User authentication on wearable devices by component analysis of heartbeat signals

Dmytro Progonov
Samsung R&D Institute Ukraine, Samsung Electronics, Kyiv, Ukraine
Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine

Oleksandra Sokol
Samsung R&D Institute Ukraine, Samsung Electronics, Kyiv, Ukraine

Abstract

Heartbeat authentication technology is suitable for use in mass-market wearable devices. Nevertheless, practical usage of the technology is limited due to low-accuracy cheap electrocardiogram sensors, signal fluctuations caused by varying usage contexts and high computational complexity. We proposed to use an innovative component analysis methods to address these issues by applying less resource-intensive algorithms. Evaluation results showed that proposed processing pipelines, based on multifractal analysis and singular spectrum analysis, achieved close to state-of-the-art Enamamu's accuracy (0.01% False Acceptance Rate and 0.77% False Rejection Rate), and outperformed it in various usage contexts. Obtained results confirm the effectiveness of using specialized signal processing methods for heartbeat authentication on constrained devices such as wearables.

1 Introduction

Today wearable devices, like smartwatches, have drawn special attentions as mobile companions in variety of use cases, such as messaging, healthcare, contactless payments to name a few. Despite rich functionality, wearable devices remain insecure for sensitive data processing due to the absence of strong authentication methods, such as fingerprint or facial recognition. Recently several companies proposed solutions to overcome this limitation, for example BioCatch and B-Secur HeartKey [7]. Their technologies are based on analysis of behavior and biometric data captured by smartwatches on-board sensors, including heartbeat signals (HBS), motion sensor, gait etc.

This paper concentrates on the case of user authentication on-device by analysis of HBS collected by electrocardiogram (ECG) sensors. The majority of modern heartbeat-based authentication (HBA) technology are based on processing of low-noised heartbeat signals captured by high-precision (medical) sensors. High cost of sensors and the necessity of dedicated users actions make this approach not suitable for mass-market. HBA solutions utilizing low-cost ECG sensors do exist, for example HeartKey SDK by B-Secur [7]. High detection accuracy in this SDK has been achieved through carefully selected handcrafted features that may be difficult to adopt quickly to various user activities (usage context). Therefore we target an appropriate transformation of HBS for robust estimation of informative features in various usage contexts without time-consuming manual feature selection. In this paper we provide evaluation of component analysis based novel signal processing methods, namely multifractal analysis (MFA), singular spectrum analysis (SSA), and dictionary learning methods for HBA.

The paper is organized as follows. The review of modern solution for HBA on smartwatches is presented in Section 2. Considered methods for heartbeat signals processing are described in Section 3. Results of performance evaluation of proposed method can be found in Section 4. Section 5 concludes this paper.

2 Related works

In last years, research communities and product companies have proposed several promising approaches for HBA on wearable devices, such as B-Secur HeartKey [7] and Keyble solution by FlyWallet [1]. Proposed methods can be categorized by used features in the following groups:

- *Fiducial points based methods* — are aimed at detection of specific (reference) points in heartbeat signals that can be used for both time alignment and feature extraction;

USENIX Symposium on Usable Privacy and Security (SOUPS) 2021. August 8–10, 2021, Vancouver, B.C., Canada.

- *Spectral transformation based methods* — involve specific transformation for robust extraction of user-specific features from HBS in various usage scenarios;
- *Artificial neural networks based methods* — utilize generalization ability of convolutional neural networks to mitigate HBS variability in various usage contexts.

The first group of methods is based on detection of fiducial points in heartbeat cycle captured by ECG sensors, such as P-QRS-T complex and Systolic/Diastolic peaks ratio [3]. Despite variety of such methods, their performance highly depends on smartwatch usage context, namely persons' movements during signals capturing.

Patil et al. have shown high efficiency of combining multiresolution wavelet analysis and convolutional neural network [10]. Their solution achieved nearly 98% accuracy for detection of HBS patterns for heart disease diagnosis. However, practical application of the solution on wearables is limited due to high computational complexity.

Spectral-based HBA methods are of special interest today. They allow for combining low-complexity decomposition methods, such as Fast Fourier Transform, with high detection accuracy. However, practical application requires careful selection of an appropriate decomposition method and basis functions. Selection of these parameters may be time-consuming and is prone to overfitting on new databases. Thus the research is needed in advanced processing methods for fast and reliable detection of user-specific features in HBS without pre-computation.

