Comput Syst Sci & Eng (2018) 1: 21–29

International Journal of
Computer Systems
Science & Engineering

Probabilistic analysis of electrocardiogram (ECG) heart signal

Amjad Gawanmeh[1,3,*], **Usman Pervez**[2] **and Osman Hasan**[2,3]

[1] *Department of Electrical and Computer Engineering, Khalifa University of Science and Technology, Abu Dhabi, UAE*
[2] *School of Electrical Engineering and Computer Science, National University of Sciences and Technology (NUST), Islamabad, Pakistan*
[3] *Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada*

Electrocardiography (ECG) is a heart signal wave that is recorded using medical sensors, which are normally attached to the human body by the heart. ECG waves have repetitive patterns that can be efficiently used in the diagnosis of heart problems as they carry several characteristics of heart operation. Traditionally, the analysis of ECG waves is done using informal techniques, like simulation, which is in-exhaustive and thus the analysis results may lead to ambiguities and life threatening scenarios in extreme cases. In order to overcome such problems, we propose to analyze ECG heart signals using probabilistic model checking, which is a formal methods based quantitative analysis approach. This work presents the formal probabilistic analysis of ECG signal abnormalities where the likelihood of abnormal patterns is studied and analyzed using the PRISM model checker

Keywords: Electrocardiography, ECG, Formal Modeling, Probabilistic Analysis

1. INTRODUCTION

Electrocardiography (ECG) [11] is a standard representation of the heart's electrical activity, which is recorded by placing biosensors on the body surface. ECG is widely used in diagnosing patients with several heart-related diseases, such as atrial and ventricular arrhythmia [6]. The diagnostic process is usually based on certain guidelines and it relies heavily on ICT healthcare services. The diagnostic guidelines are used to devise medical protocols, which in turn facilitate the understanding and developing diagnostic procedures using recorded ECGs. Due to the complexity of ECGs and the informal representation of these waves, the diagnosis may result in several ambiguities and inconsistencies when they are used by practitioners. As reported by García *et al.* [7], in order to provide correct interpretation of ECGs, practitioners must use approved standards. However, these standards are not always strictly followed, as reported by García *et al.*, which leads to many problems in medical diagnosis. Moreover, the enormous development of ICT based health systems resulted in new challenges [8]. For instance, an incomplete analysis of a medical protocol may lead to undetected bugs, which eventually can lead to disastrous consequences if these bugs cause a malfunction in the healthcare system or lead to miss a vital sign of the human subject. For example, the infamous software bug in Therac-25, i.e., a cancer therapy machine was mainly responsible for the loss of three human lives and three severe injuries [20].

Formal methods [1] utilize mathematical reasoning to conduct systematic analysis and have been successfully used to verify a variety of complex systems [5]. Hence, formal methods frequently have also been frequently used for reliability analysis of medical systems, where large and complex records of medical data are handled through various operations, such as data recording, analysis, acquisition, interpretation, and transformation. However, most of the existing methods for analyzing ECG related systems and protocols are based on informal methods, like simulation. The main reason for this gap is the complex and randomized behaviors of the ECG signals, which cannot be

*Correspondence Author. E-mail: amjad.gawanmeh@kustar.ac.ae

formally expressed in a straightforward manner. The first issue is that the ECG signal is composed of several waves. The second main complexity in formalizing the ECG waves is the relationship of the ECG waves and their normal and abnormal shape with the sinoatrial and atrioventricular nodes.

Given the randomized nature of the ECG behavior, we propose to develop formal probabilistic models for ECG behaviors and then use probabilistic model checking [4], i.e., a state-space based formal verification technique that allows expressing and verifying probabilistic temporal logic properties, for their analysis. In particular, the paper presents a formal specification of the ECG signal using the language of the PRISM Model checker [19] and the verification of some of its most desired probabilistic properties, such as the tendency of the heart to generate ECG abnormalities.

The rest of the paper is organized as follows: Section 2 presents a review of the related work and a brief overview of probabilistic model checking. The commonly used ECG specifications are given in Section 3. The formal probabilistic analysis of the ECG signal is given in Section 4, and finally Section 5 concludes the paper.

2. RELATED WORK AND PRELIMINARIES

The use of computing systems to enhance the functionality and analysis of medical systems has been an active area for quite a long time, with several emerging technologies are being used in the medical field. For instance, Wong *et al.* [22] proposed an approach to achieve real-time TCM (Traditional Chinese Medicine) telemedicine through the use of ontology and clinical intelligence discovery. In another work, Penna *et al.* [9] applied visual information extraction to biomedical applications. Such contributions shows that up to date computing based technologies can have a major impact in this safety-critical area. Similarly, modeling and analysis of medical systems has also been an important and active area of research. The authors in [3] proposed an approach for the automatic verification of the control unit of a medical biosensor that can be used for recording the ECG and medical diagnosis to detect abnormal behaviors in the heart operation. The work in [2] presented embedding ECG wave specifications in Event-B, where the authors formally analyzed ECG signal components and then verified different requirements about the signal components.

