌نگاری در فایل‌های HTML بر بستر وب ارائه کرده‌اند. ریزو و همکاران در [6] یک روش نهان‌نگاری برای مخفی‌سازی متن در شبکه‌های اجتماعی ارائه کرده‌اند. رحمان و همکاران در [7] روشی برای مخفی‌سازی اطلاعات تشخیص هویت درون دنباله DNA ارائه کرده‌اند.

یکی از روش‌های پرکاربرد نهان‌نگاری در متن، نهان‌نگاری در رایانامه است. در این روش، از بستر رایانامه برای ارسال پیام مخفی استفاده می‌شود. ایدهٔ اصلی نهان‌نگاری در مقالات [8,10,24,25,26] ارسال رایانامه به تعداد زیادی مخاطب است. این ارسال، بدون توجه به معتبربودن نشانی‌های رایانامه انجام می‌شود. مراجع [8,10,24,25]، از نشانی‌های رایانامه برای ساخت کلید نهان‌نگاری استفاده می‌کنند. احمد و همکاران در [27] برای بهبود نهان‌نگاری، اقداماتی نظیر تولید تصادفی نشانی‌های رایانامه انجام شده است. با تولید نشانی‌های رایانامه تصادفی، سطح امنیت روش، افزایش یافته است. ساتیک و ایسیک در [8] براساس فشرده‌سازی LZW، یک روش نهان‌نگاری در رایانامه معرفی شده است. این مرجع، از مجموعه‌ای شامل ۴۸ متن، برای نهان‌نگاری استفاده می‌کند. مهم‌ترین مشکل این روش، پایین‌بودن ظرفیت نهان‌نگاری است. به عبارت دیگر، این روش محدود به متن ارسالی است، در این روش، از چند کلید نهان‌نگاری استفاده می‌شود، ازاین‌رو، این روش امنیت بالایی دارد. کومار و همکاران در [24] از ترکیب الگوریتم‌های فشرده‌سازی LZW[5]، MTF[6] و BWT[7] برای افزایش ظرفیت نهان‌نگاری استفاده شده است. در [26] با استفاده از فشرده‌سازی هافمن ظرفیت نهان‌نگاری نسبت به روش [24]، بهبود یافته است. در این روش، ظرفیت نهان‌نگاری در رایانامه، از 7.03، به 7.21 افزایش یافته است. مالیک و همکاران در [25] روشی متفاوت برای نهان‌نگاری در رایانامه ارائه داده است. در این روش، با استفاده از رنگ‌کردن نویسه‌های متن پوشانه، پیام مخفی شده است. با توجه به تعدد رنگ برای هر نویسه، این روش ظرفیت

[1] Combinatory Categorial Grammar (CCG) Parser
[2] HTML
[3] Attribute
[4] Microsoft Word Document
[5] Lempel–Ziv–Welch
[6] Move to Front
[7] Burrows Wheeler Transform

[DOI: 10.52547/jsdp.19.1.59]

بالایی دارد؛ اما به‌دلیل تغییر متن ارسالی امنیت کمتری نسبت به دیگر روش‌ها دارد.

فاتح و رضوانی در [28] یک روش نهان‌نگاری در رایانامه با استفاده از تکرار نویسه‌ها ارائه شده است. این روش از ظرفیت بالاتری در مقایسه با روش‌های مشابه برخوردار است. درواقع، ظرفیت این روش نامحدود است. در این روش از فشرده‌سازی LZW برای کاهش حجم پیام استفاده شده است. همچنین در این روش هیچ تغییری روی متن پوشانه ایجاد نشده است. در این روش با افزایش نشانی‌های رایانامه، ظرفیت نهان‌نگاری افزایش می‌یابد؛ اما افزایش نشانی‌های رایانامه، شک‌برانگیز است. در این روش محدودیتی روی پیام وجود ندارد. درواقع با افزایش طول پیام تعداد نشانی رایانامه‌های ارسالی افزایش می‌یابد؛ ازاین‌رو در این مقاله، روشی برای کاهش حجم پیام و کاهش تعداد نشانی‌های رایانامه گیرنده پیشنهاد شده است تا مشکل روش مرجع [28] تا حدی مرتفع شود. روش پیشنهادی ما بر مبنای این مرجع طرح‌ریزی شده است. تفاوت عمده روش پیشنهادی در مقایسه با روش ارائه‌شده در [28]، فشرده‌سازی به کمک لغت‌نامه است. به‌کارگیری لغت‌نامه علاوه‌بر کاهش حجم پیام، عملیات رمزکردن پیام را نیز انجام می‌دهد.

در روش نهان‌نگاری پیشنهادی در این مقاله، برخلاف بسیاری از روش‌های مشهور نهان‌نگاری در متن، هیچ نوع محدودیتی روی متن ارسالی وجود ندارد. همچنین هیچ نوفه یا اطلاعات اضافی به متن ارسالی اضافه نمی‌شود. به عبارت دیگر، هیچ تغییری در متن پوشانه اعمال نمی‌شود.

۳- روش پیشنهادی

به‌طور کلی می‌توان گفت که هر شمای نهان‌نگاری از دو مرحله اصلی جاسازی و استخراج پیام محرمانه تشکیل می‌شود. در مرحله جاسازی که در سمت فرستنده پیام اجرا می‌شود، پیام محرمانه در پوشانه قرار داده می‌شود. در مقابل و در مرحله استخراج که در سمت گیرنده پیام اجرا می‌شود، پیام محرمانه از رایانامه دریافتی استخراج خواهد شد. در ادامه این بخش ابتدا مرور کلی بر روش پیشنهادی خواهیم داشت؛ سپس فرضیات مورد نیاز در شمای نهان‌نگاری پیشنهادی را بیان می‌کنیم و درنهایت مراحل جاسازی و استخراج را با جزییات شرح خواهیم داد.

۳-۱- مرور کلی

ایدهٔ اصلی مرحله جاسازی، فشرده‌سازی پیغام مخفی و سپس نگاشت پیغام فشرده‌شده به تعداد نویسه‌ها است. در ابتدا پیغام توسط یک لغت‌نامه فشرده شده و به یک رشته‌بیتی تبدیل می‌شود. با توجه به مکان هر کلمه در لغت‌نامه، یک کد به آن کلمه اختصاص می‌یابد و متن پیغام مخفی به رشته‌ای از بیت‌ها تبدیل می‌شود. این رشته‌بیتی توسط یک شاخص به بخش‌های کوچکی تقسیم می‌شود؛ سپس معادل ده‌دهی هر بخش محاسبه می‌شود؛ در ادامه نگاشتی از هر عدد ده‌دهی به تعداد نویسه‌ها در متن پوشانه انجام می‌شود. این متن پوشانه همان بدنه رایانامه است که برای گیرنده ارسال می‌شود. در نگاشت هر بخش از پیام مخفی، تعداد نویسه‌های متن پوشانه و تعداد نویسه‌های هر جمله مبنای نهان‌نگاری هستند. هر عدد ده‌دهی تعداد نویسه‌ها از ابتدای متن پوشانه را نشان می‌دهند. با شمارش تعداد نویسه‌ها از ابتدای متن پوشانه تا نویسه نهایی عدد ده‌دهی مورد نظر دست می‌آید. رشته‌بیتی حاصل از این اعداد ده‌دهی به همراه یک جدول کلید (مشترک بین دو طرف)، برای مشخص‌کردن پسوند نشانی رایانامه استفاده می‌شود؛ درنهایت تعدادی نشانی رایانامه حاوی اطلاعات مخفی تولید می‌شوند. بدنه رایانامه (متن پوشانه) هم‌زمان به تمام نشانی‌های رایانامه و نشانی رایانامه گیرنده ارسال می‌شود. گیرنده با داشتن مقادیر کلید مشترک و همچنین دریافت رایانامه، مقدار پیام مخفی را از فهرست نشانی‌های رایانامه موجود در بخش گیرندهٔ رایانامه دریافتی استخراج می‌کند.