Major deliverable of this paper is HBA technology suitable for use in various user activity contexts in mass-market wearables. We apply multifractal analysis, singular spectrum analysis and dictionary learning methods to address typical issues such as low-accuracy mass-market ECG sensor and signal fluctuations caused by varying usage contexts.

3 Advanced methods for heartbeat signals processing

The classical spectral transformations, such as Fourier and wavelet transform, are widely used in various signal enhancement and reconstruction applications. Nevertheless, these methods have limited opportunities for capturing and compact representing user-specific alterations of HBS by changing of usage context. Providing high authentication accuracy requires usage of several decomposition methods that may be inappropriate for resource-limited devices, such as wearables.

One of the methods for overcoming mentioned limitation are applying of special transformations methods for robust feature estimation from highly distorted signals. We proposed to use advanced component analysis methods for solving this task. Feature of these methods is compact (learned) representation of heterogeneous signals that simplifies detection and tracking of non-linear distortion in signals. Among known component analysis methods, special interest is taken to multifractal analysis, singular spectrum analysis and dictionary learning methods that characterize low computation complexity of signal decomposition procedure. This makes these methods attractive candidates to be used in HBA. Let us consider multifractal analysis, singular spectrum analysis and dictionary learning methods in more details in next sections.

3.1 Multifractal analysis

The multifractal analysis is widely used for signal processing in various domains, such as material science, digital signal processing, healthcare applications [8]. The MFA is aimed at evaluation of statistical parameters of heterogeneous signals (multifractals) that provide local power law dependency of adjacency elements values with singularity exponent $h(t)$ [8]:

$$(s(t+a)-s(t)) \propto a^{h(t)}, a > 0.$$

The analysis have been successfully applied for early detection of heart disease, as well as revealing HBS parameters variations under various physical activities [8]. This makes MFA a promising candidate for HBA systems.

Evaluation of statistical parameters of heterogeneous signals by MFA is done with the generalized fractal dimensions spectrum D_q and corresponding to it singularity (multifractal) spectrum $f(\alpha)$ [8]. The D_q spectrum generalizes Shannon entropy to be evaluated on different time scales by varying of scaling parameter q, $q \in \mathbb{R}$. The multifractal spectrum $f(\alpha)$ can be interpreted as the spectrum of Hausdorff-Besicovitch dimensions D_{HB} of signal components that have the same probabilities of blocks filling — $p(w) \propto w^{\alpha}$.

The MFA allows for analysing the fine structure of signal by its decomposition at different time. On the other hand, extracted components may be hard to interpret, in particular by analysis of HBS. Therefore, it represents the interest to use data processing methods that provides clear interpretation of extracted components, such as trend-like or noise-like parts.

3.2 Singular spectrum analysis

Estimation of noised signal parameters requires its decomposition into components that can be easily checked depending on theirs features. In most cases, these components are extracted using classical spectral transforms, such as Fourier Transform, Wavelet Transform etc. These transforms use predefined basis functions adapted to revealing known signal-specific features, such as signal shape or presence of specific pulses. Nina Golyandina proposed signal decomposition methods that can adaptively extract signal components even under limited prior information about signal features [6]. Proposed Singular Spectrum Analysis (SSA) is aimed at decomposition of a signal (time series) $\mathbf{S} = (s_1, \ldots, s_{n_\mathbf{S}})$ into a sum of series, so that each component can be identified as either a trend, periodic or quasi-periodic component, or a noise.

The signal decomposition by SSA includes two procedures — signal embedding and further singular value decomposition (SVD). The embedding can be regarded as a mapping of one-dimensional time series $\mathbf{S} = (s_i)_{i=1}^{n_\mathbf{S}}$ into the multidimensional series $(\mathbf{X}_j)_{j=1}^{K}$ with vectors $\mathbf{X}_j = (s_i)_{i=1}^{j+L-1}$, where $K = M - L + 1$ and window length is $L \in [2; M)$. Obtained vectors are packed into the trajectory matrix $\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_K]$.