Several other works on formal analysis of medical systems were discussed in the literature including the work in [15] for verifying the software of a medical infusion pump, the work of [12] for modeling the software of X-ray medical devices of Philips Healthcare, the work in [17] for modeling rate-adaptive pacemaker for controlling the heart rate, the work of [21] for modeling irregular conditions of both healthy and cardiac humans based on the heartbeat level of ECGs, where the segment level information of the ECG signal was used as an element of analysis. In this work, we intend to utilize the formal specification of the ECG signal, presented in [2], to conduct the formal probabilistic analysis for ECG signals. To the best of our knowledge, none of the above-mentioned research works tackled the problem of formal probabilistic analysis of ECG signals and their implications on heart disease diagnostics, which is the main scope of the current paper.

Model checking [4], sometimes called property checking, is a formal verification method that is commonly used to verify systems that can be modeled as state-transition graphs. It is commonly used to verify controllers, and communication protocols, where the verification is conducted exhaustively by exploring all the states of the model and automatically. The properties to be verified are modeled in temporal logic. Model checking tools provide the failure trace in case a property is not valid. The main disadvantage of the model checking approach is that system states can grow exponentially and thus lead to the state-space explosion problem, which is the problem of computationally handling the verification of system with a large state-space. A possible solution to this problem is to use a less complex model by using abstractions.

Probabilistic model checking [19] is a formal analysis method where the behavior of the system is described using a Markovian model and probabilistic properties can be defined and verified for this model. There are several probabilistic model checkers in the literature, such as PRISM [18], VESTA [16] YMER [24], ETMCC [13], and MRMC [14]. We have used the PRISM model checker in this work as it supports a wide range of modeling options, such as Continuous Time Markov Chains (CTMCs), Discrete Time Markov Chains (DTMCs) and Markov Decision Processes (MDPs). The system to be verified in PRISM is modeled using the PRISM modeling and specification language. PRISM also supports a wide range of property specification languages, such as PCTL, CSL, LTL and PCTL. For instance, $S_{\geq 0.5}[``normal"]$ is interpreted as the steady state probability of *normal* state $\geq$ 0.5. Once the given system is modeled with Continuous Time Markov Chain (CTMC) and is implemented in PRISM, the reliability *properties* are defined according to the needs and are verified to find the results. Another main reason for using the PRISM model checker in this work is to find failure and success probabilities about ECG specifications. Other tools, such as YMER and VESTA, are not suitable for the type of system considered in this work, as they do not support steady-state probabilities [30].

3. ELECTROCARDIOGRAM (ECG) SPECIFICATIONS

This section describes the specifications of Electrocardiography (ECG) [11], which is an electrical signal depicting the activity of heart and are recorded by medical biosensors that is attached from the body of human beings. The specifications of the ECG wave presented in this paper is based on the the formalization conducted in [2]. The record is a repetitive repetitive signal with specific patterns, and is commonly used in the medical diagnosis process, where certain ECG characteristics may lead to the diagnosis of different heart problems. In this section, the informal specifications of the fundamental parts of the ECG, as recorded by a standard 12-lead ECG, are described. In addition, the sequence of events that describes how these patterns are generated is introduced as well [23].

Figure 1 shows the process of generating ECG signal in the form of different waves by different parts of the heart. ECGs are

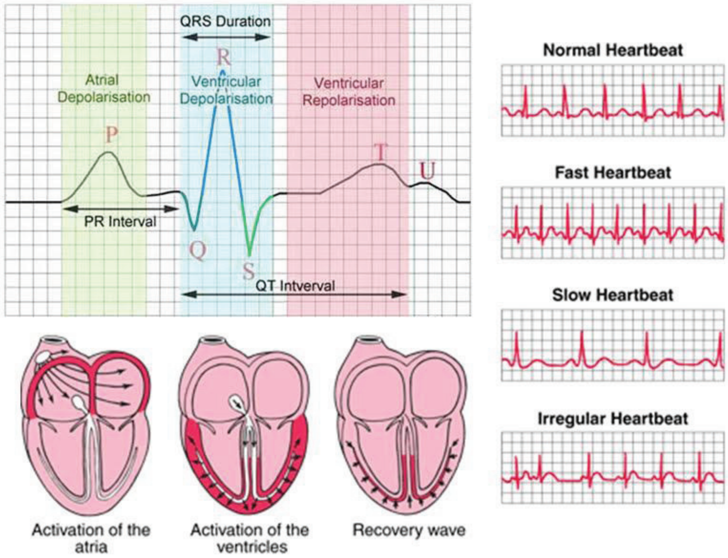

Figure 1 ECG Waves for Heart nodes [11]

produced in three phases: In the first phase, the *P* wave is generated by the sinoatrial node (SA) (the activation of the atria). In the second phase, the *QRS* wave is generated by the atrioventricular node (AV) (the activation of the ventricles). Finally, the *T* wave is generated from the heart muscles in the recovery phase, and the heart produces flat signals that indicate no-activity period between waves. These flat segments may occur between any two waves. These waves, have several characteristics, such as their frequency of occurrence, the duration of active wave, the number of times they are repeated, and amplitude. These facts, can give indications about heart's behavior and operation, such as normal or abnormal status, fast beat or slow beat, etc.