۳-۲- مفاهیم اولیه و فرضیات

در این بخش مفاهیم و علائم مورد استفاده در شمای نهان‌نگاری پیشنهادی تعریف می‌شود. همان‌طور که گفته شد، شمای پیشنهادی مبتنی بر یک لغت‌نامه است. این لغت‌نامه شامل ۶۹۸۸۸ کلمه پرتکرار در زبان انگلیسی است که در آن کلمات بر حسب میزان تکرار آن‌ها در ادبیات انگلیسی مرتب شده‌اند. به کلمات با تکرار بیشتر یک اندیس کمتر و به کلمات با تکرار کمتر یک اندیس بیشتر در لغت‌نامه اختصاص داده شده است. با توجه به تکرر هر کلمه در ادبیات انگلیسی، این لغت‌نامه کلمات را در ۳ بخش سازماندهی می‌کند. گفتنی است که هر کلمه در ادبیات انگلیسی در یک مکان از این لغت‌نامه با یک اندیس منحصربه‌فرد ذخیره می‌شود. بخش نخست

[DOI: 10.52547/jsdp.19.1.59]

لغت‌نامه شامل ۲۵۶ کلمه پرتکرار در زبان انگلیسی (با اندیس‌هایی در بازه [0-255])، بخش دوم شامل ۴۰۹۶ کلمه (با اندیس‌هایی در بازه [256-4351]) و بخش آخر شامل ۶۵۵۳۶ کلمه (با اندیس‌هایی در بازه [4352-69887]) است.

در روش پیشنهادی فرض می‌کنیم که برای هر دو طرف فرستنده و گیرنده پیام، علاوه‌بر لغت‌نامه مقادیر مشترک دیگری نیز وجود دارد که با کمک آن‌ها عملیات جاسازی و استخراج انجام می‌شود. این مقادیر به‌صورت زیر است:

- $\hat{x}$: بیشینه تعداد دفعاتی که بدنه رایانامه می‌تواند خوانده شود.
- $\hat{y}$: بیشینه تعداد جملاتی از بدنه رایانامه که می‌تواند در محاسبات قرار بگیرد.
- $\hat{z}$: بیشینه تعداد نویسه‌های بدنه رایانامه که می‌تواند در محاسبات در نظر گرفته شود.
- کلید A: شامل مجموعه‌ای هشت تایی از پسوند نشانی رایانامه‌ها است که در مرحله جاسازی و برای ساخت نشانی‌های رایانامه استفاده می‌شود. به هر پسوند نشانی رایانامه، یک عدد سه بیتی نسبت داده شده است. این مجموعه در جدول (۱) نشان داده شده است.

(جدول-۱): مجموعه مقادیر پسوند نشانی رایانامه‌ها (کلید A).
(Table-1): Prefix values used in Email addresses (Key A).

ردیف	پسوند نشانی رایانامه	کد ۳ بیتی
۱	gmail.com	**000**
۲	hotmail.com	**001**
۳	yahoo.com	**010**
۴	rediffmail.com	**011**
۵	btinternet.com	**100**
۶	aol.com	**101**
۷	msn.com	**110**
۸	verizon.net	**111**

- مجموعه S: شامل جملات موجود در بدنه رایانامه است و به‌صورت $S = \{s_1, s_2, \dots, s_{\hat{y}}\}$ نشان داده می‌شود.

۳-۳- جاسازی پیام

مرحله جاسازی پیام در سمت فرستنده پیام اجرا می‌شود و هدف از این مرحله جاسازی پیام مخفی در پوشانه است. بدیهی است که فرستنده پیام مخفی و همچنین مقادیر مشترک ارائه‌شده در زیربخش قبل را به‌عنوان ورودی دارد. در پایان این مرحله رایانامه ارسالی به‌نحوی تولید می‌شود که پیام مخفی در نشانی‌های رایانامه موجود در بخش گیرنده آن رایانامه جاسازی شده است.

مراحل جاسازی پیام مخفی در روش پیشنهادی شامل دوازده مرحله‌ی زیر است.

مرحله نخست: در ابتدا پیام مخفی با استفاده از لغت‌نامه به رشته‌ای از اعداد تبدیل می‌شود. برای این منظور هر لغت از پیام را در لغت‌نامه جستجو می‌کنیم. با توجه به نتیجه این جستجو، اندیس لغت در لغت‌نامه به‌دست‌آمده که در اینجا با j نمایش داده می‌شود. با توجه به اندیس لغت، عددی دودویی به آن نسبت داده می‌شود که این عدد از رابطهٔ زیر محاسبه شده و با p نمایش داده می‌شود:

$$p = \begin{cases} j-1 & if\ j < 256 \\ j-256-1 & if\ 256 \le j < 256+4096 \\ j-256-4096-1 & if\ j \ge 256+4096 \end{cases} \quad (۱)$$

در رابطه بالا، p می‌تواند یک عدد ۸، ۱۲ یا ۱۶ بیتی باشد. در صورت عدم وجود یک کلمه در لغت‌نامه، عدد دودویی مختص به آن کلمه از طریق اتصال کد اسکی تمامی نویسه‌های آن کلمه تولید می‌شود. بدیهی است که طول عدد دودویی در این شرایط وابسته به تعداد نویسه‌های کلمه یادشده است؛ لذا برای مقادیر p، چهار حالت متنوع متصور است. این حالت‌ها شامل مقادیر هشت بیتی، دوازده بیتی، شانزده بیتی و همچنین تعداد بیت‌های وابسته به کد اسکی هر نویسه هستند. پس در واقع، برای هر کلمه چهار حالت متصور است؛ ازاین‌رو، برای تفکیک این حالات نیاز به سرآیند دو بیتی برای هر کلمه است. این سرآیند چهار حالت مختلف دارد. این چهار حالت در جدول (۲) شان داده شده است.

(جدول-۲): چهار حالت مختلف سرآیند ابتدایی.
(Table-2): Four different values for the initial header.

مقدار هدر	حالت تعریف شده
00	عدد ۸ بیتی
01	عدد ۱۲ بیتی
10	عدد ۱۶ بیتی
11	عدد معادل کد اسکی هر نویسه

برای لغات موجود در لغت‌نامه بر اساس سرآیند می‌توان تعداد بیت متعلق به کلمه را مشخص کرد. برای مثال برای سرآیند "۱۰"، ۱۶ بیت برای لغت لحاظ می‌شود؛ اما برای لغات ناموجود در لغت‌نامه، تعداد بیت متعلق به کلمه بر اساس سرآیند قابل تشخیص نیست. در

[DOI: 10.52547/jsdp.19.1.59]

واقع، در این حالت تنها سرآیند اعلام می‌کند که کد اسکی هر نویسه باید جایگزین شود؛ اما اتمام این جایگزینی را اعلام نمی‌کند. برای رفع این مشکل یک سرآیند چهار بیتی، به سرآیند قبل اضافه می‌شود. این سرآیند تعداد نویسه‌های کلمه را مشخص می‌کند. پس برای مقدار سرآیند بین "0001" تا "1111"، تعداد نویسه‌های کلمه بین ۱ تا ۱۵ نویسه است. حالت "0000" برای کلمات با نویسه‌های بیشتر لحاظ شده است. در این حالت چهار بیت بعد نیز به‌عنوان سرآیند لحاظ می‌شود. پس برای مقدار سرآیند بین "00000001" تا "00001111"، تعداد نویسه‌های کلمه بین شانزده تا سی نویسه است. حالت "00000000" برای کلمات با نویسه‌های بیشتر لحاظ شده است. این روند برای تعیین تعداد نویسه‌های کلمه ادامه می‌یابد.

در این مرحله، فاصله خالی (White Space) کد نمی‌شود. پس از اتمام هر کلمه یک فاصله خالی مخفی وجود دارد که در زمان استخراج پیام، لحاظ می‌شود.

مرحله دوم: تعداد نویسه‌های بدنه رایانامه، شمارش و به‌عنوان پارامتر N_c لحاظ می‌شود.

