

Research Article

Function Formula Oriented Construction of Bayesian Inference Nets for Diagnosis of Cardiovascular Disease

Booma Devi Sekar and Mingchui Dong

Department of ECE, Faculty of Science & Technology, University of Macau, Avenue Padre Tomas Pereira S.J.S, Macau

Correspondence should be addressed to Booma Devi Sekar; sboomadevi@yahoo.com

Received 6 February 2014; Revised 15 July 2014; Accepted 31 July 2014; Published 27 August 2014

Academic Editor: Aparup Das

Copyright © 2014 B. D. Sekar and M. Dong. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An intelligent cardiovascular disease (CVD) diagnosis system using hemodynamic parameters (HDPs) derived from sphygmogram (SPG) signal is presented to support the emerging patient-centric healthcare models. To replicate clinical approach of diagnosis through a staged decision process, the Bayesian inference nets (BIN) are adapted. New approaches to construct a hierarchical multistage BIN using defined function formulas and a method employing fuzzy logic (FL) technology to quantify inference nodes with dynamic values of statistical parameters are proposed. The suggested methodology is validated by constructing hierarchical Bayesian fuzzy inference nets (HBFIN) to diagnose various heart pathologies from the deduced HDPs. The preliminary diagnostic results show that the proposed methodology has salient validity and effectiveness in the diagnosis of cardiovascular disease.

1. Introduction

Cardiovascular diseases (CVD) are known as the silent killers and often they may develop over time without being noticed until a critical stage is reached. Early diagnosis, care, and continuous monitoring are crucial in preventing heart failures. Thus, exploiting the benefit of multiple technological advancements, research over the past decade has focused on the development of various intelligent tools, to support healthcare professionals and promote CVD self-monitoring. In the same vein, our research team has also been devoted to the research and development (R&D) of e-home healthcare system for CVD self-monitoring [1–5]. One of our key developments is the hemodynamic analysis of sphygmogram (SPG) signal [3], which derives 32 critical vital signs/hemodynamic parameters (HDPs). These HDPs such as cardiac output (CO), stroke volume (SV), systematic arterial compliance (SAC), total peripheral resistance (TPR), and so forth can serve as indices to monitor the health status of cardiovascular system [3, 6, 7].

Though the benefits of SPG and hemodynamic analysis have been well documented [8–11], we believe that an intelligent CVD diagnosis system based on the derived HDPs would benefit e-home healthcare. In this paper, we

propose to apply artificial intelligence (AI) technology to develop such an intelligent CVD diagnosis system. Literature review shows that among various AI technologies, expert system (ES), in particular Bayesian inference nets (BIN), has emerged as one of the most successful intelligent tools in various applications [12–14]. Especially, BIN with its ability to execute staged decision process and provide reasoned conclusions has established a long track record in medical informatics [15–19], leading to the development of various clinical decision support systems (CDSS) [20–24]. To support a doctor's approach of diagnosis with staged decision process, a BIN is adopted in this paper in order to design an intelligent CVD diagnosis system based on HDPs.

However, difficulty arises in constructing the BIN and quantifying the inference nodes to compute the inference through the nets and solve uncertainties. Many renowned researchers including Pearl [25], Chickering [26], Heckerman [27], Friedman and Koller [28], and so forth have contributed towards addressing such bottleneck problem. Some of the key contributions in this regard are worth mentioning.

In constructing a BIN, researchers first developed algorithms that learn the parameters from a large data set to optimally construct the graphical model. These were generally referred to as the learning models and were further

distinguished into search and score based methods [26, 29] and conditional dependence analysis methods [30]. In such methods, although both the graphical topology and the joint probability distribution could be learnt and defined from the data set, there were many shortcomings as implementation of such approach required large amount of qualified data.

The alternative to the data-driven approach was the manual construction of BIN through knowledge acquisition from domain experts using various knowledge elicitation techniques [31]. Though, initially, it was the preferred approach for developing CDSS, it also suffered subsequently from various challenges. It became challenging to systematically analyze the acquired knowledge to construct a hierarchical multistage BIN. Moreover, parameter estimation by different experts faced cognitive biases, often leading to ad hoc estimation of a large number of statistical parameters (e.g., prior probabilities, likelihoods, etc.). Moreover, manual construction required the prior specification of graphical structure between domain variables.

