Hiding Data in a QImage File

Gabriela Mogos *

Abstract—**The idea of embedding some information within a digital media, in such a way that the inserted data are intrinsically part of the media itself, has aroused a considerable interest in different fields. One of the more examined issues is the possibility of hiding the highest possible amount of information without affecting the visual quality of the host data. As opposed to the traditional hiding methods, which are using different mathematical methods in order to hide the information, steganography in quantum informatics implies an approach based on the laws of the quantum physics. This work has as a purpose to expand the field of applicability of the embedded information within a digital media from the classical informatics to the quantum one.**

Keywords: qubit, qutrit, entanglement, steganography

1 Introduction

Steganography involves hiding data in an overt message and doing it in such a way that it is difficult for an adversary to detect and difficult for an adversary to remove. Based on this goal, three core principles can be used to measure the effectiveness of a given steganography technique: amount of data, difficulty of detection, and difficulty of removal. Steganography has become a hot topic on the Internet in the context of electronic privacy and copyright protection. Just as with encryption, two people using steganography must agree on the algorithm they are going to use and exchange this algorithm prior to communicating. Steganography is a technology where modern data compression, information theory, spread spectrum, and cryptography technologies are brought together to satisfy the need for privacy on the Internet. Steganography methods themselves are rapidly evolving and becoming increasingly sophisticated. Steganography has become a hot topic on the Internet in the context of electronic privacy and copyright protection.

Amount of data suggests that the more data you can hide, the better the technique.

Difficulty of detection relates to how easy it is for somebody to detect that a message has been hidden. There is usually a direct relationship between how much data can be hidden and how easy it is for someone to detect it. As you increase the amount of information that is hidden in a file, you increase the chance that someone will be able to detect that there is information hidden in the file.

Difficulty of removal involves the principle that someone intercepting your file should not be able to remove the hidden data easily.

Algorithmic steganography often uses as support audio or video files, speculating the imperfections of the human sense organs. The files hiding information will continue to perform their mission: the image will be able to be seen normally, the music will be listened normally, the documents will be read normally. Somebody who does not know in advance that these files dissimulate supplementary information, will not realize it even after using these files. The advantage of steganography as compared to cryptography is that the hidden message does not draw attention towards itself. Taking into account the advantages of steganography in classical informatics, and starting from the premise that a classical color image can be seen as a register of "color" qutrits in quantum informatics, this work has as purpose to expand the field of applicability of the steganography from the classical informatics to the quantum one.

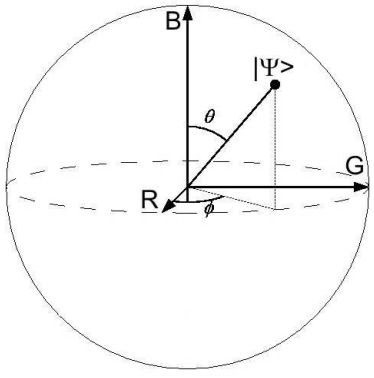

Figure 1: Qutrit's representation in terms of the R, G and B axes.

In the classical informatics, image is a row of pixels. Zizzi[1] proved that a qubit can be associated to a pixel. In quantum informatics, we can look at the image as at a registry of ternary systems which has 3 basis states ((R,G,B) for the color space) which can exist in superpositions of that three basis states. This 3-level system

*Al.I.Cuza University, Computer Science Department, General Berthelot no.16, 700483, Iasi, Romania, Email: gabi.mogos@gmail.com

is called a qutrit (figure 1). Consequently, the image is a registry of qutrits, their "colored" state being determined by the combination in different proportions of the colors Red, Green, Blue, or, in other words, by their linear combination. The state of such a qutrit is expressed:

$$|\Psi\rangle = \sin\theta\cos\phi|R\rangle + \sin\theta\sin\phi|G\rangle + \cos\theta|B\rangle \quad (1)$$

