NeuroPass: A secure neural password based on EEG

Abhejit Rajagopal[1], Anthony C. Nguyen[2], Dennis M. Briggs
Department of Electrical Engineering, University of California, Los Angeles
[1] abhejit@ucla.edu , [2] anthony.c.nguyen@ucla.edu

***Abstract*—The NeuroPass development project is a novel application of brain-computer interfaces (BCI) in an everyday computing environment. Using the Emotiv EEG headset, we demonstrate how salient brain-wave features can be used to encode neural passwords. Because such EEG waveforms suffer from tremendous amounts of quantization and transient noise, a careful consideration of possible feature comparisons must be made to properly interpret such signals. Unlike most popular EEG tools, NeuroPass v1.0 does not require the use of any external software dependencies and has been fully automated to demonstrate feasibility and robust nature of such an encryption scheme in a machine-learning environment. The neural passwords demonstrated increase security by orders of magnitude, while decreasing the complexity of a user's input.**

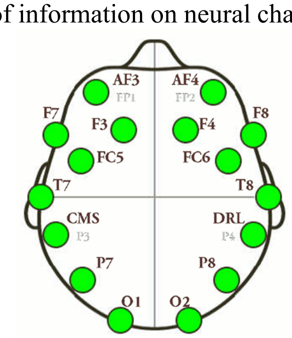
***Index Terms*— automatic recognition, biometrics, brain-computer interfaces, EEG, neural channels.**

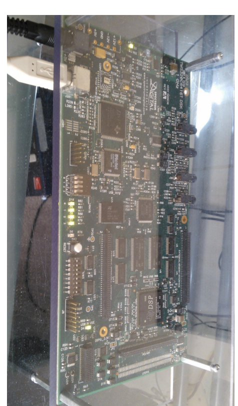
I. INTRODUCTION

THE electroencephalogram (EEG) signal is a measurement of the electrical activity present on an individual's scalp. EEG sensors measure voltage fluctuations that typically arise from ionic current flows within neurons of the brain. Although such types of evoked potentials have been measured and studied since Vladimir Vladimirovich Pravdich-Neminsky's 1912 public demonstration on a household canine, concrete analysis of such measurable electronic potentials still remains in active development [1, 2].

The reason for this is primarily due to the nature of the EEG signal itself. Since the surface scalp potential is the result of a ensemble of neural oscillations, it is extremely difficult to correlate particular voltage measurements with any particular neuron. There has been much work in this field, known as brain mapping, but the problem is almost intractable without a sufficient computational framework [6, 7] to allow for the unveiling of intimately connected neural activity. Many research groups are taking this approach to develop analysis methodologies [4, 5] .

The work presented here uses a functional understanding of EEG brain-waves to model a framework that allows for coding of information on neural channels.





In particular, by detecting the salient, or even non-obvious higher-dimensional, signal features on the numerous channels of an EEG, we demonstrate that we are able to code and decode information. Examining the extent to which information is able to be coded on these noisy neural channels will be instrumental to understanding how the brain itself handles ensemble neural signaling to the body's muscles and cells, and may provide the type of framework to fundamentally progress the field of brain mapping.

II. Detection on a Noisy Channel

Perhaps the most pressing challenge in analyzing arbitrary biological data is making dependable decisions from fundamentally noisy data. Consider **Figure 1b**, where the AF3 EEG channel is measured using the Emotiv EEG sensor headset. The electronic imprint of an individual's "blink" motor function is un-deniable; yet, it is both remarkable and baffling how small electronic events causes such a deterministic outcome. While it is hard to determine exactly the amplitude of the neural signal, it is clear that the surface potential (EEG) channel is quite noisy and thus requires significant processing before useful information may be readily extracted.

A. Removing the DC offset from sensor measurement

We developed several filtering functions that allow us to process EEG data more effectively. First, in order to remove the arbitrary DC offset introduced by the EEG measurement, we constructed a first-order 0.178Hz high-pass filter as instructed by the Emotiv headset manual. The resulting filter (**Figure 2a**) removed the DC offset of approximately 4200µV.

B. Countering transient signal noise

The sampling rate of the Emotiv headset is 128 samples/sec. Either due to inaccuracies in the measurement hardware, or stochastic potential variations, the measured EEG signal has significant high-frequency noise on nearly every channel. Such a high signal-to-noise ratio impairs feature detection accuracy and must be curtailed. We wrote a set of intuitive first-order filters that we applied to the channel data (**Figure 2b**). The derivation of such filters and coefficients is given in the appendix.