The second step of decomposition is SVD transform of obtained trajectory matrix $\mathbf{X}$. The singular value decomposition represents matrix $\mathbf{X}$ as a sum of rank-one bi-orthogonal elementary matrices. Denote by $(\lambda_i)_{i=1}^{L}$ the eigenvalues of $\mathbf{X}\mathbf{X}^T$ sorted in decreasing order. Then, $(\mathbf{U}_j)_{j=1}^{L}$ is the orthogonal system of normalized eigenvectors corresponding to these eigenvalues. Denote by $\mathbf{V}_i = \mathbf{X}^T \mathbf{U}_i / \sqrt{\lambda_i}$, then SVD of the trajectory matrix $\mathbf{X}$ can be written as:

$$\mathbf{X} = \sum_{i=1}^{d} \mathbf{X}_i = \sum_{i=1}^{d} \sqrt{\lambda_i} \mathbf{U}_i \mathbf{V}_i^T .$$

The matrices $\mathbf{X}_i$ have rank 1, so they are elementary matrices. The vectors $\mathbf{U}_i$ and $\mathbf{V}_i$ stand for the left and right eigenvectors of the trajectory matrix. Then, the collection $(\sqrt{\lambda_i}, \mathbf{U}_i, \mathbf{V}_i)$ is called $i-$th eigentriple of matrix $\mathbf{X}$, where $\{\sqrt{\lambda_i}\}_{i=1}^{d}$ is spectrum of the matrix $\mathbf{X}$.

The SSA method provides theoretically proved approach to decompose the signal into "explainable" components, such as trends, trend-like and noise-like signals. Analysis of extracted components allows for selecting appropriate components to be used as features during user authentication. Nevertheless, these components should be manually selected, which may be inappropriate for processing a huge dataset. Therefore it can be useful for decomposition methods that "learn" appropriate representation from the signal itself. One well-known example of such "trainable" decomposition methods is dictionary learning presented in the next section.

3.3 Dictionary learning from signals

Sparse signal decomposition in redundant dictionaries is widely applied to noise removal, signal compression, and pattern recognition [2]. This approach allows for learning user-specific features from captured HBS, namely explained by physiology-related features.

Let us discuss the widespread learning methodology for constructing a dictionary $\mathbf{A}$. Assume that a training database $\{\mathbf{S}_i\}_{i=1}^{n_\mathbf{S}}$ is given. This database was generated with some fixed but unknown dictionary $\mathbf{A}$ that may be learned by solving of the following optimization problem:

$$\min_{\mathbf{A}, \{\mathbf{x}_i\}_{i=1}^{n_\mathbf{S}}} \sum_{i=1}^{n_\mathbf{S}} ||\mathbf{x}_i||_0 \text{ subject to } ||\mathbf{S}_i - \mathbf{A}\mathbf{x}_i||_F \leq \varepsilon, i \in [1; n_\mathbf{S}], \quad (1)$$

where $\|\cdot\|_0$ is L_0-norm of a vector, and $\|\cdot\|_F$ is Frobenius norm of a matrix [6].

This optimization problem describes each given signal $\mathbf{S}_i$ as the sparsest representation $\mathbf{x}_i$ over the unknown dictionary $\mathbf{A}$, and aims to jointly find the proper signal representation and the dictionary.

The optimization problem (1) may be solved with well-known K-SVD algorithm [2]. Keeping all the columns fixed apart from the j_0-th one, the $\mathbf{a}_{j_0}$ column can be updated along with the coefficients that multiply it in $\mathbf{X}$. We isolate the dependency on $\mathbf{a}_{j_0}$ by re-writing (1) as [2]:

$$||\mathbf{S} - \mathbf{A}\mathbf{X}||_F^2 = \left\| \mathbf{E}_{j_0} - \mathbf{a}_{j_0} \mathbf{x}_{j_0}^T \right\|_F^2 .$$

Here, $\mathbf{x}_j^T$ stands for the j-th row of $\mathbf{X}$, while $\mathbf{E}_{j_0} = (\mathbf{S} - \sum_{j \neq j_0} \mathbf{a}_j \mathbf{x}_j^T)$ is a pre-computed error matrix.