2 Work method

The procedure of hiding a message inside an image made of qutrits is based on the entanglement of the states of three qubits: the first one belongs to the image, the second one belongs to the secret key, and the third one belongs to the message itself. Andrzej Grudka and Antoni Wjcik [2] showed that from a qutrit can be reconstructed a qubit with a state randomly chosen. A qubit is represented in an orthogonal base $\{|i\rangle,|j\rangle\}$, thus:

$$|\Psi\rangle = \sin\varphi|i\rangle + \cos\varphi|j\rangle \quad (2)$$

The first step in the implementation of the procedure is to reconstruct a qubit with the generic state given by the equation (2) starting from a qutrit of the image. From a qutrit with the state (1) two qubits can be reconstructed, of one's own free choice, with the states:

$$|\Psi\rangle = \sin\phi|R\rangle + \cos\phi|G\rangle \quad (3)$$

respectively

$$|\Phi\rangle = \sin\theta|B\rangle + \cos\theta|K\rangle \quad (4)$$

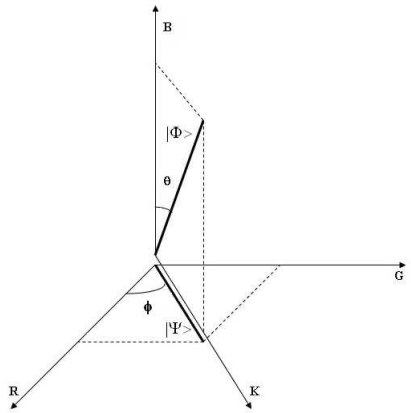

Figure 2: From a qutrit can be reconstructed two qubits: $|\Psi\rangle$ or $|\Phi\rangle$.

2.1 Obtaining the secret key

In 1984, Bennett and Brassard developed a protocol in order to obtain the secret key using qubits. The protocol is already known: the sender will send a qubit to the receiver (using a quantum channel), which will measure its state by projecting it randomly on one of the two possible bases. He will make public (using a classical channel) only the chosen base, and not the obtained result. If the choice was good, then the sender and the receiver will share the same qubit, thus they will eliminate it completely. The sender and the receiver will realize a number of iterations until they will obtain a key with a corresponding length.

The bases used at the measurement will be $\{|R\rangle,|G\rangle\}$ and $\{|B\rangle,|K\rangle\}$. After getting the final key, the bites of the key will be represented in the work basis used at the measurements in the protocol of key distribution.

2.2 Hiding the information

Hiding the information inside an image involves:

1. Getting the secret key;

2. Encrypting the message and hiding it inside the image.

The phenomenon of entanglement stays on the basis of the encrypting procedures and hiding the information. The entanglement is realized among the states of three qubits: one belongs to the message, one to the secret key, for the encrypting, and one is obtained from a qutrit of the image, for the hiding. In order to obtain the qubit of the image, we need to reconstruct it starting from a qutrit. We assume that the state of the qubit obtained is:

$$|\Psi\rangle_i = \sin\phi|R\rangle + \cos\phi|G\rangle \quad (5)$$

The qubit belonging to the secret key has the state::

$$|\Psi\rangle_k = \sin\gamma|R\rangle + \cos\gamma|G\rangle \quad (6)$$

The qubit belonging to the secret message has the state:

$$|\Psi\rangle_m = \sin\delta|R\rangle + \cos\delta|G\rangle \quad (7)$$

The maximally entanglement obtained has the state:

$$|\Psi\rangle_{mki} = \sin\phi\sin\gamma\sin\delta|RRR\rangle_{mki} + \\ + \cos\phi\cos\gamma\cos\delta|GGG\rangle_{mki} \quad (8)$$

We can also write the state of the entanglement as it follows:

$$|\Psi\rangle_{mki} = \sin\chi|\mathbf{R}\rangle + \cos\chi|\mathbf{G}\rangle \quad (9)$$

where:

$$\sin\chi = \sin\phi\sin\gamma\sin\delta; \quad \cos\chi = \cos\phi\cos\gamma\cos\delta \quad (10)$$

It is very important to consider the fact that we work with "colored" qutrits of the image, and the alteration of their shade as a result of the procedure of hiding the information can be perceived visually, which is not advisable at all.