C. Feature detection

After processing our data into signals with discernible neural events, we began the process of determining the information present on the various channels. Our approach, while simple, is intuitive and analyzes EEG signals from a functional point of view; i.e. our method does not employ a biological model but, rather, determines signal features from obvious signal deviations.

Using such techniques, and for the purposes of our project, we identified three such signal features that appeared on a single noisy channel (**Figure 3**). Although the peaks are obvious to the human eye, it is an algorithmic challenge to correctly identify such events. Using our own masking methods, we generated pulse codes that describe the sequence of peaks present in any EEG channel (**Figure 4**). By applying pulse code and pulse width modulation encoding schemes, we demonstrated the ability to decode information on a neural channel.

It should be noted here that the entire data process thus far is entirely automatic and requires no human supervision. That is, our algorithm is capable of detecting events directly from recorded data, and can be directly ported to a remote-node embedded application where such neural activity is detected and reported, such as for authentication purposes [3, 8].

III. Mathematics Used

A. Bilinear Transform

We developed the equation for analog filters, and then applied the following bilinear transform on the Laplacian s variable. T is the sampling period, $1/F_s$.

$$s = \left(\frac{2}{T}\right)\left(\frac{1-z^{-1}}{1+z^{-1}}\right)$$

B. High-pass filter

Derivation of first-order high-pass RC digital filter:

$$\frac{V_o}{V_i}(s) = \frac{1}{\frac{1}{sC}+R} = \frac{1}{1+sRC}$$

Applying the bilinear transform . . .

$$y(n) = (\tfrac{1}{1+\frac{2RC}{T}})x(n) + (\tfrac{1}{1+\frac{2RC}{T}})x(n-1) - (\tfrac{1-\frac{2RC}{T}}{1+\frac{2RC}{T}})y(n-1)$$

C. Low-pass filter

Derivation of first-order low-pass CR digital filter:

$$\frac{V_o}{V_i}(s) = \frac{R}{\frac{1}{sC}+R} = \frac{sRC}{1+sRC}$$

Applying the bilinear transform . . .

$$y(n) = (\tfrac{1}{1+\frac{T}{2RC}})x(n) - (\tfrac{1}{1+\frac{T}{2RC}})x(n-1) - (\tfrac{1+\frac{T}{2RC}}{1-\frac{T}{2RC}})y(n-1)$$

D. Thresholding:

Mathematical formalism of a threshold with parameter θ:

$$t(x) = \operatorname{sgn}\left(\sum_{i=1}^{n} w_i x_i - \theta\right)$$

E. Masking:

Pseudocode of producing a mask of finite width. Such an operation can be understood as multiplying a signal by numerous, offset, rectangle functions.

Output pulse code of the logical filter = Mask (input).
C=Mput(x) if:set[j=0-3 input(x-j)]==logical code, else: 0.

IV. Embedded Processing

In order to increase the application of our algorithms, we ported our Matlab library to C and have tested it on a TI DSK6416.

We optimized our memory usage to work with the limited hardware resources. In this case, additional heap memory from the SD RAM was required to record the sampled stream and store every channel for co-current processing on the TI board. Such code can easily be tweaked to record fewer channels, but our application maintains a quick portability to other platforms because of the way we dynamically allocated the data. Our algorithmic complexity is at most $O(n^2)$ for the logical steps, but several of the functions have been optimized for $O(n\log_n)$ performance. The methodology is representative of more complex algorithms and coding schemes that may require multiple channels.

V. EEG Challenges

A. Sensors

We found that the electrical contacts provided with the EPOC headset degraded significantly over time. A green patina build-up, found to significantly limit the sensitivity and detectability of extracted features, was abated with citric acid and household vinegar treatments (**Figure 5**).

B. Consistent channel identification

Many errors in the automated PCM encoding scheme were due to the irregularity of recorded EEG channels. Repeatable placement of EEG electrodes is necessary, but not always possible. An inclusion of channel identification would be useful in this regard.