The heart activities, reflected by the ECG components and their sequence can be divided into two main parts: *active waves* and *idle segments*. The active waves represent three main signals: *P*, *QRS* and *T* waves. The idle part of the ECG signal includes three periods: PQ-segment, ST-segment and TP-segment. While the idle segments exhibit similar behavior and different durations, that active ones may exhibit different behaviors, repetition, existence and durations. Figure 2 illustrates the ECG signal and the components described above.

Figure 3 shows how ECGs are generated from basic comports and all possible scenarios from which ECG signal is composed of. The heart starts its activity from the idle segment, which is indicated by label 1. All possible ECG alternatives generated by heart nodes can be achieved by traversing from one component to another through numbered labels. As per the ECG specifications, there are four distinct cases [2]:

Case 1: In this case, neither the *SA* nor the *AV* nodes are working. This is where we get no pulse from the heart and only a repeated occurrence of the idle segment is recorded. This is illustrated in Block 1 in Figure 3.

Case 2: When only the *SA* node is working. In this scenario, the heart can generate only *P* waves and idle segments. This is the abnormal behavior of the heart, where *P* waves are repeatedly generated idle segments. This is illustrated by Blocks 2 and 3 in Figure 3. In addition, if the *AV* node becomes active, the ECG will have a different pattern, while the inactivity of the *SA* node will result in generating an ECG according to teh Case 1 above.

Case 3: The *AV* node is working in this case only, therefore, the heart generates the *QRS* waves, which can be followed by a *T* wave, idle segments, or both. This case can be further divided into three situations: the first is called the fatigued ventricles activation, where *QRS* is generated repeatedly (Block 5 in Figure 3). The second is called the less fatigued ventricles activation, where the heart generates the idle flat segment and *QRS* wave alternatively (Blocks 5 and 6 in Figure 3). The third situation is called the recovery state, which allows the heart to recover from the ventricles activation. This may occur when the heart generates *QRS* and *T* waves along with flat segments in between (blocks 5, 7, 8, and 9 in the Figure 3). The heart system can be be in any one of these three situations while only the *AV* node is working.

Case 4: The *SA* and *AV* nodes are working is the normal case of the heart, where all waves occur in the generated signal in the order: the *P*, *QRS* then *T* wave. The idle segments can occur anywhere in between. This is illustrated by the sequence of blocks. The status of the ECG can be transferred from this case into any other only if *SA* or *AV*, or both stop working.

4. PROBABILISTIC ANALYSIS OF ECG HEART SIGNAL

In order to formally analyze the probabilistic behavior of a heart based on its ECG, the ECG behavior needs to be formally modeled first. We have chosen to use a CTMC to develop this model as ECG is recorded over a period of continuous time intervals. In addition, using CTMCs allows us to define the state transition probabilities probabilistic decisions. The objective of the proposed probabilistic analysis is to find the probabilities of occurrence of different behaviors of the heart, such as normal, abnormal, dead, fatigued and less fatigued. In turn, these probabilities can be used to diagnose the problems in the heart as well