مرحله سوم: در این مرحله یک پنجره از بیت‌های پیام تولید می‌شود. بیت‌ها از ابتدای رشته پیام مخفی خوانده‌شده، به پنجره بیتی اضافه شده و هم‌زمان عدد ده‌دهی معادل بیت‌های موجود در پنجره محاسبه می‌شود. این فرآیند تا جایی دنبال می‌شود که عدد ده‌دهی حاصل از بیت‌های موجود در پنجره از مقدار $\hat{x} \times N_c$، بزرگ‌تر شود. عدد حاصل d نام‌گذاری می‌شود. گفتنی است که با توجه به ساخت قدم‌به‌قدم این عدد ده‌دهی، بیت سمت چپ در رشته‌بیتی پیام مخفی، کم‌ارزش‌ترین بیت و بیت سمت راست در آن، پرارزش‌ترین بیت در عدد ده‌دهی d است. جزییات این فرآیند برای تولید عدد ده‌دهی d در الگوریتم (۱) نشان داده شده است.

مرحله چهارم: همان‌طور که گفته شد، هر بیت از پیام مخفی به یک نویسه از بدنه رایانامه نگاشت می‌شود. به‌عبارت دیگر تعداد نویسه‌های مورد نیاز برای مخفی‌سازی عدد d با توجه به طول بدنه رایانامه باید استخراج شود. بدیهی است که عدد ده‌دهی d با مقدار $\hat{x}$ قابل نمایش نیست؛ ازاین‌رو برای محاسبهٔ عدد ده‌دهی d، شاخص موقعیت کنونی بر روی رشته‌بیتی پیام مخفی، یک واحد به عقب برمی‌گردد. در صورت صفربودن بیت کنونی، تا رسیدن به بیت یک، به عقب برگشت داده می‌شود. در ادامه، تعداد بیت‌های صفر پیمایش شده شمارش و در N_z ذخیره می‌شود و تعداد بیت‌های باقی مانده مقدار d جدید را تولید می‌کند. از این مرحله به بعد با مقدار جدید d عملیات دنبال می‌شود. دلیل ذخیره صفرها، بی‌نیازکردن الگوریتم به نگهداری تعداد بیت‌های پیمایش شده است. اگر تعداد این صفرها ذخیره نشوند، باید تعداد بیت‌های پیمایش‌شده ذخیره شوند و در صورت عدم ذخیرهٔ این اطلاعات، امکان بازیابی پیام میسر نیست.

(الگوریتم-۱): انتخاب عدد ده‌دهی مورد نیاز از رشته‌بیتی پیام مخفی.

(Algorithm-1): Choosing a decimal value from the bitstream of the secrect text.

```
1.  SecretBlockSelection(M, x̂, N_c)
2.    d ← 0
3.    while d < x̂ × N_c do
4.      // get the next bit in M
5.      b ← the least significant bit in M
6.      // shift d one bit to the left
7.      d ← d ≪ 1
8.      if (b == 1) then
9.        // use bitwise OR for adding b to
          d
10.       d ← d or b
11.     end if
12.   end while
13.   return d
14. end SecretBlockSelection
```

در اعداد دودویی، بیت سمت راست، کم‌ارزش‌ترین بیت است؛ اما در مرحلهٔ سوم، بیت سمت راست، پرارزش‌ترین بیت لحاظ شده بود. دلیل این تصمیم‌گیری، جلوگیری از صفرشدن پر ارزش‌ترین بیت است. در صورت صفرشدن پرارزش‌ترین بیت، امکان بازیابی آن وجود ندارد. اگر بیت سمت چپ، پرارزش‌ترین بیت باشد، امکان صفربودن آن وجود دارد. در صورت صفربودن این بیت، با توجه به الگوریتم پیشنهادی، امکان بازیابی آن وجود ندارد و پیام درست استخراج نخواهد شد.

مرحله پنجم: در این مرحله مقادیر x، y و z محاسبه می‌شوند. برای این منظور ابتدا مقادیر x و m با کمک روابط (۲) و (۳) محاسبه می‌شوند:

$$x = \frac{d}{N_c} \qquad (۲)$$

$$m = \frac{d}{N_c} \bmod 10 \qquad (۳)$$

[DOI: 10.52547/jsdp.19.1.59]

گفتنی است که مقدار m برابر با باقیمانده تقسیم d بر N_c به پیمانه ۱۰ است. از ابتدای متن، خواندن جمله‌به‌جمله انجام می‌شود. طول جملات خوانده‌شده، تا بزرگ‌ترشدن از مقدار m، با هم جمع می‌شوند. مقدار y برابر با تعداد جملات منهای یک است؛

$$\sum_{i=1}^{y}|s_i| \le m \le \sum_{i=1}^{y+1}|s_i| \qquad (۴)$$

مقدار z از اختلاف بین مقدار m و مجموع تعداد نویسه‌های خوانده‌شده در y جمله حاصل می‌شود؛ لذا به‌وسیلهٔ رابطه (۵) محاسبه می‌شود:

$$z = m - \sum_{i=1}^{y}|s_i| \qquad (۵)$$

مرحله ششم: مقدار $\hat{z}$، بیشینه تعداد نویسه است. ممکن است، مقدار z از مقدار $\hat{z}$، بیشتر باشد. در این حالت نیاز است که از مفهومی به‌عنوان دسته (category) برای رفع این مشکل استفاده کنیم. با استفاده از این مفهوم، امکان جاسازی مقادیر z بزرگتر از $\hat{z}$ فراهم می‌شود. برای نمایش هر دسته از یک مقدار و یک نماد (Symbol) استفاده می‌شود. مقدار هر دسته که با category نمایش داده می‌شود، از طریق رابطه (۶) محاسبه، سپس این مقدار با کمک مقادیر موجود در جدول (۳) به یک نماد نگاشت می‌شود. نمادهای ارائه‌شده در این جدول به‌طور متداول در نشانی‌های رایانامه استفاده می‌شوند. این نمادها، شامل نقطه (.)، خط فاصله (-) و زیرخط (_) هستند. در صورت عدم وجود هیچ‌کدام از نمادهای جدول (۳) در رایانامه مربوطه، مقدار $category$ برابر با صفر است. اکنون می‌توانیم به کمک این نمادها مقدار z را تا چهار برابر مقدار $\hat{z}$ لحاظ نماییم. مقدار نهایی z با استفاده از رابطهٔ (۷) محاسبه می‌شود:

$$category = \frac{z}{\hat{z}+1} \qquad (۶)$$

$$z = \frac{z}{\hat{z}+1} \bmod 10 \qquad (۷)$$

(جدول-۳): نمادهای در نظر گرفته شده برای هر دسته.
(Table-3): Corresponding symbol of each category.

نماد (Symbol)	دسته (Category)
.	1
_	2
-	3

مرحله هفتم: رشته‌بیت برای مقادیر x، y و z با توجه به بیشینه تعداد بیت‌های درنظرگرفته‌شده برای آن‌ها، ساخته می‌شوند. این رشته‌ها به‌ترتیب کنار هم قرار گرفته و برای مرحله بعد ارسال می‌شود.

مرحله هشتم: سه بیت انتهایی رشته‌بیت تولیدشده در مرحلهٔ قبل برای مشخص‌کردن پسوند نشانی رایانامه جدا شده و با استفاده از کلید A پسوند مربوطه استخراج می‌شود. لازم به یادآوری است، همان‌طور که در بخش ۳-۲ گفته شد، در کلید A برای هر پسوند یک کد سه بیتی منحصربه فرد تعریف شده است که کمک می‌کند در این مرحله با کمک کد سه بیتی حاصله به یک پسوند مشخص دست یافت.

مرحله نهم: بیت‌های باقی‌مانده به دسته‌های چهاربیتی تقسیم شده و به هر دسته به ترتیب یک اندیس داده می‌شود. مقادیر این اندیس‌ها از صفر شروع شده و به‌ترتیب اضافه می‌شود. بنابراین برای دسته‌ای با اندیس i مقدار ده‌دهی حاصل از چهار بیت آن دسته را با k_i نشان می‌دهیم. با این تعریف و با کمک رابطه زیر برای هر دسته نظیر دسته i، یک نویسه از نویسه‌های انگلیسی استخراج می‌گردد.

$$C_i = (i \times 16 + k_i) \bmod 26 \qquad (۸)$$

که مقدار C_i یک عدد بین صفر تا ۲۵ بوده و نمایان گر یکی از نویسه‌های انگلیسی برای دسته iام است. برای نمونه برای مقدار صفر نویسه 'a' و برای مقدار ۲۵ نویسه 'z' نگاشت می‌شود.