In order to minimize the term $\mathbf{E}_{j_0}$ while keeping the cardinalities of all the representations fixed, a subset of the columns of $\mathbf{E}_{j_0}$ should be taken — those that correspond to the signal from the training set that are using the j_0-th atom. Therefore, we define a restriction operator, $\mathbf{P}_{j_0}$ that remove the non-relevant columns. Therefore, $(\mathbf{x}_{j_0}^R)^T = \mathbf{x}_{j_0}^T \mathbf{P}_{j_0}$ as the restriction on the row $\mathbf{x}_{j_0}^T$, choosing the non-zero entries only.

The update step targets both $\mathbf{a}_{j_0}$ and $\mathbf{x}_{j_0}^T$ by the block-coordinate Least-squares approach:

$$\mathbf{x}_{j_0}^R = \left(\mathbf{P}_{j_0}^T \mathbf{E}_{j_0}^T \mathbf{a}_{j_0} \right) / \|\mathbf{a}_{j_0}\|_2^2, \mathbf{a}_{j_0} = \left(\mathbf{E}_{j_0} \mathbf{P}_{j_0} \mathbf{x}_{j_0}^R \right) / \|\mathbf{x}_{j_0}^R\|_2^2. \quad (2)$$

In most cases, a few rounds of updates for (2) are sufficient for getting the desired dictionary $\mathbf{A}$ [2].

In our research, we used dictionary-learning methods for creating user-specific basis for HBS decomposition. The main idea is that heartbeat signals of target user may be sparsely represented in learned dictionary, while signals of another person require much "denser" representation. Let us compare decomposition vectors $\mathbf{x}$ for target and another users heartbeat signals by usage of learned dictionary $\mathbf{A}$ (Table. 1).

Table 1: Decomposition coefficients for heartbeat signals captured for pseudo randomly selected users from in-house dataset for predefined usage context (users are lying).

	Coefficient #1	Coefficient #2	Coefficient #3	Coefficient #4
Target user				
Signal #1	0.000	-0.455	20.736	0.000
Signal #2	-19.758	-1.199	0.748	0.000
Signal #3	-19.233	0.000	2.264	0.000
Another user				
Signal #1	-4.833	-3.083	12.747	2.451
Signal #2	-12.744	-2.343	6.621	1.927
Signal #3	-6.849	-3.610	11.030	1.161

Learned dictionary $\mathbf{A}$ allows for obtaining quasi-sparse decomposition for heartbeat signals of target user (Table. 1). Applying the same dictionary to HBS of another person leads to non-zeros decomposition coefficients. In addition, decomposition coefficients for another person preserve similar signs (Table. 1) that differ from corresponding signs for target user. These features may be used for building an HBA systems.

Let us analyze the performance of presented methods on both publicly open and in-house datasets. The corresponding results are presented in the next section.

4 Experiments

Proposed solution was implemented for Samsung Galaxy Watch 3 smartwatch. Evaluation has been done for two cases — processing of HBS captured in fixed (supine) position of persons, as well as signals collected in various usage contexts. The former one corresponds to main use cases for state-of-the-art solutions, where distortions of collected signals can be minimized. The latter one allows us evaluating performance of proposed solutions in use cases closed to real usage situations, e.g. during persons walking.

The examples of HBS for fixed (supine) position of persons were taken from publicly available datasets presented in PhysioNet service:

- *Combined measurements of ECG, Breathing and Seismocardiograms* [5] – the dataset includes measurements of 20 presumably healthy volunteers. Approximately, 867 signals per-user were captured.
- *Wilson Central Terminal ECG Database* [9] – the dataset includes ECG signals recorded from 92 patients (14 signal-per-user on average) with I – III Leads.

The in-house dataset includes ECG-signals captured with Samsung Galaxy Watch 3 by presumably healthy volunteers (25 logs-per-user with 30 heartbeat-per-log on average) in widespread usage contexts (during lying, sitting, still standing and walking). Due to COVID-19 quarantine limitations, the in-house dataset was collected for small group of six persons (two men and four women) that represents the target auditorium of smartwatch users (20-30 years old). The HBS were extracted from captured ECG-signals using well-known Pan-Tompkins algorithm. Processed signals were aligned to R-to-R points of QRS complex of heartbeat signals.