C. Robustness among different individuals

In this study, we used ourselves as test subjects for a total of two test subjects. Originally we had available to us seven volunteer test subjects, however this course was eventually abandoned. Because the equipment being used is property of the University of California, there are strict safety, privacy, and ethical requirements for conducting experiments with live test subjects – human subjects or otherwise. Passwords from each user was recorded as a single-trial event. The collection of data from more individuals using the EPOC would facilitate us to improve the robustness of our algorithms and determining the broadness of the application applicability.

VI. Encoding the Password

The signal features and pulse codes generated by NeuroPass v1.0 alone, however, do not provide the user with a complete, secure authentication system. This is because the signal features used were not unique to individual users i.e. from user to user, the present system distinguishes only the occurrence of a user's 'blink' and not the identity of the user blinking. Although NeuroPass v1.0 may not replace existing password systems, it instead augments current passwords against brute-force attacks quickly and easily.

A. *Current password systems*

At present, existing password systems typically use a combination of alphanumeric or special characters to authenticate users. For the sake of comparison, let us consider a regime with X possible characters. If the password stream is n characters long, there are, X^n possible combinations. A hacker attempting to falsely authenticate using a brute-force method would potentially need to try all 37^n combinations--an enormous amount that is virtually impossible to break with present computing resources. In practice however, passwords are often guessed due to their semblance to relatable objects, words, or numbers (such a birthday or jersey number).

B. *Password system service*

A user of NeuroPass is first instructed to create a user profile. Then the user must chose an encoding scheme for their feature events. In v1.0, this corresponds to blinks, clenches, and eye-rolling neural events. Using a combination of these recorded events (e.g. PCM, PWM), user input, and random hashing algorithms, we encode a complex password stream that is unique to each user (Table 1). The NeuroPass password-system service, which operates very in a manner very similar to key-rings, serves this encoded password to authenticated users. When a user is logged into their local NeuroPass service, the Neuropass password can be provided as a password input for logins instead of taking conventional input from the keyboard. While the input stream provided by the user can be virtually very simple, such as a series of blinks and clenches, the resulting Neuropass can be very complex (Table 2); this effectively enables the integration of neural passwords with present password systems. If X is the number of characters in the set, n is the length of a sequence, and m is the length of the X^{3mn} possible values, as opposed to the X^n values provided by present systems.

	Logical Level	Example Encoding
Blink	0	A!Rv34a
Clench	1	U8if*9
Eye-Rolling	2	B42rr5
Full Encoding	A!Rv34aU8if*9B42rr5	

Table 1: Potential User Encoding Demonstration

User's Authentication Sequence	011212
Neuropass Encoding-Authentication Combination	A!Rv34aU8if*9U8if*9 B42rr5U8if*9B42rr5

Table 2: Potential NeuroPass stored password

VII. Work Division

A. *Abhejit Rajagopal*

A. Rajagopal was responsible for the original neural password concept. Using his experience in data processing and arbitrary signal analysis, he drafted the initial code base that was used to demonstrate a proof-of-concept. Following this work, his efforts were focused on bottom-up efficient and portable algorithm design; this included optimized digital filters, masking, and encoding functions that can be ported to any platform. He was also the primary creator of the PowerPoint which was presented December 9th 2013.

B. *Anthony C. Nguyen*

A. C. Nguyen was responsible for the data collection methodology and identification of viable EEG channels. He maintained contact with Emotiv during the development cycle to resolve any technical issues with the EPOC. His software development efforts were focused on creating the user interface and password system in MATLAB. He was also primarily responsible for porting code written by A. Rajagopal into C while effectively managing memory on the TI DSK6416 embedded platform.

C. *Shared Work*

Both A. Rajagopal and A. Nguyen were involved in the logical design and coding of the NeuroPass secure password system.

VIII. Conclusion

In conclusion, our project automatically extracted neural events corresponding to an individual's *blinking*, *jaw-clenching*, and *eye-rolling* activities **with accuracy ranging from 67% to 95% on single-trial inputs, depending on the quality of the recorded data.** The security was increased by a factor of X^m, while reducing the complexity of user inputs siginificantly. By identifying more features on noisy EEG channels [9], we can increase the strength of our biometric and prevent side-channel attacks, or use our knowledge to further the applications of brain-mapping.

IX. Future Work

In future work we are investigating the extent to which information can be coded on such neural channels. Currently, we are examining various possible automated generic mapping techniques, applicable to computer vision systems.

