Demonstrating feasibility of a Trojan-horse attack on a commercial QKD system

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

[1] Max Planck Institute for the Science of Light, Günther-Scharowsky-Str. 1, Bau 24, 91058 Erlangen, Germany

[2] Universität Erlangen-Nürnberg, Staudtstraße 7/B2, 91058, Erlangen, Germany

[3] Institute for Quantum Computing, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1 Canada

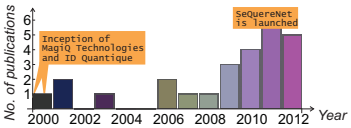
Quantum hacking: Motivation

- To investigate loopholes in practical QKD implementations that could be exploited by an eavesdropper Eve.

→ Imperfections & vulnerabilities mainly arise from technological deficiencies or operational weaknesses.

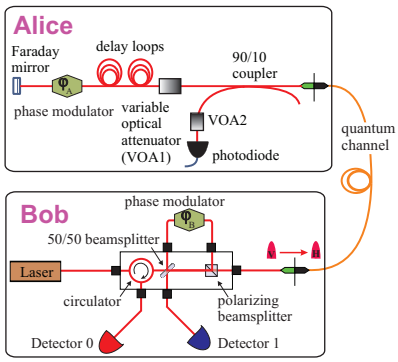
→ Looking for loopholes is a normal, iterative process during development of a secure communication system.

→ Responsible disclosure provides a positive feedback to the field. The eventual goal is to make practical implementations more secure!



P. Jouguet et al., Phys. Rev. A 86, 032309 (2012)
N. Jain, arXiv:1206.7019v1
G. Leuchs, ICQI, Ottawa (2011)
V. Makarov, CLEO/Europe-EQEC, Munich (2011)
V. Scarani and C. Kurtsiefer, arXiv:0906.4547v1

Clavis2 – Commercial QKD system from ID Quantique



Schematic of Alice and Bob in Clavis2

D. Stucki, N. Gisin, O. Guinnard, G. Ribordy, and H. Zbinden, New J. Phys. 4, 41 (2002)

C. H. Bennett and G. Brassard, Proc. IEEE Int. Conference on Computers, Systems, and Signal Processing, 175 (1984)
V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Phys. Rev. Lett. 92, 057901 (2004)

QKD protocols implemented

BB84 and ***SARG04*** protocols with highly attenuated laser pulses (typical mean photon number, $\mu < 1.0$).

Operating principle: Plug-n-Play

Asymmetric Mach-Zehnder interferometric configuration operating in double pass (both source and detector in Bob).

State preparation, basis application and detection

Alice modulates $\varphi_A = \{0, \pi/2, \pi, 3\pi/2\}$ to prepare the state. Bob modulates $\varphi_B = \{0, \pi/2\}$ to apply the basis choice and measures the outcome using avalanche photodiodes (APDs).

Exploitable vulnerability

If Eve can surreptitiously read Bob's modulation during the operation of ***SARG04***, then she obtains the raw key exchanged by Alice and Bob without their knowledge.

Trojan-horse attack: from ideas to implementation

Concepts

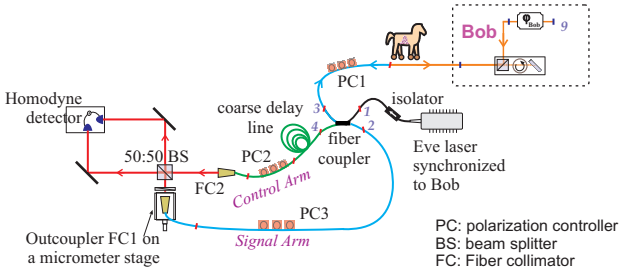
- Quantum channel is not just for Alice and Bob, it is also a potential open door for an eavesdropper.
- No optical component transmits or absorbs perfectly; some light is inevitably reflected back towards the input.
- Optical time/frequency domain reflectometry (OTDR/OFDR) can be used to generate maps of Alice and Bob from the quantum channel, allowing Eve to know when and with what intensity to attack.

N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, Phys. Rev. A 73, 022320 (2006)
A. Vakhitov, V. Makarov, and D. R. Hjelme, J. Mod. Opt. 48, 2023 (2001)

List of questions for Eve

- *When* exactly should the Trojan-horse pulse be launched?
- *How* to choose the suitable wavelength, brightness, and polarization for this pulse?
- *What* properties of the back-reflected pulse should be analyzed?
- *Which* components in the QKD system could be adversely affected by the attack?
- *What* percentage of the final secret key may be obtained from this attack?

