Investigating the feasibility of a practical Trojan-horse attack on a commercial quantum key distribution system

Nitin Jain[1,2], **Elena Anisimova**[3,4], **Christoffer Wittmann**[1,2], **Christoph Marquardt**[1,2], **Vadim Makarov**[3,4] **and Gerd Leuchs**[1,2]

[1] *Max Planck Institute for the Science of Light, Erlangen, Germany*
[2] *Institut für Optik, Information und Photonik, University of Erlangen-Nürnberg, Germany*
[3] *Institute for Quantum Computing, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1 Canada*
[4] *Department of Electronics and Telecommunications, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway*

As of today, quantum key distribution (QKD) is the most promising and pervasive application of quantum information technology. It offers unconditional security based on the laws of quantum mechanics: an eavesdropper *Eve* introduces errors while listening to the key-exchange between two legitimate parties, *Alice* and *Bob*, which disclose her presence. However, if the theoretical model is not properly implemented or if it fails to provide a complete description of the implementation, loopholes may arise (such as from technological deficiencies or operational vulnerabilities), that allow Eve to successfully breach the security.

An optical component inside a QKD system may be probed from the quantum channel by sending in sufficiently-intense light and analyzing the back-reflected light. This forms the basis of a Trojan horse attack [1]. We experimentally review the feasibility of such an attack on Clavis2, a commercially available QKD system from ID Quantique [2]. The objective is to read Bob's phase modulator (PM) to acquire knowledge of his basis choice, as this information suffices for constructing the raw key in the Scarani-Acin-Ribordy-Gisin 2004 (SARG04) protocol [3].

The principal idea is to send in a bright coherent pulse at a suitable wavelength λ and appropriately-chosen time τ, such that its back-reflection would've traversed through Bob's PM when it was activated (with frequency $f_{\text{Clavis2}} = 5$ MHz). This back-reflection, essentially a weak coherent state $|\alpha(\lambda)\rangle$, carries an imprint of Bob's randomly-chosen phase of 0 or $\frac{\pi}{2}$. Eve's task is then to be able to distinguish between two weak coherent states with some angle $\theta(\lambda)$ between them. This can be accomplished by, e.g., homodyne detection. The prior information that Eve requires is: when to send in the pulse (τ), and how many photons on average ($|\alpha(\lambda)|$) to expect. These can be readily estimated by techniques such as optical time domain reflectometry (OTDR) [1, 4].

We first prepared OTDR maps of the Bob module at three different wavelengths: 806, 1310 and 1550 nm. In fig. 1(a), we present the reflection maps at two of them. We find that the highest back-reflection level that could be utilized for an attack is only around -60 dB, implying that Eve needs to send in a bright pulse to obtain *at least* a few photons in the back-reflected pulse (a higher $|\alpha|$ would reduce probability of discrimination error).

With such a chosen intensity of Eve's input light, we find that strong afterpulsing occurs in Bob's detectors (see fig. 1(b)). Since this would cause a high QBER that would stop the QKD exchange, we are currently exploring the long wavelength ($1600 - 2000$ nm) regime where we conjecture that a low detector sensitivity and/or high back-reflection level would mitigate the afterpulsing effects. The idler output of an optical parametric oscillator, or supercontinuum light serve as two possible light sources to perform such a broadband spectral characterization. We report on the first results obtained with these sources and the feasibility to craft and execute a successful attack.

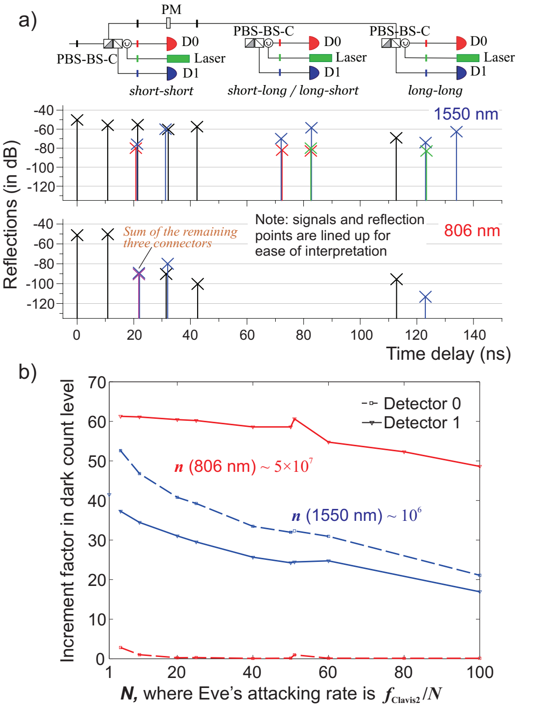

Figure 1: a) Reflection maps obtained using the OTDR technique at 806 nm and 1550 nm. b) Increase in dark count rate of Bob's detectors due to afterpulsing effects.

Several technical countermeasures such as watchdog detectors, optical isolators/filters, etc. have been proposed and need to be taken into account for security proofs.

References

[1] N. Gisin *et al.*, Phys. Rev. A **73**, 022320 (2006).

[2] Datasheet of Clavis2, available at ID Quantique website `http://www.idquantique.com`

[3] V. Scarani, A. Acin, G. Ribordy and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).

[4] A. Vakhitov, V. Makarov, and D. R. Hjelme, J. Mod. Opt. **48**, 2023 (2001).