Appendix

Our code is available for public download at http://eeg.abhe.info .

Acknowledgment

We would like to thank Dr. Dennis Briggs for supporting us throughout the quarter and for encouraging us to dare mighty things. We would also like to thank Dr. William Kaiser for providing the Emotiv EEG headset.

REFERENCES

[1] Awang, Saidatul Ardeenawatie, M. P. Paulraj, and Sazali Yaacob. "Implementing Eigen Features methods/neural network for EEG signal analysis." *Intelligent Systems and Control (ISCO), 2013 7th International Conference on*. IEEE, 2013.

[2] Übeyli, Elif Derya. "Analysis of EEG signals by implementing eigenvector methods/recurrent neural networks." *Digital Signal Processing* 19.1 (2009): 134-143.

[3] Al-Hudhud, Ghada, et al. "Using brain signals patterns for biometric identity verification systems." *Computers in Human Behavior* 31 (2014): 224-229.

[4] Jiang, Derong, and Jianfeng Hu. "Research of Computing in EEG Password based on Wavelet." *BioMedical Information Engineering, 2009. FBIE 2009. International Conference on Future*. IEEE, 2009.

[5] Zhao, Qibin, and Liqing Zhang. "Temporal and spatial features of single-trial EEG for brain-computer interface." *Computational intelligence and neuroscience* 2007 (2007).

[6] Delorme, Arnaud, and Scott Makeig. "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis." *Journal of neuroscience methods* 134.1 (2004): 9-21.

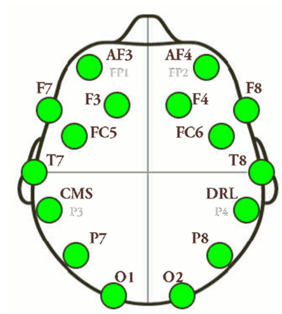
[7] Onton, Julie, and Scott Makeig. "Information-based modeling of event-related brain dynamics." *Progress in brain research* 159 (2006): 99-120.

[8] Chuang, John, Charles Wang Hamilton Nguyen, and Benjamin Johnson. "I think, therefore i am: Usability and security of authentication using brainwaves." *Proceedings of the Workshop on Usable Security, USEC*. Vol. 13. 2013.

[9] Blankertz, Benjamin, et al. "Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis." *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 11.2 (2003): 127-131.

Figure 1: Data collection channels.

A) 14 EEG channels measured on Emotiv EPOC headset



B) Example of noisy AF3 channel signal showing contamination and non-smooth data.

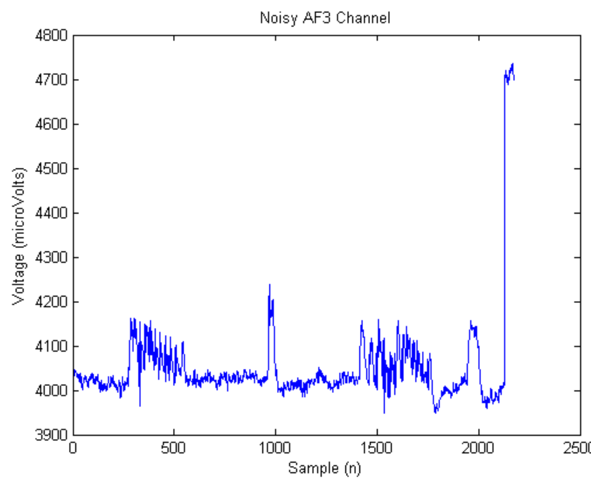
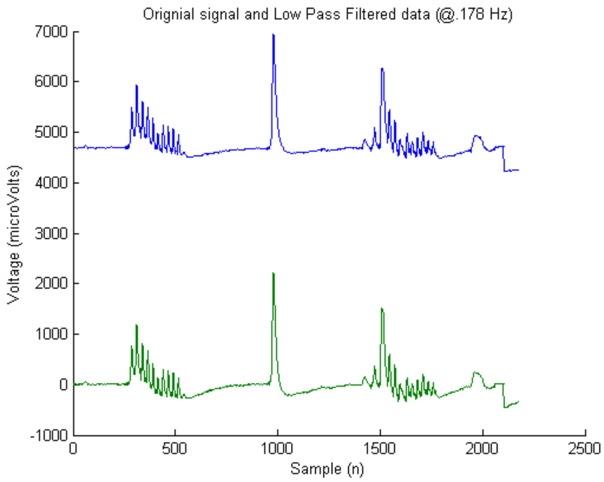