Starting from the state of a qutrit (the color shade) which is represented by the equation (1) we wanted to determine

the maximum variations of the two angles θ respectively ϕ to which the modification of the color should not be visually perceived, as well as the influence produced by the entanglement of the states over these two angles. In the same time, we check if there are constraints imposed in the choice of the states of the qubits (the key and the message) so as the "resulted" entanglement should not determine variations of the angles θ respectively ϕ higher than the ones previously determined.

Starting from a randomly chosen color combination: $R = 0,231; G = 0,835; B = 0,125$ (the green color) the two angles: θ and ϕ, have been calculated.

$$\begin{aligned} R &= \cos\phi \sin\theta \\ G &= \sin\phi \sin\theta \\ B &= \cos\theta \end{aligned}$$

We obtained the values: $\theta = 82^o 13^{'}$ respectively $\phi = 74^o 31^{'}$.

In order to follow the maximum admitted variations of the angles ϕ and θ for which the alteration of the shade of the color is not perceived visually, each angle was modified individually, as long as the other angle was maintained constant. The two angles ϕ and θ characterize the states (3) respectively (4) of the qubits which are obtained from the qutrits of the image. Maintaining a constant angle ϕ, the maximum variation obtained for the angle θ is of $2^{'}$, and for the constant angle θ, we obtained a maximum variation of $1^{'}$ of the angle ϕ.

Our purpose is to realize if, in order to obtain the final state of the entanglement (9), the selection of the states of the qubits which will be entangled is randomly made, or not. In other words, we want to see if it is necessary to impose some rules in the selection of the angles γ and δ (states of the qubits of the key, respectively of the message) so as the angle χ ($\phi - 1^{'} \leq \chi \leq \phi$) of the state of the entanglement G.H.Z. - $|\Psi\rangle_{mki}$, to meet the condition:

$$\begin{aligned} \sin(\phi - 1^{'}) \leq \sin\chi \leq \sin\phi \\ \cos(\phi - 1^{'}) \leq \cos\chi \leq \cos\phi \end{aligned} \tag{11}$$

For the sinus function, we have:

$$\sin(\phi - 1^{'}) \leq \sin\phi \sin\gamma \sin\delta \leq \sin\phi \tag{12}$$

For $z = \sin\gamma \sin\delta$, we have:

$$\sin(\phi - 1^{'}) \leq \sin\phi \cdot z \leq \sin\phi$$

consequently:

$$0 \leq z \leq 1 \tag{13}$$

Replacing z, the equation (13) becomes:

$$0 \leq \sin\gamma \sin\delta \leq 1 \tag{14}$$

That is:

$$0 \leq \sin\delta \leq 1 \quad \texttt{for} \quad 0 \leq \sin\gamma \leq 1 \tag{15}$$

Taking into account the cosine function, the condition (15) is fulfilled: $\delta, \gamma \in [0, \frac{\pi}{2}]$.

We can notice that the entanglement of the two states (6) and (7) with the state (5) does not determine the variation of the angle ϕ so as the shade of the qubit to be altered.

The encrypting and the hiding of the message are represented by the entanglement of the states of the three qubits: of the key $|\Psi\rangle_k$ and of the message $|\Psi\rangle_m$ (for encrypting) and of the image $|\Psi\rangle_i$ (for hiding). A first step towards hiding the information is the dividing of the message and of the key in sets with equal number of qubits, so as to each qubit of the message corresponds a qubit of the key. Due to the fact that the message contains a high number of qubits from the image, hiding the information involves the use of a number of qutrits equal to the number of the qubits in the message. In order that the message is reconstructed correctly, the states given by the equation (9) must be displayed on the surface of the image under the form of an oriented graph (figure 3). The states represented by the equation (9) correspond to the nodes in the graph, and the link between two nodes is made either by an entanglement of states, or using a controlled-phase gate.

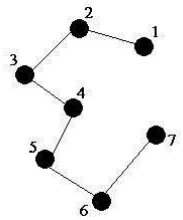

Figure 3: The states given by the entanglement of the states of the three qubits: of the key, of the message and of the image will be displayed on the surface of the image under the form of an oriented graph.