مرحله دهم: با استفاده از رابطه (۸) نویسه مرتبط به مقدار N_z استخراج شده و یک نویسه پیش از پسوند نشانی رایانامه ساخته می‌شود. در این بخش، i برابر با صفر و k_i برابر N_z آن نشانی رایانامه است.

مرحله یازدهم: با نویسه‌های تولیدشده و پسوند نشانی رایانامه، نشانی رایانامه معنادار ساخته می‌شود.

مرحله دوازدهم: برای جاسازی بیت‌های باقیمانده پیام مخفی، مراحل سوم تا یازدهم به‌ترتیب تکرار می‌شوند. بدیهی است که در هر دور از مراحل سوم تا یازدهم، تعدادی از بیت‌های پیام مخفی (که در پنجره بیتی قرار داده می‌شوند) جاسازی می‌شود. این مراحل تا زمانی که تمامی بیت‌های پیام مخفی جاسازی نشده‌اند، تکرار خواهند شد.

با انجام مراحل بالا، تمام نشانی رایانامه‌های حاوی اطلاعات مخفی، تولید می‌شوند. بدنه رایانامه (متن پوشانه)

[DOI: 10.52547/jsdp.19.1.59]

هم‌زمان به تمام نشانی‌های رایانامه و نشانی رایانامه گیرنده ارسال می‌شود. بسیاری از رایانامه‌های تبلیغاتی به افراد متعددی ارسال می‌شوند؛ بنابراین تعداد زیاد نشانی‌های گیرنده در یک رایانامه، موجب مشکوک‌شدن گیرنده نمی‌شود. باید چند نمونه نشانی رایانامه در دنیای واقعی موجود باشد؛ زیرا در غیر اینصورت موجب مشکوک شدن شبکه امنیتی می‌شود.

۴-۳- استخراج پیام

بعد از دریافت رایانامه توسط گیرنده، مرحله استخراج پیام اجرا می‌شود. گفتنی است که گیرنده مقادیر مشترک ارائه‌شده در زیربخش ۳-۲ را در اختیار دارد. همچنین فرض می‌کنیم که رایانامه ارسالی توسط گیرنده به‌طور کامل و صحیح دریافت شده است؛ لذا هدف از این مرحله، استخراج پیام مخفی از رایانامه دریافتی و با کمک مقادیر مشترک است. برای این منظور گیرنده باید مراحل زیر را برای هر نشانی رایانامه موجود در فهرست نشانی‌های گیرنده در رایانامه دریافتی اجرا کند.

مرحله نخست: ابتدا مقادیر x، y و z از نشانی رایانامه مورد نظر استخراج می‌شوند. با توجه به قرارداد بین گیرنده و فرستنده، تعداد نویسه‌های حاوی پیام از نشانی رایانامه استخراج می‌شوند. برای نمونه گیرنده و فرستنده می‌توانند برای جاسازی و استخراج این تعداد نویسه در سه نویسه ابتدایی نشانی رایانامه توافق کنند. هر نویسه iام از نشانی رایانامه که حاوی پیام است (که با C_i نشان داده می‌شود)، با توجه به اندیس آن نویسه در نشانی رایانامه (مقدار i) و کد اسکی آن نویسه، از طریق رابطهٔ زیر به عدد ده‌دهی k_i تبدیل می‌شود:

$$k_i = (C_i - 16i) \bmod 26 \qquad (۹)$$

اعداد ده‌دهی تولیدشده به‌صورت رشته‌بیت درآمده و در کنار هم قرار می‌گیرند.

مرحله دوم: پسوندهای نشانی رایانامه از طریق کلید A به رشته‌های سه‌بیتی تبدیل و به انتهای رشته تولیدشده در مرحله قبل متصل می‌شوند.

مرحله سوم: از رشته‌بیت تولیدشده، مقادیر x، y و z استخراج می‌شوند. مقدار z می‌تواند بیشتر از مقدار $\hat{z}$ باشد؛ ازاین‌رو، با استفاده از نمادهای به‌کاررفته در نشانی رایانامه، تعداد دسته‌های $\hat{z}$ محاسبه و با مقدار z تولیدشده، جمع می‌شود.

مرحله چهارم: نویسه موجود پیش از پسوند نشانی رایانامه، با استفاده از رابطهٔ (۹) به عدد ده‌دهی N_z تبدیل می‌شود.

مرحله پنجم: تعداد نویسه‌های x تکرار متن، y جمله ابتدایی و z نویسه باقی‌مانده محاسبه و با هم جمع می‌شوند؛ سپس این مقدار به عدد دودویی تبدیل می‌شوند. به تعداد N_z به ادامهٔ رشته‌بیت صفر اضافه و رشته‌بیت جدید ساخته می‌شود.

مرحله ششم: مرحله‌های یک تا پنج، برای کلیه نشانی‌های رایانامه تکرار می‌شوند و رشته‌های بیتی جدید در کنار یکدیگر قرار می‌گیرند.

مرحله هفتم: رشته تولیدشده را با استفاده از لغت‌نامه از حالت فشرده خارج و پیام استخراج می‌شود.

۵-۳- تحلیل امنیتی

همان‌گونه که در مقدمه بیان شد، عدم تشخیص توسط چشم انسان، عدم تشخیص توسط روش‌های آماری و ظرفیت نهان‌نگاری از مهم‌ترین پارامترهای امنیتی روش‌های نهان‌نگاری است که در روش پیشنهادی تا حد زیادی ملاحظه شده است. ظرفیت روش پیشنهادی نامحدود است؛ ولی با افزایش طول پیام، تعداد نشانی رایانامه‌های ساخته‌شده افزایش می‌یابد. در روش پیشنهادی مهم‌ترین پارامتر امنیتی، تعداد نشانی‌های رایانامه است. با کاهش این تعداد، سطح امنیت نهان‌نگاری افزون خواهد شد.

در روش پیشنهادی تعداد نشانی‌های رایانامه ساخته‌شده وابسته به تعداد بیت‌های پیام پس از فشرده‌سازی با لغت‌نامه، $\hat{x}$، $\hat{y}$ و $\hat{z}$ است. با افزایش بیت‌های پیام، تعداد نشانی‌های رایانامه ساخته‌شده افزایش می‌یابد. همچنین با افزایش $\hat{x}$، $\hat{y}$ و $\hat{z}$ تعداد نشانی‌های رایانامه ساخته‌شده کاهش می‌یابد. افزایش این پارامترها، وابسته به تعداد نویسه‌های اضافه‌شده به هر نشانی رایانامه است. اگر تعداد نویسه‌های اضافه‌شده به هر نشانی رایانامه بیش از چهار نویسه باشد، آن نشانی را غیر متعارف می‌کند. اگر این تعداد نویسه‌ها برابر با سه لحاظ شود، نشانی رایانامه متعارف و متداول خواهد بود. با لحاظ‌کردن این سه نویسه و پسوند هر رایانامه، پانزده بیت برای این سه پارامتر می‌توان در نظر گرفت. با کاهش دو پارامتر $\hat{y}$ و $\hat{z}$ پارامتر $\hat{x}$ قابل افزایش است. وجود جملات کوتاه و با نویسه‌های اندک در متن پوشانه، پارامتر $\hat{z}$ را کاهش می‌دهد و امکان افزایش پارامتر $\hat{x}$ را فراهم می‌کند؛ اما از طرفی تعداد نویسه‌های متن پوشانه کاهش می‌یابد که این

[DOI: 10.52547/jsdp.19.1.59]

مسأله مطلوب نیست. کاهش جملات نیز پارامتر $\hat{y}$ را کاهش می‌دهد و امکان افزایش پارامتر $\hat{x}$ را فراهم می‌کند؛ اما از طرفی تعداد نویسه‌های متن پوشانه کاهش می‌یابد که این مسأله مطلوب نیست.