Accuracy analysis was done according to 10-times cross-validation procedure by splitting HBS collection sessions between training (90%) and testing (10%) sets. Evaluation was performed using binary classifiers, such as Random Forest (RF), Decision Tree (DT), k-nearest neighbors (kNN) and Naïve Bayes (NB). False Acceptance Rate (FAR) and False Rejection Rate (FRR) were used as accuracy metrics.

During performance analysis, the following parameters of considered signal processing methods were used:

- *Multifractal analysis*: the scaling parameter q was changed from $q_{min} = (-25)$ to $q_{max} = 25$ with step $\Delta_q = 0.1$.
- *Singular spectrum analysis*: the parameter L during signal embedding stage was changed from $L_{min} = 2$ to $L_{max} = 1000$ with step $\Delta_L = 2$.
- *Dictionary learning*: the lasso-lars algorithm was used for components (dictionary atoms) extraction. The number of extracted components was varied from $N_{min} = 5$ to $N_{max} = 20$ with step $\Delta_N = 5$.

Obtained multifractal spectrum, SSA spectrum and decomposition coefficients in learned dictionary were used as features for HBS classification during user authentication.

Proposed solution was compared with state-of-the-art Discrete Wavelet Transform (DWT) based method proposed by Enamamu et al. [4]. The Enamamu's solution uses statistical parameters of DWT coefficients for ECG-signals as features during user authentication. Heartbeat signals features were extracted using widespread wavelet basis, such as Daubechies, biorthogonal, Meyer to name a few.

The estimated FAR and FRR values for considered signals transformation methods and state-of-the-art Enamamu's solutions are presented in Table 2.

Table 2: Error rate for considered signal transformation methods and state-of-the-art Enamamu's solution by usage of ECG-signals from open databases.

Database	Used features	Classifier	FAR	FRR
Wilson Central Terminal ECG dataset	Multifractal analysis	Naive Bayes	4.97	8.16
	Singular spectrum analysis	kNN	0.08	11.95
	Dictionary learning	Naive Bayes	70.95	6.33
	Enamamu's solution	bior1.5 + NB	3.61	2.08
ECG, Breathing, Seismo-cardiograms dataset	Multifractal analysis	Random Forest	0.03	1.77
	Singular spectrum analysis	Random Forest	0.01	0.68
	Dictionary learning	Decision Tree	0.42	51.62
	Enamamu's solution	dmey + RF	0.01	0.77

Let us note large values of FAR (FAR $\sim 0.5\%$) and FRR (FRR $\sim 50\%$) metrics for Dictionary Learning method (Table 2) in comparison with other methods. This can be explained by high similarity of captured HBS that leads to high similarity of learned dictionaries $\mathbf{A}$ for users. On the other hand, MFA and SSA methods allow for achieving authentication accuracy on open datasets that is close to the state-of-the-art Enamamu's solution (Table 2).

On the next stage False Acceptance Rate and False Rejection Rate were calculated by using HBS captured for various usage contexts (user's physical activities). Evaluated metrics for considered signal processing methods and Enamamu's solution are presented in Table 3.

Obtained results for different usage context (Table 3) are close to obtained for open datasets (Table 2). Using Dictionary Learning techniques leads to considerable increase of FRR values (up to 76%) by preserving FAR values close to state-of-the-art solutions. Usage of multifractal and singular spectrum analyses allowed to achieve close to Enamamu's method authentication accuracy. Note that MFA has shown quite bigger error levels in comparison with SSA (Table 3). This is because of better adaptation to analyzed signals – presence of noise-like parts in HBS modelled by several monofractal components, that can be suppressed by grouping procedure on reconstruction stage of SSA.

Table 3: Error rate for considered signal transformation methods and state-of-the-art Enamamu's solution by usage of ECG-signals from in-house dataset in various usage context.