In our case, the link between the nodes is realized by the entanglement of the states of the qubits of the message which will be hidden inside the image (figure 4). Then each of these qubits (m), together with a qubit of the key (k), are entangled with a qubit from the image (i).

Taking into account the representation of the states of the three qubits, the state of a node is as it follows:

$$\begin{aligned} |\Psi\rangle_{mki} = \sin\phi \sin\gamma \sin\delta |R_m R_k R_i\rangle + \\ + \cos\phi \cos\gamma \cos\delta |G_m G_k G_i\rangle \end{aligned} \tag{16}$$

or:

$$\begin{aligned} |\Psi\rangle_{mki} = \sin\phi \sin\gamma \sin\delta |RRR\rangle_{mki} + \\ + \cos\phi \cos\gamma \cos\delta |GGG\rangle_{mki} \end{aligned} \tag{17}$$

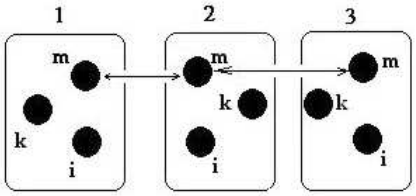

Figure 4: Inside of the node - the entanglement of three qubit's states.

2.3 Recovery of the hidden information

At the destination, after receiving the image, the information must be extracted. This means that the receiver must know first the position on the image where the message begins, and this can be found only talking with the sender.

Thus, the position on the image where the first node of the graph is found, and from where the hidden information actually starts is communicated to the receiver. The sender communicates the position of the first node using a classical communication channel. In order to extract every qubit used in the states entanglement from the image, the receiver should have a clean original image, without any hidden information. The receiver will extract, for the beginning, from the clean image (using the same base as the one from the protocol of distribution of the key and the one of the qubits of the message) the qubit $|\Psi\rangle_i$.

Extracting the first qubit of the message, in the same way of the other qubits belonging to the hidden information, is realized starting from the maximal entanglement from the nodes of the oriented graph, entanglement with the state as it follows:

$$|\Psi\rangle_{mki} = \alpha|RRR\rangle_{mki} + \beta|GGG\rangle_{mki} \qquad (18)$$

where: $\alpha = \sin\phi \sin\gamma \sin\delta$ and $\beta = \cos\phi \cos\gamma \cos\delta$.

The first qubit means in fact to obtain the state $|\Psi\rangle_m$. For this, the receiver will apply over the qubits of the image and of the key of the gate C-NOT and Hadamard in order to obtain the Bell bases:

$$|\Psi_\pm\rangle_{ik} = \frac{1}{\sqrt{2}}(|RR\rangle_{ik} \pm |GG\rangle_{ik})$$
$$|\Phi_\pm\rangle_{ik} = \frac{1}{\sqrt{2}}(|RG\rangle_{ik} \pm |GR\rangle_{ik})$$

The equation (18) becomes:

$$|\Psi\rangle_{mki} = \frac{1}{2}[|\Psi_+\rangle_{ik}(\sin\delta|R\rangle_m + \cos\delta|G\rangle_m) +$$
$$+|\Psi_-\rangle_{ik}(\sin\delta|R\rangle_m - \cos\delta|G\rangle_m) +$$
$$+|\Phi_+\rangle_{ik}(\cos\delta|R\rangle_m + \sin\delta|G\rangle_m) +$$
$$+|\Phi_-\rangle_{ik}(-\cos\delta|R\rangle_m + \sin\delta|G\rangle_m)]$$

The measurement of the state $|\Psi\rangle_{mki}$ using the bases $|\Psi_\pm\rangle_{ik}$ and $|\Phi_\pm\rangle_{ik}$ determines its collapsing in the state $|\Psi\rangle_m$, i.e. to obtain the state of the first qubit belonging to the hidden information. The important thing is that the receiver takes into account that the basis in which the reconstruction of the qubit of the image is made $|\Psi\rangle_i$, the basis which will be identical to the one used in the protocol of distribution of the key, as well as of the qubits of the message.