اگر مقادیر $\hat{x}$، $\hat{y}$ و $\hat{z}$ با هم برابر و پنج بیتی باشند، مقدار $\hat{x} \times N_c$ بیشینه مقدار خود را خواهد داشت و تعداد نشانی رایانامه‌های ساخته‌شده کمینه می‌شود. پس متن پوشانه با ۳۱ جمله و بیشینه ۳۱ نویسه در هر جمله مناسب‌ترین متن پوشانه برای نهان‌نگاری با روش پیشنهادی است.

۴- مثالی از جزییات روش پیشنهادی

در این بخش، برای درک بهتر روش پیشنهادی، مثالی برای نمایش جزییات هر دو مرحله جاسازی و استخراج آورده شده است. از"Behind using" و شکل (۶) به‌ترتیب به‌عنوان پیام مخفی و متن پوشانه استفاده شده است. $\hat{x}$، $\hat{y}$ و $\hat{z}$ به‌ترتیب برابر ۱۲۷، ۳ و ۶۳ هستند. بیشینه نویسه‌های یک جمله در متن پوشانه دویست است. پس بیشینه مقدار z می‌تواند دویست باشد. همان‌گونه که در مرحله ششم بخش نهان‌نگاری شرح داده شد، با استفاده از مقدار Category می‌توان مقدار z را تا چهار برابر بیشتر از $\hat{z}$ در نظر گرفت. پس، کمینه مقدار $\hat{z}$ باید برابر پنجاه باشد و از طرفی باید این مقدار معادل $2^n - 1$ باشد. ازاین‌رو، $\hat{z} = 63$ لحاظ می‌شود. تعداد جملات متن چهار است؛ اما برای اختصاص چهار جمله نیاز به سه بیت است. با انتخاب $\hat{y} = 3$ تنها دو بیت نیاز است و برای استفاده از نویسه‌های جمله آخر می‌توان از پارامتر z استفاده کرد. مجموع $\hat{x}$، $\hat{y}$ و $\hat{z}$ تنها ۱۵ بیت است و ازاین‌رو تنها هفت بیت برای تعداد جملات باقی می‌ماند که ناگزیر مقدار $\hat{x}$ برابر با ۱۲۷ خواهد بود. پیام در سه نویسه ابتدایی و یک نویسه انتهایی نشانی‌های رایانامه تعبیه می‌شود.

مرحله نخست: رشته‌بیتی پیام مخفی با استفاده از لغت‌نامه به صورت زیر مشخص می‌شود:

"10110001010100100001110010110000"

مرحله دوم: مقدار N_c که نشان‌دهنده تعداد نویسه‌های کل متن پوشانه است، برابر ۶۰۱ است.

مرحله سوم: بیت‌ها از ابتدای رشته‌پیام خوانده و هم‌زمان عدد ده‌دهی آن محاسبه می‌شود. این فرآیند تاجایی دنبال می‌شود که عدد حاصل از ۷۶۳۲۷ (۱۲۷×۶۰۱) بزرگ‌تر شود. رشته‌بیت در این حالت برابر است با:

"10110001010100100001"

البته همان‌گونه که بیان شد، بیت سمت راست پرارزش‌ترین بیت است؛ ازاین‌رو رشته‌بیت بالا معادل با رشته‌بیت پایین است:

"10000100101010001101"

مرحله چهارم: مقدار عدد ده‌دهی محاسبه شده با مقدار $\hat{x}$ قابل نمایش نیست؛ ازاین‌رو شاخص موقعیت کنونی یک بیت به عقب برمی‌گردد. در ادامه مقدار چهار در N_z ذخیره می‌شود. این مقدار برابر با تعداد صفرها تا رسیدن به نخستین یک است. رشته‌بیت حاصل برابر با "100101010001101" است. عدد ده‌دهی معادل این رشته ۱۹۰۸۵ است.

مرحله پنجم: با توجه به رابطه‌های (۲ تا ۵) مقادیر $x = 31$، $y = 2$ و $z = 140$ به‌دست می‌آید.

محاسبه این پارامترها به شرح زیر است:

$$19085 - 18631 = 454 < 601$$

که این مقدار به معنای ۳۱ بار خواندن کل متن است. طول جمله نخست دویست نویسه، جمله دوم ۱۱۴، جمله سوم ۱۴۵ و جمله چهارم ۱۴۳ نویسه است. 454<145+114+200 در این حالت مقدار $y = 2$ است. 140=454-314 نویسه‌های باقی‌مانده هستند و در این حالت مقدار $z = 140$ می‌شود.

مرحله ششم: مقدار z از مقدار $\hat{z}$ بیشتر است. به همین دلیل، با توجه به رابطهٔ (۶) مقدار $category = 2$ است؛ بنابراین نماد مربوطه برای نمایش این مقدار از جدول (۳) استخراج می‌شود. این نماد "_" است. همچنین مقدار z جدید با توجه به رابطهٔ (۷) محاسبه می‌شود که برابر با مقدار دوازده است.

مرحله هفتم: مقادیر دودویی x، y و z محاسبه و به‌ترتیب در کنار یکدیگر قرار می‌گیرند. این رشته‌بیت معادل شکل (۱) است.

$$\underbrace{0011111}_{x}\underbrace{10}_{y}\underbrace{001100}_{z}$$

(شکل-۱): رشته‌بیتی حاصل از xyz

(Figure-1): The bitstream obtained from x, y, and z values.

مرحله هشتم: پسوند "@btinternet.com" از کلید A با توجه به سه بیت انتهایی یعنی "100" استخراج می‌شود.

مرحله نهم: رشته‌بیت باقی‌مانده به دسته‌های چهارتایی "0001"، "1111" و "0011" تقسیم می‌شود. شاخص i در دستهٔ نخست، صفر است، ازاین‌رو با استفاده از فرمول (۸)، از دستهٔ نخست نویسه b استخراج می‌شود. شاخص i برای دسته دوم برابر یک می‌شود و نویسه f

[DOI: 10.52547/jsdp.19.1.59]

استخراج می‌شود. برای دسته آخر با مقدار $i = 2$ نویسه j استخراج می‌شوند.

مرحله دهم: مقدار N_z برابر چهار است؛ ازاین‌رو، برای نویسه انتهایی نشانی رایانامه مطابق فرمول (۸)، e استخراج می‌شود.

مرحله یازدهم: با نویسه‌ها، نماد و پسوند نشانی رایانامه، نشانی رایانامهٔ معنادار ساخته می‌شود.

مراحل سوم تا یازدهم برای تمام بخش‌ها تکرار می‌شود. رایانامه ارسالی تولیدشده در این مثال در شکل (۲) آمده است.

همان‌گونه که اشاره شد، تعداد نشانی‌های رایانامه ساخته‌شده وابسته به تعداد بیت‌های پیام پس از فشرده‌سازی با لغت‌نامه، تعداد نویسه‌های متن پوشانه و $\hat{x}$ است. در این مثال، مقدار $\hat{x} \times N_c$ برابر با ۷۶۳۲۷ می‌شود و قابلیت پنهان‌سازی متوسط ۱۶٫۵ بیت در یک نشانی رایانامه را می‌دهد. تعداد بیت‌های پیام پس از فشرده‌سازی به کمک لغت‌نامه برابر با ۳۲ بیت است. به‌طور متوسط برای پنهان‌سازی هر ۱۶٫۵ بیت نیاز به یک نشانی رایانامه است، ازاین‌رو به‌طورتقریبی به دو نشانی رایانامه برای نهان‌نگاری نیاز است.

اگر متن پوشانه دارای ۳۱ جمله با بیشینه $= ۱۲۴$ ۴×۳۱ نویسه باشد، با احتساب فاصله‌ها به‌طورتقریبی متن پوشانه شامل ۵۵۰۰ نویسه است. برای این متن میزان $\hat{x} = 31$ است. پس مقدار $\hat{x} \times N_c$ برابر با ۱۷۰۵۰۰ می‌شود و به‌طور متوسط برای پنهان‌سازی هر هجده بیت نیاز به یک نشانی رایانامه است.

پس اگر مقادیر $\hat{x}$، $\hat{y}$ و $\hat{z}$ با هم برابر باشند، مقدار $\hat{x} \times N_c$ بیشینه مقدار خود را خواهد داشت و تعداد نشانی رایانامه‌های ساخته‌شده کمینه می‌شود.