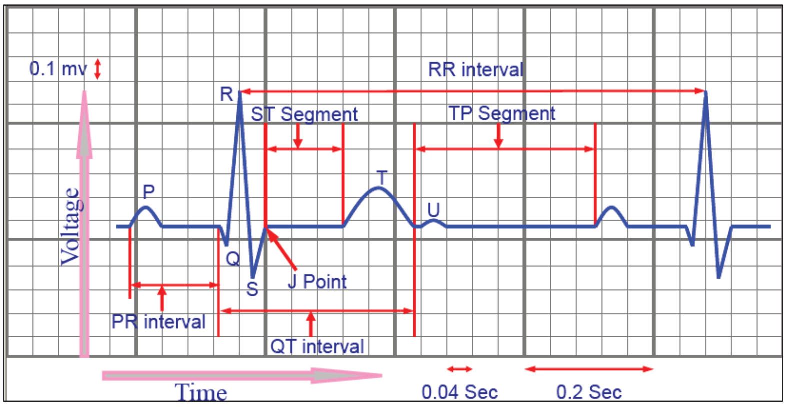

Figure 2 ECG Signal Specifications

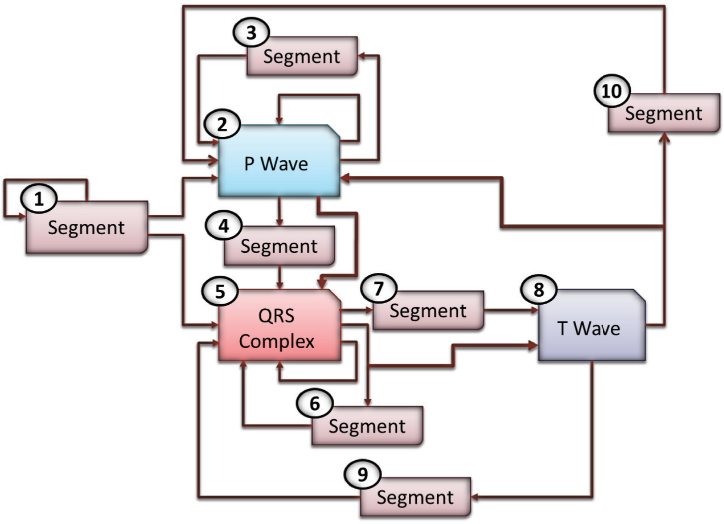

Figure 3 Behavior of ECG Waves

Table 1 State Transitional Probabilities

State Transitions	Probabilities	State Transitions	Probabilities
λ_0	0.01	λ_7	0.165
λ_1	$1-\lambda_0$	λ_8	0.2
λ_2	0.01	λ_9	$1-\lambda_6-\lambda_7-\lambda_8$
λ_3	0.05	λ_{10}	0.25
λ_4	0.70	λ_{11}	0.25
λ_5	$1-\lambda_2-\lambda_3-\lambda_4$	λ_{12}	$1-\lambda_{11}-\lambda_{10}$
λ_6	0.035	λ_f	1
λ_d	0.0		

as in the body. The Markov Chain (MC) of the underlying ECG wave behavior is presented in Figure 4 and its transitional probabilities, which present the events of generating ECG waves, such as P, T, or QRS by the SA and AV nodes, are initially given as mentioned in Table 1. The value λ_f represents the forced transition, with value 1, and λ_d represents dead heart situation and is assumed to be 0.

According to the standard 12-lead ECG representation, as discussed earlier in this section, each behavior of a heart is referred to as the occurrence of a specific pattern repetition within the ECG waveform. For instance, the normal behavior of the heart refers to the occurrence of P, QRS and T waves sequentially, in addition to the occupance of idle flat segments anywhere. For illustration, these sequences are presented in Figure 5 by using different labels, i.e., the dead behavior is labeled as Π, the abnormal behavior is labeled as Δ, the fatigued behavior is labeled as δ, the less fatigued behavior is labeled as Ω, the recovery behavior is labeled as $\diamond$ and the normal behavior is labeled as Σ.

Since, we aim to find the probability of occurrence of these behaviors within the ECG waveform, we define these behaviors in the form of *properties*, which are in turn formalized and verified in PRISM for the CTMC of the ECG. These *properties* are defined using the language of the PRISM model checker. For instance, the probability of occurrence of an abnormal behavior, ehich reflects the occurrence of the SA node within the ECG waveform, is defined as follows:

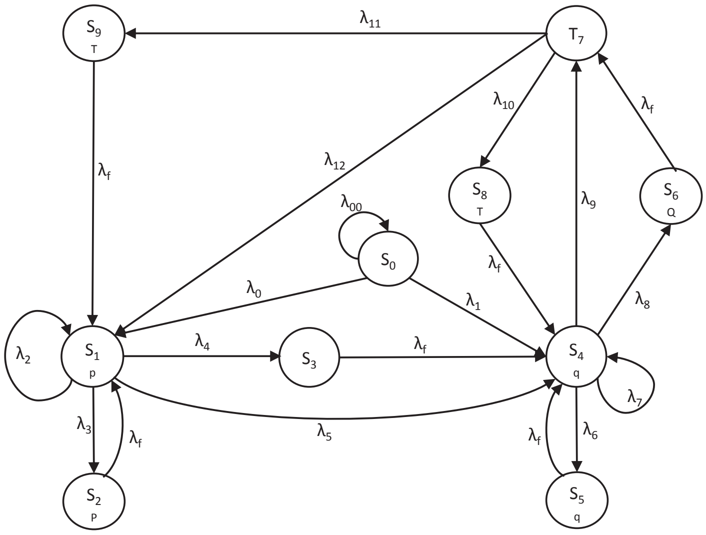

Figure 4 ECG State Machine using CMDP

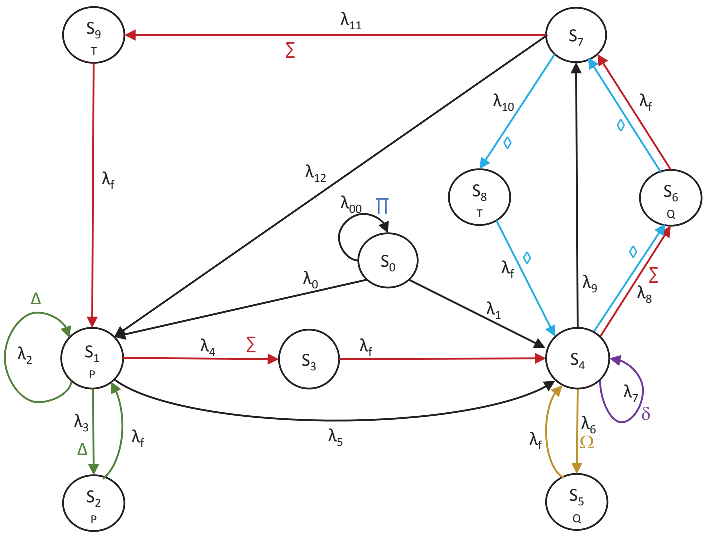

Figure 5 ECG Specifications with MDP

$$P =? [F <= T \ sa\_node >= K] \tag{1}$$

where P represents probability, F refers to future, T is used as a variable to set the future time, sa_node is a variable whose value is updated when the sequence of an abnormal behavior is completed successfully, as presented by the label Δ, in Figure 5, and K is a variable with values ranging from 0 to 100. The state machine, as presented in the Figure 5, is implemented in PRISM and the *property* of each behavior is verified by executing the model for one unit of time and by using different values of K and failure rates.

Figure 6 provides the probability of a given number of times the SA node is in the abnormal state based on initial values, given in Table 1 above.

We ran the experiment for finding the probability of occurrence of K number of sequences of abnormal behaviors for different values of λ_4, which represents the tendency of the heart to generate normal ECG wave, illustrated as Σ in Figure 5 above. The results are shown in Figure 7 for different values of λ_4, given in the percentage form.

Similarly, the other *properties* that present the probability of occurrence of fatigued, less fatigued and normal behavior are as follows:

$$P =? [F <= T \ fatigued = K] \tag{2}$$

$$P =? [F <= T \ less\_fatigued = K] \tag{3}$$

$$P =? [F <= T \ normal = K] \tag{4}$$

The experiment for the probability of occurrence of K number of fatigued behaviors until one unit of time is shown in Figure 8 below. This experiment was conducted for the probabilities, given in the above-mentioned Table. We then conducted an experiment for finding the probability of occurrence of K number of sequences of fatigued behavious until one unit of time for different values of λ_8, which represents the tendency of the heart to continue in the normal ECG wave behavior, illustrated as Σ in Figure 5 above. The results are illustrated in Figure 9 for different values for of λ_8, given in the percentage form.