Figure 2: Digital Filtering and Signal Processing

A) Removing DC offset with a low-pass Filter



B) Applying first-order high-pass filters to enunciate local deviations

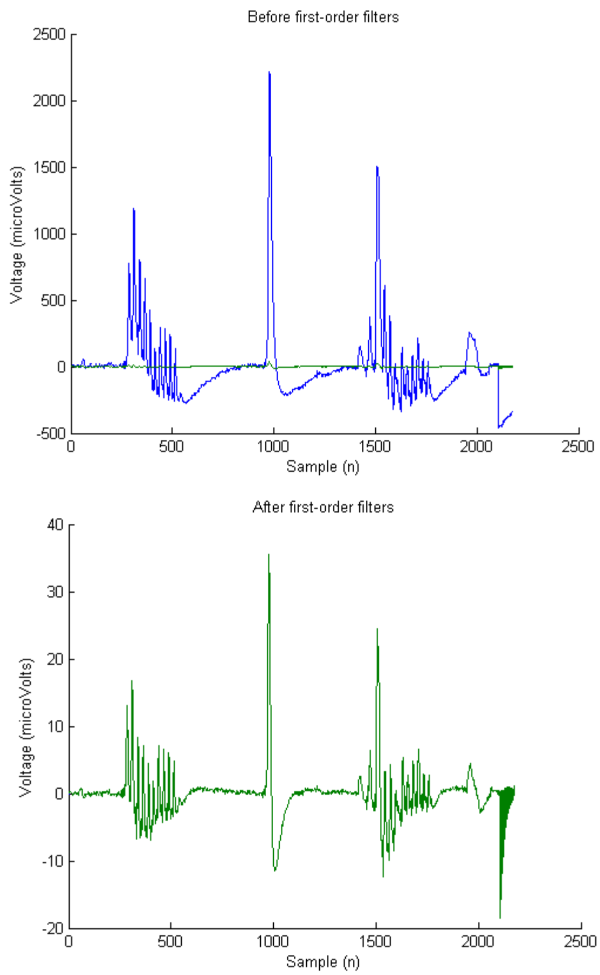


Figure 3: Detected features: blinks, clenches, eye-rolling

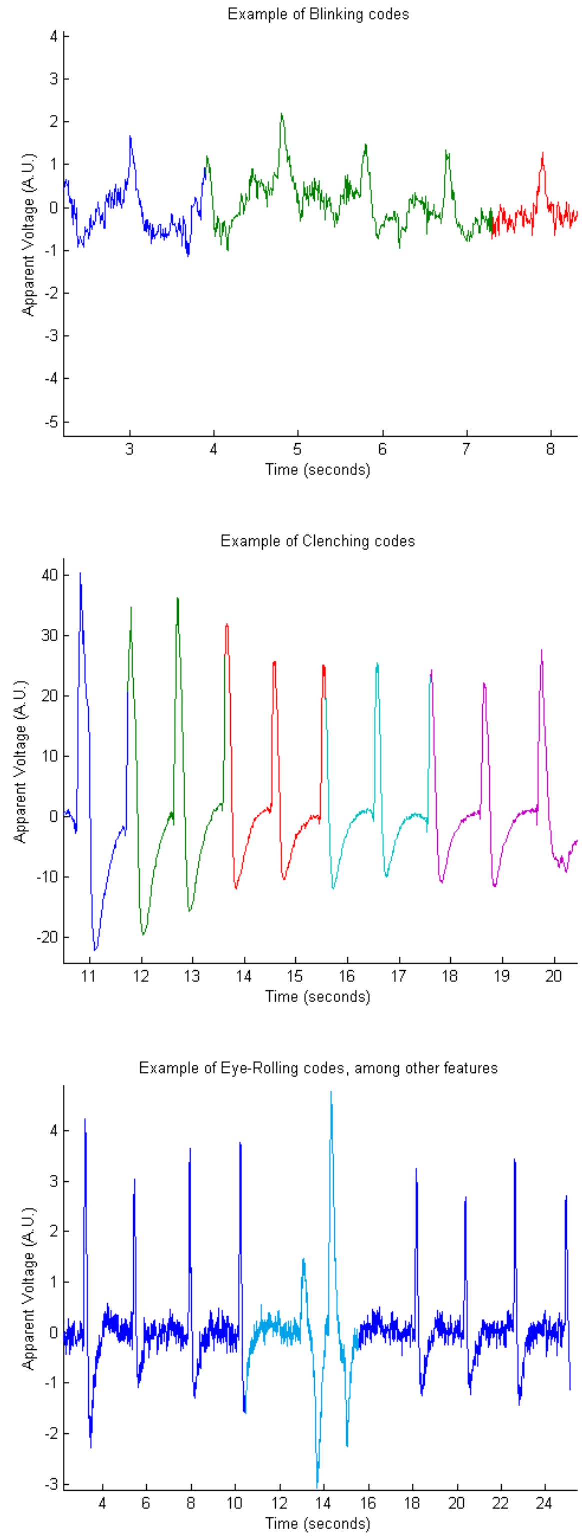
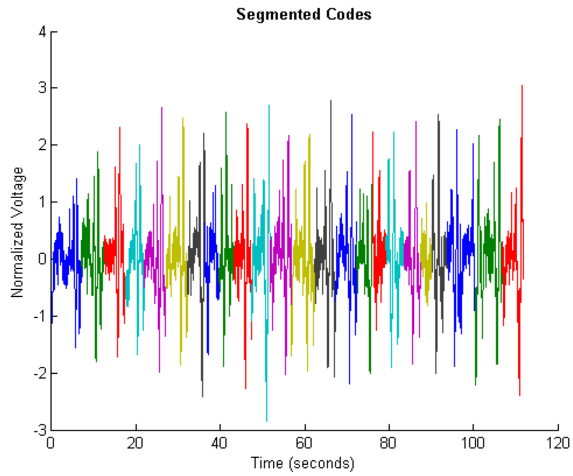
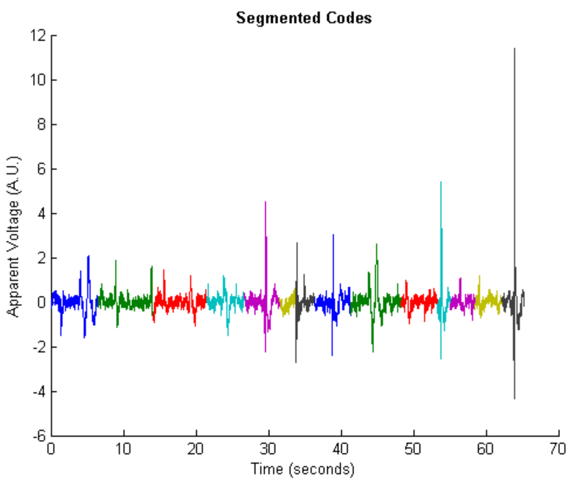


Figure 4: Coding schemes demonstrated

A) Pulse Code Modulation (PCM): Each signal "beat" was extracted. By performing analysis and statistics on each beat, the logical code was determined.