برای مثالی از استخراج پیام از رایانامه نخست استفاده و بخش نخست رشته‌بیت پیام تولید می‌شود.

مرحله نخست: به‌دلیل آن که دسته‌های بیت چهارتایی در سه نویسه نخست پنهان شده‌اند، در ابتدا باید رشته‌بیت مربوط به آن‌ها را استخراج کرد. رشته‌بیت مربوط به هر نویسه با استفاده از رابطه (۹) حاصل می‌شود. در جدول (۴) نتایج آورده شده است.

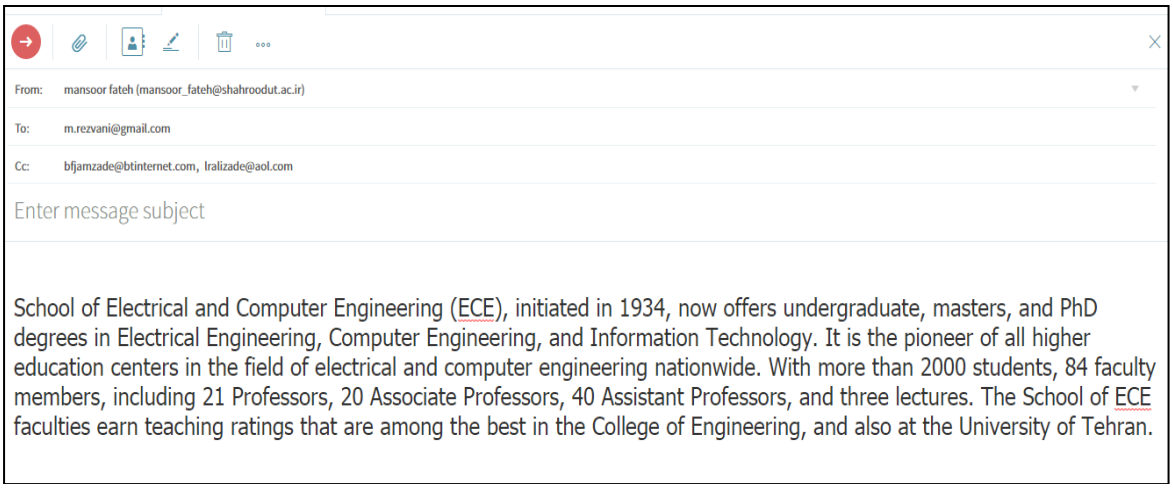
مرحله دوم: با استفاده از کلید A با توجه به پسوند "@btinternet.com"، رشته‌بیت "100" استخراج می‌شود. این رشته‌بیت به انتهای رشته‌بیت حاصل در مرحلهٔ نخست متصل می‌شود.

(جدول-۴): تولید رشته‌بیتی از نویسه‌های نخست در رایانامه.
(Table-4):: The bitstream generated from first three characters in the email.

C_i	i	k_i	رشتهٔ بیتی
b	۰	1	0001
f	1	15	1111
j	2	3	0011

مرحله سوم: مقادیر x، y و z از رشته تولیدشده، استخراج می‌شوند (x="0011111", y="10", z="001100"). با استفاده از جدول (۳) تعداد دسته‌های $\hat{z}$ برابر با یک و مقدار نهایی z، 140 است.

مرحله چهارم: مقدار N_z با استفاده از فرمول (۹) از نویسه آخر استخراج می‌شوند. این مقدار برابر چهار است.



(شکل-۲): رایانامه ارسالی توسط روش نهان‌نگاری پیشنهادی برای یک مثال.
(Figure-2): An example of the email generated using the proposed method.

مرحله پنجم: رشته‌های x، y و z به عدد ده‌دهی تبدیل می‌شوند. تعداد نویسه‌های x دور متن، y جمله ابتدایی و z نویسه باقی‌مانده محاسبه و با هم جمع می‌شوند. اکنون این مقدار به عدد دودویی تبدیل می‌شود.

[DOI: 10.52547/jsdp.19.1.59]

با توجه به N_z، چهار صفر پشت رشته‌بیت گذاشته می‌شود. این مقدار برابر با "0000101010001101" است؛ سپس این رشته‌بیت معکوس و درنهایت برای نشانی رایانامه ابتدایی رشته‌بیت "1011000101010000" ساخته می‌شود.

مرحله ششم: مراحل یک تا پنج برای دیگر نشانی‌های رایانامه به‌ترتیب تکرار می‌شوند و رشته‌های بیت جدید کنار هم قرار می‌گیرند.

مرحله هفتم: رشته‌بیت با استفاده از لغت‌نامه از حالت فشرده خارج و پیام استخراج می‌شود.

۵- نتایج ارزیابی

در این بخش نتایج ارزیابی روش پیشنهادی ارائه می‌شود. برای این منظور ابتدا محیط ارزیابی و سپس نتایج ارزیابی معیارهای مهم برای روش پیشنهادی شرح داده می‌شود.

۵-۱- محیط ارزیابی

کارایی روش نهان‌نگاری پیشنهادی با کمک دو پارامتر ظرفیت و تعداد نشانی‌های رایانامه تولیدشده مورد ارزیابی قرار می‌گیرد. برای این منظور ظرفیت شمای نهان‌نگاری پیشنهادی را با روش پیشنهادی با مراجع [1]، [3]، [6]، [20] و [28] مقایسه می‌کنیم. شکل‌های (۴ و ۳) به‌ترتیب پیام مخفی و متن ارسالی مشترک استفاده‌شده در این مراجع را نشان می‌دهد؛ لذا برای ارزیابی ظرفیت در این مقاله از پیام مخفی و متن ارسالی یادشده استفاده خواهیم کرد.

ظرفیت یک روش نهان‌نگاری از طریق رابطه (۱۰) محاسبه می‌شود. در این رابطه، N_m تعداد بیت‌های پیام و N_c تعداد بیت‌های متن پوشانه است.

$$capacity = \frac{N_m}{N_c} \qquad (۱۰)$$

همان‌گونه که در مقدمه بیان شد، تعداد نشانی‌های رایانامه از دیگر پارامترهای امنیتی مهم در روش‌های نهان‌نگاری مبتنی بر نشانی رایانامه است. نهان‌نگاری با تعداد نشانی رایانامه کمتر از امنیت بالاتری برخوردار است؛ ازاین‌رو در ادامه تعداد نشانی‌های رایانامه ساخته‌شده در روش پیشنهادی با مرجع [28] مقایسه شده است. گفتنی است که روش پیشنهادی و مرجع [28] ظرفیت نامحدود برای نهان‌نگاری ارائه کرده‌اند.

۵-۲- ظرفیت نهان‌نگاری

کارایی الگوریتم‌های نهان‌نگاری به عوامل مختلفی مانند عدم تشخیص توسط چشم انسان، عدم تشخیص توسط روش‌های آماری و غیره وابسته است؛ علاوه‌بر این موارد، مهم‌ترین عامل کارایی الگوریتم نهان‌نگاری، ظرفیت است. ظرفیت روش پیشنهادی محدود نیست و می‌توان هر حجم از پیام را در متن پوشانه ذخیره کرد. درواقع با افزایش تعداد نشانی رایانامه‌های ارسالی می‌توان حجم بیشتری از پیام را مخفی کرد، ازاین‌رو، ظرفیت روش پیشنهادی با دیگر روش‌های نهان‌نگاری در رایانامه، مقایسه شده است. برای مقایسه روش‌ها، از یک متن پوشانه، پیغام مخفی و تعداد نشانی رایانامه مشترک استفاده شده است. شکل‌های (۴ و ۳) به‌ترتیب پیام مخفی و متن ارسالی مشترک استفاده‌شده در مراجع مربوط به نهان‌نگاری در رایانامه را نشان می‌دهد.