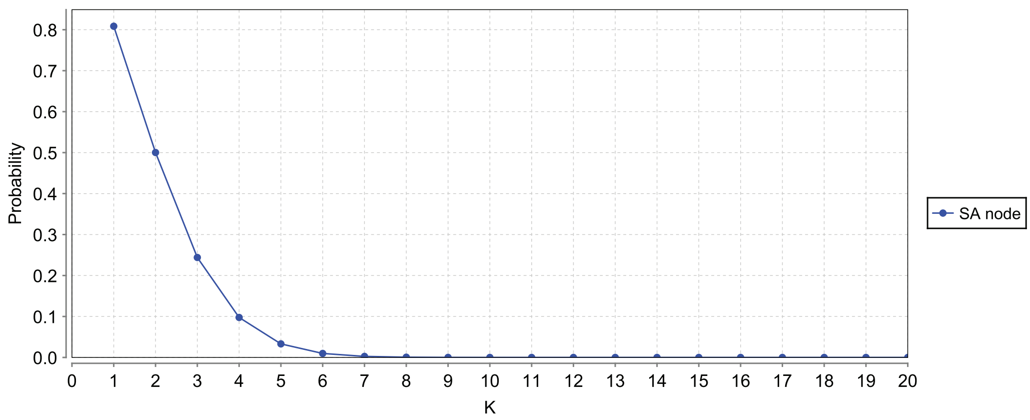

Figure 6 Occurrence of SA node in abnormal state

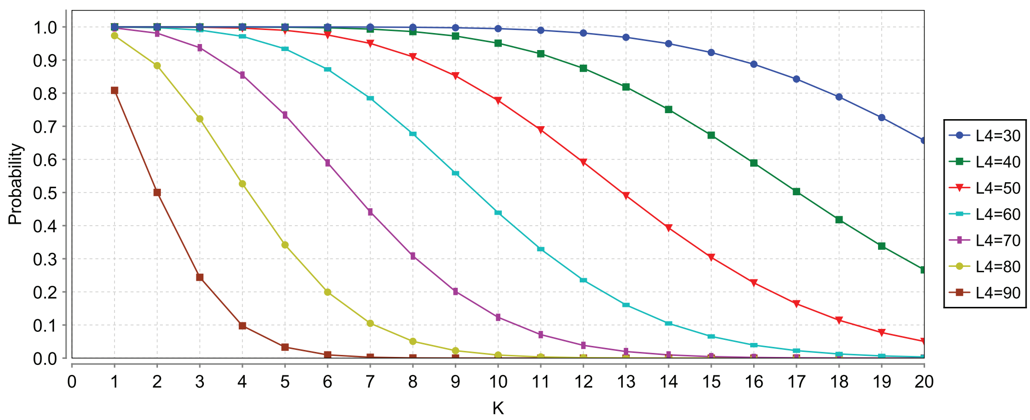

Figure 7 SA node occurrence probability for different values of λ_4

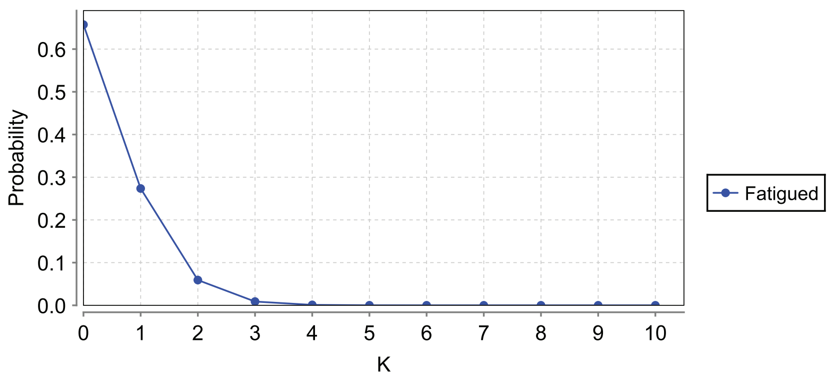

Figure 8 Occurrence probability for fatigued case

The experiment for the probability of occurrence of K number of less fatigued behaviors until one unit of time is shown in Figure 10 below. This experiment was also conducted for the given probabilities in the above Table. We also conducted the experiment for the probability of occurrence of K number of sequences of fatigued behaviors until one unit of time for different values of λ_8. The results are illustrated in Figure 11 for different values for of λ_8, given in the percentage form.

Finally, we conducted the experiment to test the behavior of the heart nodes while in the normal case. The experiment was conducted vs the ability of the heart to go into the recovery state, i.e., generating the QRS wave repetitively. Figure 12 below shows the probability of generating the normal ECG wave, vs different values of λ_{10}, given in the percentage format.

The obtained results can be helpful in identifying tendencies for specific cardialic and abnormal behaviors. It is reported in [6] that the presence of any abnormality, in the presence of an abnormal ECG, increases the probability of different diseases to happen in relatives. Based on this, the findings in this work can be helpful in identifying such probabilities based on any statistical data that can be obtained from the ECG records of the patient. In fact, it is intended that this issue be addressed in the