B) Pulse Width Modulation (PWM): All amplitude peaks were determined. The logical code was extracted by performing sliding window averages and measuring periods of high-event-frequency and low-event-frequency.

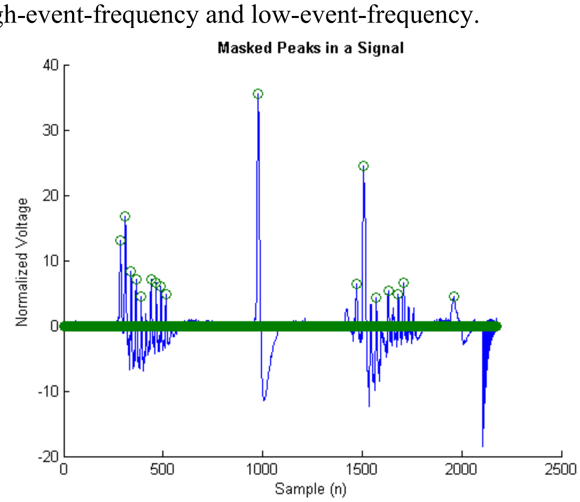


Figure 5: Copper oxidation and treatment

A) Green patina on copper (gold-plated) contacts

B) Treatment with citric acid and vinegar as solvents.

C) Clean contacts after cleaning procedure