نتایج ارزیابی ظرفیت نهان‌نگاری در روش‌های یادشده برای نهان‌نگاری رایانامه با کمک پیام مخفی و متن ارسالی مشترک یادشده در جدول (۵) ارائه شده است. همان‌طور که در این جدول مشاهده می‌شود، روش پیشنهادی و روش مرجع [28] هر دو دارای ظرفیت نامحدود در نهان‌نگاری هستند. پس این روش‌ها از ظرفیت بالاتری در مقایسه با روش‌های مراجع [1]، [3]، [6] و [20] برخوردار هستند. علاوه‌بر ظرفیت بالای روش [28]، وابسته‌نبودن به متن پوشانه از دیگر ویژگی‌های مهم این روش است. روش پیشنهادی در این پژوهش نیز مستقل از متن پوشانه و ویژگی روش ارائه‌شده در [28] را حفظ کرده است.

> "in the research area of text steganography, algorithms based on font format have advantages of great capacity, good imperceptibility and wide application range. However, little work on steganalysis for such algorithms has been reported in the literature. based on the fact that the statistic features of font format will be changed after using font-format-based steganographic algorithms, we present a novel support vector machine-based steganalysis algorithm to detect whether hidden information exists or not. this algorithm can not only effectively detect the existence of hidden information, but also estimate the hidden information length according to variations of font attribute value. as shown by experimental results, the detection accuracy of our algorithm reaches as high as 99.3% when the hidden information length is at least 16 bits."

(شکل-۳): متن ارسالی (پوشانه) مورد استفاده در مرجع [18].
(Figure-3): Cover text used in [18].

[DOI: 10.52547/jsdp.19.1.59]

“behind using a cover text is to hide the presence of secret messages the presence of embedded messages in the resulting stego text cannot be easily discovered by anyone except the intended recipient”

(شکل-۴): پیام مخفی مورد استفاده در مرجع [18].

(Figure-4): Secret message used in [18].

(جدول-۵): ظرفیت روش‌های مختلف نهان‌نگاری در متن.

(Table-5): Capacity of different text-steganography approaches.

مرجع	ظرفیت
[1]	**7.017**
[3]	**6.92**
[6]	**7.03**
[20]	**7.21**
[28]	**نامحدود**
روش پیشنهادی	**نامحدود**

۳-۵- تعداد نشانی‌های رایانامه تولیدشده

در روش ارائه‌شده در [28]، با افزایش نشانی‌های رایانامه، ظرفیت نهان‌نگاری نیز افزایش می‌یابد. درواقع با افزایش نشانی‌های رایانامه، امکان مخفی‌سازی تعداد بیشتری از بیت‌های پیام وجود دارد؛ بنابراین ظرفیت نهان‌نگاری در این روش نامحدود است. برای مقایسه روش‌های نهان‌نگاری با ظرفیت نامحدود، نیاز به یک پارامتر ارزیابی دیگر است؛ لذا به‌ازای یک پیغام یکسان تعداد نشانی رایانامه‌های ساخته‌شده در این روش را استخراج می‌کنیم و روشی با تعداد نشانی رایانامه‌های کمتر روشی مناسب‌تر است.

در جدول (۶) تعداد نشانی رایانامه‌های ساخته‌شده به‌ازای پیغام شکل (۵) و پوشانه شکل (۶) برای روش پیشنهادی با مرجع [28] مقایسه شده است. همان‌طور که در این جدول دیده می‌شود، تعداد نشانی‌های تولیدشده توسط روش پیشنهادی کمتر از روش ارائه‌شده در مرجع [28] است.

تعداد نشانی‌های تولیدشده کمتر به‌دلیل استفاده از نوع فشرده‌سازی در روش پیشنهادی است که مبتنی بر ایده لغت‌نامه است. برای بررسی تأثیر نوع فشرده‌سازی پیام مخفی، تعداد بیت‌های تولیدی بعد از فشرده‌سازی پیام مخفی شکل (۵) را برای هر دو روش پیشنهادی و [28] ارزیابی کردیم. مقدار دودویی حاصل از فشرده‌سازی این پیام در شکل (۷) به‌ازای هر دو روش مختلف آورده شده است.

در جدول (۷) طول پیام پس از فشرده‌سازی، به‌ازای پیغام شکل (۵) برای هر دو روش پیشنهادی و مرجع [28] مقایسه شده است. همان‌گونه که مشاهده می‌شود، طول پیام در روش پیشنهادی برابر با ۱۰۴ بیت و طول پیام برای روش پیشنهادی در]28[برابر ۱۵۰ بیت است؛ لذا روش پیشنهادی حدود ۴۴٪ ($\frac{150-104}{104}\times 100 = 44.2\%$) کمتر از طول پیام در روش ارائه‌شده در [28] است. کاهش طول پیام، امکان نهان‌نگاری امن‌تر را فراهم می‌کند. با کاهش طول پیام، تعداد نشانی رایانامه‌های ساخته‌شده کاهش یافته و درنتیجه امنیت نهان‌نگاری افزایش می‌یابد؛ لذا می‌توان نتیجه گرفت که امنیت نهان‌نگاری در روش پیشنهادی بالاتر از روش ارائه‌شده در [28] است.

Shahrood University of technology

(شکل-۵): پیام مخفی برای ارزیابی اندازه خروجی الگوریتم فشرده‌سازی.

(Figure-5): Secret message for evaluating the length of the compression algorithm.

School of Electrical and Computer Engineering (ECE), initiated in 1934, now offers undergraduate, masters, and PhD degrees in Electrical Engineering, Computer Engineering, and Information Technology. It is the pioneer of all higher education centers in the field of electrical and computer engineering nationwide. With more than 2000 students, 84 faculty members, including 21 Professors, 20 Associate Professors, 40 Assistant Professors, and three lectures. The School of ECE faculties earn teaching ratings that are among the best in the College of Engineering, and also at the University of Tehran.

(شکل-۶): متن ارسالی (پوشانه).

(Figure-6): The cover text.

(جدول-۶): تعداد نشانی رایانامه‌های ساخته‌شده به‌ازای پیغام شکل (۵) و پوشانه شکل (۶).

(Table-6): Number of email addresses generated for the secret message and cover text shown in Figure 5 and Figure 6, respectively.

مرجع	تعداد نشانی‌های رایانامه
[28]	10
روش پیشنهادی	7

10010 00111 00000 00111 10001 01110 01110
00011 10100 01101 01000 10101 00100 10001
10010 01000 10011 11000 01110 00101 10011

[DOI: 10.52547/jsdp.19.1.59]

00100 00010 00111 01101 01110 01011 01110 00110 11000

(الف) دودویی پیام مخفی با روش LZW.

A) Binary form of the secret message compressed by LZW.

11 1000 1010011 1101000 1100001 1101000 1110010 1101111 1101111 1100100 01 111000110011 00 00011100 10 1111100001110000

(ب) دودویی پیام مخفی به‌کمک لغت‌نامه.

B) Binary form of the secret message compressed by Dictionary.

(شکل-۷): نتایج فشرده‌سازی متن مخفی شکل (۵) با کمک روش LZW و لغت‌نامه.

(Figure-7): Compression results of the LZW and Dictionary algorithms over the secret message shown in Figure 5.

(جدول-۷): تعداد بیت‌های پیام بعد از فشرده‌سازی برای روش‌های با ظرفیت نامحدود.

(Table-7): Number of bits after compressions for two methods providing unlimited capacity.

مرجع	تعداد بیت‌های پیام بعد از فشرده‌سازی
[28]	150
روش پیشنهادی	104

۶- نتیجه‌گیری

در این مقاله، روشی نوین بر پایه لغت‌نامه و تعداد نویسه‌های موجود در متن پوشانه، برای نهان‌نگاری در رایانامه ارائه شده است. روش پیشنهادی محدود به زبان و نوع متن پوشانه خاص نیست و با هر متن پوشانه‌ای می‌توان نهان‌نگاری را انجام داد. در روش پیشنهادی، ابتدا به کمک لغت‌نامه متن پیام مخفی فشرده و به رشته‌بیتی تبدیل می‌شود. رشته‌بیت تولیدی به بخش‌هایی شکسته می‌شود. با توجه به عدد دهدهی هر بخش و تعداد نویسه‌های موجود در متن پوشانه، نشانی‌های رایانامه تولید می‌شوند. متن پوشانه همزمان به گیرنده و تمام نشانی‌های رایانامه ارسال می‌شود. از آنجایی که متن پوشانه در این فرایند تغییر نمی‌کند، روش پیشنهادی برای چشم انسان قابل تشخیص نیست. همچنین به‌دلیل ارسال همزمان رایانامه به چندین نشانی تولیدشده (که یکی از آن‌ها گیرنده اصلی است)، گیرنده پیام مشخص نیست. مزیت دیگر روش پیشنهادی در مقایسه با دیگر روش‌ها، تولید تعداد نشانی رایانامه معقول و کمتر نسبت به سایر روش‌های مشابه است. ظرفیت روش پیشنهادی در مقایسه با دیگر روش‌های نهان‌نگاری در متن، بیشتر است و درواقع این روش یک ظرفیت نامحدود برای نهان‌نگاری ارائه می‌کند.