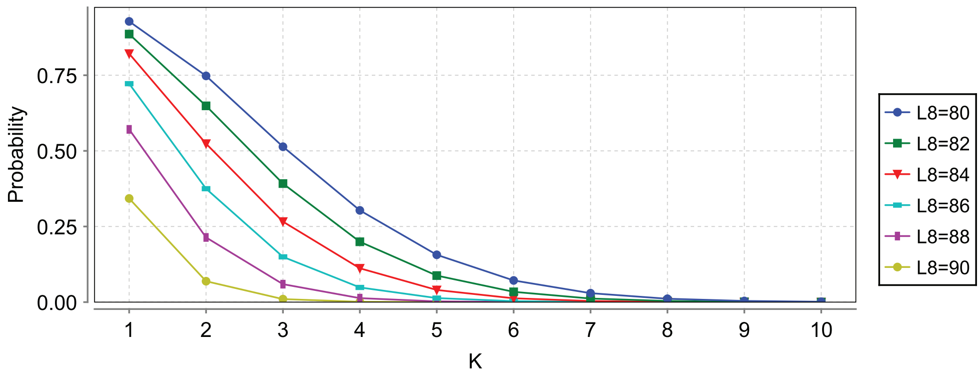

Figure 9 Fatigued vs for different values of λ_8

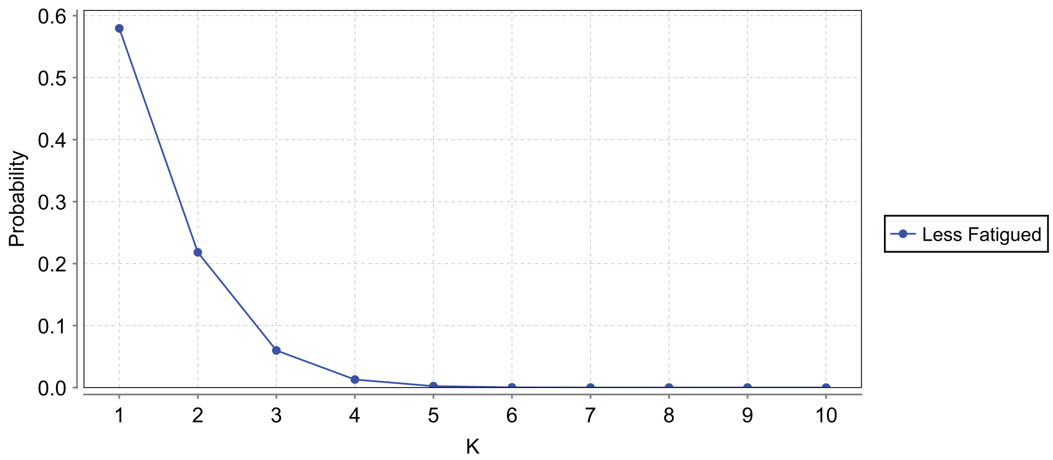

Figure 10 Occurrence probability for less fatigued case

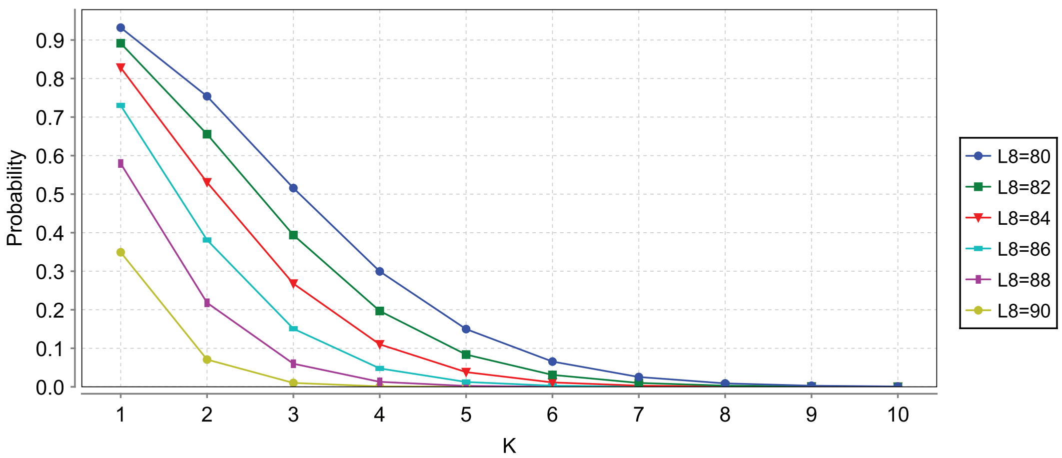

Figure 11 Fatigued vs for less different values of λ_8

next step, where ECG records can be obtained and studied for statistics on abnormalities [10], and then the tendencies for other behaviors can be validated using the method proposed here.

While the results obtained here only reflect a high level view of the cardiac tenancy for specific medical conditions, such as normal or abnormal behavior, this model can be further extended by identifying certain probabilistic measures from the operation of different heart nodes under particular patients conditions, and then, based on this, the model can be useful for predicting specific heart problems. However, this step may require further consultations with cardiologists who can identify these measures.

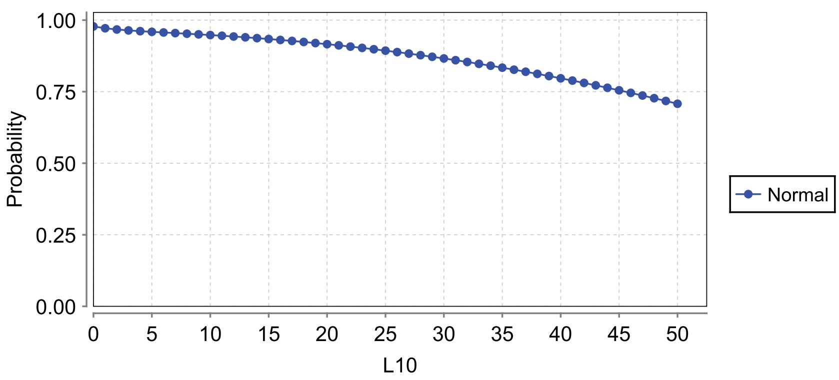

Figure 12 Occurrence probability for normal case vs λ_{10} (Recovery)

5. CONCLUSIONS

This paper presents a formal probabilistic modeling of ECG behaviors and the formal verificaiton, using probabilistic model checking, for analyzing the behavior of ECG heart signals. The first step in this regard us to formalize the specifications of the ECG signals as a CTMC using the language of the PRISM Model checker. This model can then be used to formally verify some of its most desired probabilistic properties, such as the tendency of the heart to generate ECG abnormalities. The results presented in this paper can be useful for the medical analysis and diagnosis based on ECG signals, where tendencies towards specific medical diseases, such as atrial and ventricular arrhythmia, can be observed from ECG abnormalities [6]. Hence, it is essential to have a method that identifies quantitative results on these tendencies based on ECG behaviors. As future work, we are working to obtain statistical results about using ECG signal for diagnosis and then validate ECG abnormalities tendencies using probabilistic model checking. Moreover, we plan to identify correlations between specific deceases and more complex ECG behaviors, and use probabilistic model checking to analyze these correlations.