Due to the fact that the distribution of these maximal entanglement has the shape of an oriented graph, the receiver can determine the following positions of the nodes of the graph on the "map" of the image, the extraction of the qubits of the message being made in the same way with the extraction of a qubit, as it was presented above. The hidden information is reconstituted placing in order the qubits obtained from every node numbered from 1 to n (n=number of the nodes of the graph), its reading being also made from 1 to n.

3 Analysis

Steganalysis is the process of analyzing various media such as digital photos, video, audio and other file formats in order to find the existence of a secret message or watermark and respond appropriately to the find. The ethical nature of the "appropriate" response may be debated but we are concerned with the technology. A steganalysis "attack" represents the technique with which the steganalyst attempts to recover, modify or remove a stego message. In classical cryptography, there exist 5 steganalysis attacks which are incidentally derived from 4 cryptanalysis techniques: stego-only, known-cover, known message, chosen stego and chosen message. In the stego-only method the steganalyst only has available the stego medium or the finished stego product. This is by far the most difficult attack approach since there is no starting point from which to start extracting the hidden message. So typically the steganalyst will scan by steganalysis algorithm type first. The "known message attack" assumes either a part of or the entire hidden message is available to the steganalyst. An efficient approach is to begin in parallel an effort to decrypt the message and an effort to detect other hidden messages based on the signature of the known message. The "chosen stego attack" asserts the steganography algorithm and the cover data are known. In this case the key, if the message is encrypted, and the hidden message are unknown. "Chosen message attack" refers to the steganalyst's knowledge of the hidden message with the goal of effectively detecting stego messages. This attack assumes the hidden message is known but a community has no knowledge of which container is hiding it. In this effort the steganalyst will generate various stego messages using various stego algorithms in an attempt to find consistent patterns and improve detection of the hidden message. In the case of the quantum cryp-

tography, the steganalysis "attack" also implies the use by the steganalyst of some techniques by which a stego message can be recovered, modified or removed, but the ways to realize this are different from the classical case.

An eavesdropper can wiretap the public channel, but that won't do him any good. He gets information on the bases and not on the outcome of the measurement. In case the eavesdropper attempts to measure part of the Quantum Channel he betrays herself by a high Quantum Bit Error Rate (QBER) and the sender and the receiver are warned. The Quantum Bit Error Rate (QBER) is the ratio of an error rate to the key rate and contains information on the existence of an eavesdropper and how much he knows.

On one hand, in order to be able to interfere in any of the ways (recover or modify) over the hidden message, it is necessary to know the encrypting key. An eavesdropper can try to rebuild the encrypting key obtained as a result of the protocol Quantum Key Distribution (QKD). On the other hand, trying to remove the message from the digital image file involves a series of steps. Then, considering that the message suffered a series of "transformations" (the state entanglement between the qubits of the message and then again the entanglement between the message's qubits, the key's qubits and the qubits belonging to the image support) the qubits which compose the message will be difficult to extract and the message will be difficult to recover completely, without knowing the encrypting key, and, of course, the state of the support qutrit.

4 Conclusions

With steganography we can send messages without anyone having knowledge of the existence of the communication. Steganography can be a solution which makes it possible to send news and information without being censored and without the fear of the messages being intercepted and traced back to you. Multimedia provides excellent covers for hidden information. Selecting the proper combination of steganography tools and covers is key to successful information hiding.

References

[1] P.A. Zizzi, *Holographym quantum geometry and quantum information theory*, Entropy, 2000, pp. 39-69 .

[2] Andrezj Grudka, Antoni Wojcik, *How to encode the states of two non-entangled qubits in one qutrit*, quant-ph/0303168, 2003.

[3] Ashley Montanaro, *Quantum walk on directed graphs*, quant-ph/0504116, 2005.

[4] Ingemar Cox, Matthew Miller, Jeffrey Bloom, Jessica Fridrich, Ton Kalker, *Digital Watermarking and Steganography*, M.K. Series in Multimedia Information and Systems, 2007.

[5] Chun-Shien Lu, *Multimedia Security: Steganography and Digital Watermarking Techniques for Protection of Intellectual Property*, Idea Group Publishing, 2005.