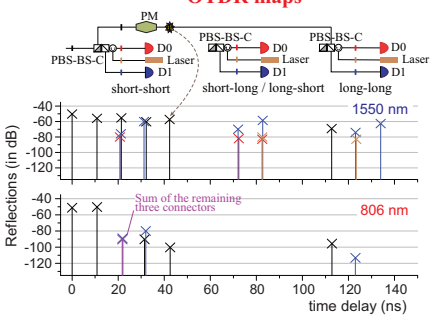
A Trojan-horse eavesdropper



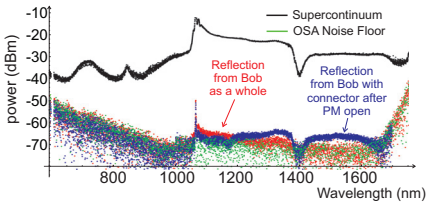
PC: polarization controller
BS: beam splitter
FC: Fiber collimator

Schematic of Eve's apparatus to perform a readout of Bob's phase modulation

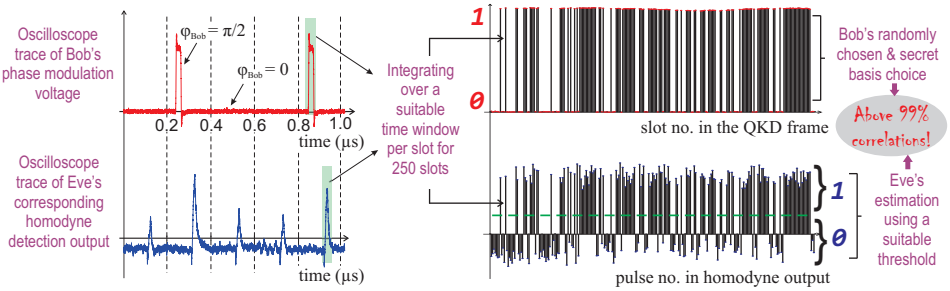
OTDR maps



Spectral characterization of reflectivity



Experimental results and issues for Eve

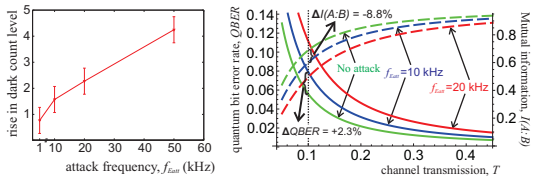


Results of phase readout; mean photon number, μ (Signal\Control arm) ≈ 100\10^8 and attack frequency, f_{Eatt} = 5 MHz

Eve's bane: Afterpulsing effects in Bob's APD based single-photon detectors

- → A back-reflection level of -60 dB, as per the OTDR map, implies Eve needs to send millions of photons (bright pulses) into Bob to get just a few photons back.
- → Bright pulses cause ***afterpulsing*** in the APD due to filling of traps by charge carriers. High attack frequency worsens the problem.
- → Afterpulsing results in substantial increase of dark counts. This raises the QBER and thus, may disclose Eve's presence.

C. Branciard, N. Gisin, B. Kraus, V. Scarani, Phys. Rev. A 72, 032301 (2005)



Increase in dark counts due to afterpulsing and corresponding changes in quantum bit error rate, QBER [solid lines] and mutual information of Alice and Bob, I(A:B) [dashed lines].

R. H. Haitz, J. Appl. Phys. 36, 3123 (1965); S. Cova, A. Lacaita, and G. Ripamonti, IEEE Electron. Dev. Lett. 12 685 (1991)

Future directions

Attack strategy and further improvements

- Afterpulsing is a blessing in disguise for security.
- We are devising an attack strategy in which Eve manipulates the QKD frame sent by Alice to Bob so that:
  - → the incurred QBER is below the abort-threshold of Clavis2,
  - → the privacy amplification step is insufficient in preventing Eve from knowing a partial amount of the final secret key, and
  - → the detection rate observed by Bob is within the bounds of the expected detection rate.
- Preliminary simulation results show this strategy works for low channel transmission values ($T \leq 0.2$) at least.
- Spectral sensitivity of the APD is four orders smaller at 1700 nm than at 1550 nm. We conjecture the afterpulsing might be significantly reduced in this wavelength regime.

Possible countermeasures

- ✓ Use interfaces such as FC/APC with reduced reflections.
- ✓ Reducing the duty cycle of the phase modulation voltage.
- ✓ Technical countermeasures, e.g. watchdog detector, are possible but need to be adapted to security proofs.