REFERENCES

1. Jean-Raymond Abrial. Faultless systems: Yes we can! *IEEE Computer Journal*, 42(9):30–36, 2009.
2. H. Al-Hamadi, A. Gawanmeh, and M. Al-Qutayri. Formalizing electrocardiogram (ecg) signal behavior in event-b. In *IEEE 16th International Conference on e-Health Networking, Applications and Services, Healthcom*, pages 55–60, 2015.
3. H. Al-Hamadi, A. Gawanmeh, M. Al-Qutayri, and M. Ismail. A framework for the verification of an ecg biosensor algorithm. *Analog Integrated Circuits and Signal Processing*, 90(3):523–538, 2017.
4. Christel Baier and Joost-Pieter Katoen. *Principles of Model Checking*. The MIT Press, Cambridge, MA, USA, 2008.
5. Bowen J.P. Boca and J. Siddiqi. *Formal methods: state of the art and new directions*. Springer-Verlag, London Limited, 2010.
6. Perry M Elliott, Aris Anastasakis, Michael A Borger, Martin Borggrefe, Franco Cecchi, Philippe Charron, Albert Alain Hagege, Antoine Lafont, Giuseppe Limongelli, Heiko Mahrholdt, et al. 2014 ESC guidelines on diagnosis and management of hypertrophic cardiomyopathy. *European heart journal*, 35(39):2733–2779, 2014.
7. Javier García-Niebla, Pablo Llontop-García, Juan Ignacio Valle-Racero, Guillem Serra-Autonell, Velislav N Batchvarov, and Antonio Bayés De Luna. Technical mistakes during the acquisition of the electrocardiogram. *Annals of Noninvasive Electrocardiology*, 14(4):389–403, 2009.
8. A. Gawanmeh, H. Al-Hamadi, M. Al-Qutayri, S.-K. Chin, and K. Saleem. Reliability analysis of healthcare information systems: State of the art and future directions. In *2015 17th International Conference on E-Health Networking, Application and Services, HealthCom 2015*, pages 68–74, 2015.
9. Daniele Magazzeni Giuseppe Della Penna and Sergio Orefice. Extending visual information extraction to biomedical applications. *Computer Systems Science and Engineering*, 31(5), 2016.
10. AL. Goldberger, LAN. Amaral, L. Glass, JM. Hausdorff, PCh. Ivanov, RG. Mark, JE. Mietus, GB. Moody, C-K. Peng, and HE. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, 2000. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215].
11. Ken Grauer. *A practical guide to ECG interpretation*. Mosby Inc, 1998.
12. J.F. Groote, A. Osaiweran, and J.H. Wesselius. Analyzing the effects of formal methods on the development of industrial control software. In *IEEE Int. Conf. on Software Maintenance*, pages 467–472, 2011.
13. J. Meyer-Kayser H. Hermanns, J. Katoen and M. Siegle. ETMCC: Model Checking performability properties of Markov Chains. In *Dependable Systems and Networks*, pages 673–673, 2003.
14. M. Khattri J. Katoen and I.S. Zapreev. A Markov reward Model Checker. In *Quantitative Evaluation of Systems*, pages 243–244, 2005.
15. R. Jetley, S. Purushothaman Iyer, and P.L. Jones. A formal methods approach to medical device review. *Computer*, 39(4):61–67, 2006.
16. M. Viswanathan K. Sen and G. Agha. Vesta: A statistical Model-Checker and analyzer for probabilistic systems. In *Quantitative Evaluation of Systems*, pages 251–252, 2005.
17. M. Kwiatkowska, H. Lea-Banks, A. Mereacre, and N. Paoletti. Formal modelling and validation of rate-adaptive pacemakers. In *IEEE Int. Conf. on Healthcare Informatics*, pages 23–32, 2014.
18. M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: Verification of Probabilistic Real-time Systems. In *Computer Aided Verification*, volume 6806 of *LNCS*, pages 585–591. Springer, 2011.
19. Marta Kwiatkowska, Gethin Norman, and David Parker. Prism: probabilistic model checking for performance and reliability analysis. *ACM SIGMETRICS Performance Evaluation Review*, 36(4):40–45, 2009.
20. Nancy G. Leveson and Clark S. Turner. An investigation of the

Therac-25 accidents. *Computer Journal*, 26(7):18–41, 1993.

21. Ming Li and Xin Li. Verification based ECG biometrics with cardiac irregular conditions using heartbeat level and segment level information fusion. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pages 3769–3773, May 2014.
22. Jackei Wong, Allan Wong, Tharam Dillon, and Wilfred Lin. A novel approach to achieve real-time tcm (traditional chinese medicine) telemedicine through the use of ontology and clinical intelligence discovery. *Computer Systems Science and Engineering*, 24(4):219–240, 2009.
23. Frank G Yanowitz. Introduction to ECG interpretation, v8.0. *Intermountain Healthcare*, 2012.
24. H. Younes. Ymer: A Statistical Model Checker. In *Computer Aided Verification*, volume 3576 of *LNCS*, pages 429–433. Springer, 2005.