با توجه به نتایج امیدبخش در شمای نهان‌نگاری پیشنهادی، به‌عنوان کار آینده می‌توان با بهبود روش فشرده‌سازی، حجم پیام مخفی و درنهایت تعداد نشانی‌های رایانامه تولیدشده را کاهش داد؛ علاوه‌بر این با کمک ماژول‌های طراحی‌شده در روش پیشنهادی، می‌توان یک شمای نهان‌نگاری امن با ظرفیت بالا برای شبکه‌های اجتماعی پیشنهاد داد.

7- References ۷- مراجع

[1] M. Taleby Ahvanooey, Q. Li, J. Hou, AR. Rajput, C. Yini, "Modern Text Hiding, Text Steganalysis, and Applications: A Comparative Analysis," *Entropy*, 2019 Apr; 21(4):355.

[2] M. Taleby Ahvanooey, Q. Li, HJ. Shim, Y. Huang, "A comparative analysis of information hiding techniques for copyright protection of text documents," Security and Communication Networks, 2018.

[3] B. Gupta Banik, SK. Bandyopadhyay, "Novel Text Steganography Using Natural Language Processing and Part-of-Speech Tagging", *IETE Journal of Research*, vo. 13, pp. 1-2, 2018.

[4] NS. Kamaruddin, A. Kamsin, LY. Por, H. Rahman, "A Review of Text Watermarking: Theory, Methods, and Applications," *IEEE Access*, vol. 6:80, pp. 11-28, 2018.

[5] M. Taleby Ahvanooey, H. Dana Mazraeh, SH. Tabasi, "An innovative technique for web text watermarking (AITW)," *Information Security Journal: A Global Perspective*, 1;25(4-6):191-6. 2016.

[6] SG. Rizzo, F. Bertini, D. Montesi, C. Stomeo, "Text watermarking in social media," In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, 2017, vol. 31, pp. 208-211, ACM.

[7] MS. Rahman, I. Khalil, X. Yi, "A lossless DNA data hiding approach for data authenticity in mobile cloud based healthcare systems," *International Journal of Information Management*, vol. 1, no. 45, pp. 276-88, 2019.

[8] E. Satir and H. Isik, "A Huffman compression based text steganography method," *Multimedia tools and applications*, vol. 70, no. 3, pp. 2085-2110, 2014.

[9] C.C. Chang, "A reversible data hiding scheme using complementary embedding strategy," *Information Sciences*, vol. 180, no. 16, pp. 3045-3058, 2010.

[10] E. Satirand H. Isik. "A compression-based text steganography method," *Journal of Systems and Software*, vol. 85, no. 10, pp. 2385-2394, 2012.

[DOI: 10.52547/jsdp.19.1.59]

[11] S. Bhattacharyya, P. Indu, and G.Sanyal, "Hiding Data in Text using ASCII Mapping Technology (AMT)," *International Journal of Computer Applications*, vol. 70, no. 18, 2013.

[12] R. Kumar, A. Malik, S. Singh, B. Kumar, and S. Chand, "A space based reversible high capacity text steganography scheme using font type and style," *In International Conference on Computing, Communication and Automation (ICCCA)*, pp. 1090-1094, 2016.

[13] S.A. Al-Asadi and W.Bhaya, "Text Steganography in Excel Documents Using Color and Type of Fonts," *Research Journal of Applied Sciences*, vol. 11, no. 10, pp. 1054-1059, 2016.

[14] S. Roy and M.Manasmita, "A novel approach to format based text steganography," *In Proceedings of the 2011 International Conference on Communication, Computing & Security*, pp. 511-516, 2011.

[15] B.K. Ramakrishnan, P.K.Thandra, and A.V. Srinivasula, "Text steganography: a novel character-level embedding algorithm using font attribute," *Security and Communication Networks*, vol. 9, no. 18, pp. 6066-6079, 2016.

[16] A.M. Hamdan and A.Hamarsheh, "AH4S: an algorithm of text in text steganography using the structure of omega network," *Security and Communication Networks*, vol. 9, no. 18, pp.6004-6016, 2016.

[17] M. Shirali-Shahreza, "Text steganography by changing words spelling," In Advanced Communication Technology, 10th International Conference on, vol. 3, pp. 1912-1913, 2008.

[18] J. Gardiner, "StegChat: A Synonym-Substitution Based Algorithm for Text Steganography," PhD Thesis, School of Computer Science University of Birmingham, pp. 1-64, 2012.

[19] C.Y. Chang and S. Clark, "Linguistic steganography using automatically generated paraphrases," *In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 591-599, 2010.

[20] T.P. Nagarhalli, "A New Approach to SMS Text Steganography using Emoticons," *In International Journal of Computer Applications (0975–8887), National Conference on Role of Engineers in Nation Building (NCRENB-14)*, pp. 1-3, 2014.

[21] M. Garg, "A novel text steganography technique based on html documents," *International Journal of Advanced Science and Technology*, vol. 35, pp. 129-138, 2011.

[22] A. Majumder and S. Changder, "A novel approach for text steganography: Generating text summary using Reflection Symmetry," *Procedia Technology*, vol. 10, pp. 112-120, 2013.

[23] L.Y. Por, K. Wong, and K.O. Chee, "UniSpaCh: A text-based data hiding method using Unicode space characters," *Journal of Systems and Software*, vol. 85, no. 5, pp. 1075-1082, 2012.

[24] R. Kumar, S. Chand, and S. Singh, "An Email based high capacity text steganography scheme using combinatorial compression," *In Confluence The Next Generation Information Technology Summit (Confluence), 5th International Conference*, pp. 336-339, 2014.

[25] A. Malik, G. Sikka, and H.K. Verma, "A high capacity text steganography scheme based on LZW compression and color coding," *Engineering Science and Technology, an International Journal*, vol. 20, no. 1, pp.72-79, 2016.

[26] R. Kumar, A. Malik, S. Singh, and S. Chand, "A high capacity email based text steganography scheme using Huffman compression," *In Signal Processing and Integrated Networks (SPIN), 3rd International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 53-56, 2016.

[27] T. Ahmad, M.S. Marbun, H. Studiawan, W. Wibisono, and R.M.Ijtihadie, "A Novel Random Email-Based Steganography," *International Journal of e-Education, e-Business, e-Management and e-Learning*, vol. 4, no. 2, pp. 129-134, 2014.

[28] M. Fateh, M. Rezvani, "An email-based high capacity text steganography using repeating characters," *International Journal of Computers and Applications*, pp. 1-7, 2018.

[29] Chang CY, Clark S. "Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method," *Computational linguistics*, vol, 40, no. 2, pp. 403-48, 2014

محسن رضوانی مدرک دکترای خود را در حوزه امنیت شبکه و از دانشگاه UNSW استرالیا دریافت کرد و در حال حاضر دانشیار دانشگاه صنعتی شاهرود است.

نشانی رایانامه ایشان عبارت است از:

mrezvani@shahroodut.ac.ir.

[DOI: 10.52547/jsdp.19.1.59]

منصور فاتح مدرک دکترای خود را در حوزه هوش مصنوعی و از دانشگاه تربیت مدرس دریافت کرد و در حال حاضر دانشیار دانشگاه صنعتی شاهرود است.

نشانی رایانامه ایشان عبارت است از:

mansoor_fateh@shahroodut.ac.ir.

[DOI: 10.52547/jsdp.19.1.59]