Database	Used features	Classifier	FAR	FRR
Users are lying	Multifractal analysis	Random Forest	2.66	39.90
	Singular spectrum analysis	Decision Tree	5.49	28.06
	Dictionary learning	Random Forest	0.21	60.95
	Enamamu's solution	dmey + RF	1.76	22.15
Users are sitting	Multifractal analysis	Random Forest	2.14	30.28
	Singular spectrum analysis	Decision Tree	4.12	20.53
	Dictionary learning	kNN	2.68	58.62
	Enamamu's solution	dmey + RF	0.95	18.79
Users are standing	Multifractal analysis	Decision Tree	6.54	29.51
	Singular spectrum analysis	Decision Tree	3.61	17.95
	Dictionary learning	Random Forest	0.56	63.06
	Enamamu's solution	dmey + RF	0.88	14.25
Users are walking	Multifractal analysis	Naive Bayes	34.69	20.50
	Singular spectrum analysis	Decision Tree	7.70	36.98
	Dictionary learning	kNN	3.41	76.66
	Enamamu's solution	dmey + RF	2.03	45.37

5 Conclusion

Heartbeat-based user authentication provides spoofing-proof, transparent and convenient way for person on-device authentication without using additional trusted smartphone. Modern solutions for HBA are based on utilization of embedded sensors for capturing heartbeat signals, applying spectral decomposition and artificial neural network based authentication methods. These methods provide high authentication accuracy by using redundant set of basis functions that complicates authentication process. In the paper, we presented HBA technology based on multifractal analysis, singular spectrum analysis and dictionary learning methods, that allow representing of heterogeneous signals in a compact way.

Results of accuracy evaluation has shown that considered methods allows for achieving close to state-of-the-art authentication accuracy by using less computationally-intensive algorithms. The MFA and SSA-based processing pipelines for heartbeat signals achieved similar to state-of-the-art Enamamu's solution for open dataset (0.01% FAR and 0.77% FRR), and outperformed it in various usage contexts (7.70% FAR and 36.98% FRR). The Dictionary Learning based techniques have shown unsatisfactory performance (up to 50% FRR values) for all considered cases, which makes them inapplicable for heartbeat-based authentication systems.

Therefore, the SSA may be recommended as promising alternative to known solutions by preserving similar or higher accuracy. On the other hand, this method may be inappropriate for some GPU-limited wearables due to extensive usage of matrix operations. So, the MFA may be used as a good candidate for HBA on these devices.

Finally, we would like to mention that presented results may provide estimation of error levels for main use cases only (during persons lying, sitting, still standing and walking) of target audience of smartwatch users (20-30 years old). Still, performance of considered solutions for wider consumers groups (i.e. children and elderly persons), as well as in specific usage context, such as in transports, during workout, needs additional research that we would like to aim in the papers.

References

[1] CES 2021: Wearable Device with Biometric Authentication. https://www.eetimes.eu/ces-2021-wearable-device-with-biometric-authentication/. Accessed: 2021-03-02.

[2] M. Aharon, M. Elad, and A. Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.*, 54, 2006.

[3] Se Young Chun, Jae-Hwan Kang, Hanvit Kim, Chungho Lee, Ian Oakley, and Sung-Phil Kim. Ecg based user authentication for wearable devices using short time fourier transform. In *Int. Conf. on Telecommun. and Signal Process. (TSP)*. IEEE, June 2016.

[4] Timibloudi Enamamu, Abayomi Otebolaku, Jims Marchang, and Joy Dany. Continuous m-health data authentication using wavelet decomposition for feature extraction. *Sensors (Basel)*, 20, 2020.

[5] Miguel A. Garcia-Gonzalez, Ariadna Argelagos-Palau, Mireya Fernandez-Chimeno, and Juan Ramos-Castro. A comparison of heartbeat detectors for the seismocardiogram. In *Int. Conf. on Comput. and Cardiology*. IEEE, September 2013.

[6] Nina Golyandina and Anatoly Zhigljavsky. *Singular Spectrum Analysis for Time Series*. Springer, 2020.

[7] Antuan Goodwin. B-Secur HeartKey tech unlocks your car with unique rhythm of your heartbeat. https://www.cnet.com/roadshow/news/b-secur-ekg-heartkey-tech-unlocks-car-with-heartbeat/. Accessed: 2020-12-04.

[8] David Harte. *Multifractals: Theory and Applications*. Chapman and Hall/CRC, 2001.

[9] H. Moeinzadeh and G. Gargiulo. Wilson Central Terminal ECG Database (version 1.0.1). PhysioNet. doi: 10.13026/f73z-an96.

[10] Dinesh D. Patil and R.P. Singh. Ecg classification using wavelet transform and wavelet network classifier. In *Advances in Intelligent Systems and Computing*, pages 289–303. Springer, 2018.