To overcome the critical challenges, some unique benefits from data-driven and knowledge elicitation techniques are availed in this paper and a new approach to construct hierarchical multistage BIN and quantify the inference nodes is proposed. Function formulas in first order predicate logic form are derived to guide in constructing the hierarchical multistage BIN. Further, the FL technology is used to quantify dynamic statistical parameters to inference nodes. The proposed methodology is then applied to construct hierarchical multistage Bayesian fuzzy inference nets (HBFIN) to diagnose various heart pathologies based on HDPs. HBFIN is finally validated using site-measured medical data acquired from two hospitals in China.

2. Hemodynamic Parameters

HDPs derived from hemodynamic analysis of SPG signal can serve as powerful indices for prognosis of CVDs. There are various approaches to hemodynamic analysis [32, 33]. In this paper, hemodynamic analysis is computed based on elastic cavity theory [3], in which the point and area based morphological features of SPG signal as shown in Figure 1 are used to deduce 32 HDPs. The following will show the derivation of some of the important HDPs.

Blood flow continuous equation is

$$Q_{in} = Q_{out} + \frac{dV}{dt_1}, \quad (1)$$

$$Q_{out} + \frac{dV}{dt_2} = 0, \quad (2)$$

where Q_{in} is the volume of blood flowing into the artery and Q_{out} is the volume of blood flowing into the vein. t_1 and t_2 are the systolic and diastolic time period, respectively.

Relation between pressure and blood flow is

$$Q_{out} = \frac{p - p_v}{R}, \quad (3)$$

where p is the arterial pressure, p_v is the venous pressure, and R is the peripheral resistance of cardiovascular system.

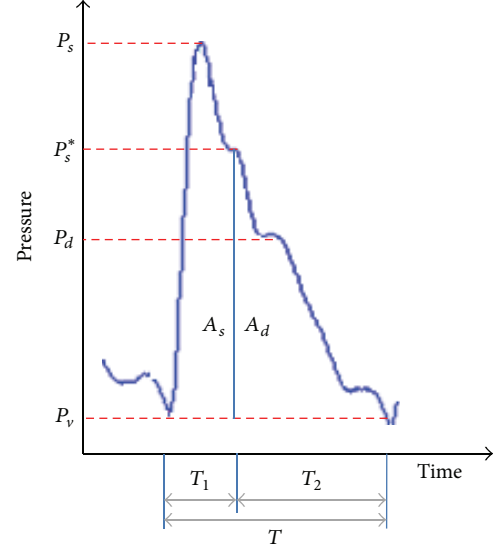


FIGURE 1: Point and area based morphological features of a typical SPG signal.

Arterial pressure-volume equation is

$$AC = \frac{dV}{dp}, \quad (4)$$

where AC is the arterial compliance.

Now, with (1)~(4), the analytic equation of elastic cavity can be computed as follows:

$$Q_{in} = AC \frac{dp}{dt_1} + \frac{p - p_v}{R}, \quad (5)$$

$$AC \frac{dp}{dt_2} + \frac{p - p_v}{R} = 0. \quad (6)$$

Computing the integral of (5) and (6),

$$SV = AC (P_s^* - P_d) + \frac{A_s}{R}, \quad (7)$$

$$AC (P_d - P_s^*) + \frac{A_d}{R} = 0,$$

where SV is the stroke volume and the parameters A_s , A_d , P_s^* , and P_d are the morphological features obtained from SPG as in Figure 1. Thus, the HDPs-SV, AC, R, and so forth can be computed with following equations.

Auxiliary blood pressure index is

$$k = \frac{\int_0^T \bar{P} dt}{T (P_s - P_d)} = \frac{\bar{P} - P_d}{P_s - P_d}. \quad (8)$$

Stroke volume is

$$SV = \frac{0.28}{k^2} T (P_s - P_d). \quad (9)$$

Auxiliary sphygmogram index is

$$\eta = \frac{A_s + A_d}{A_d} = 1 + \frac{A_s}{A_d}. \quad (10)$$

Arterial compliance is

$$AC = \frac{SV}{\eta(P_s - P_d)}. \quad (11)$$

Peripheral resistance is

$$R = \frac{\bar{P} - P_v}{SV} \cdot T \approx \frac{\bar{P}}{SV} \cdot T. \quad (12)$$

Similarly, with morphological features and deduced HDPs, various other HDPs can be generated.

3. Define Function Formulas and Statistical Parameters

3.1. Medical Data Acquisition. The site-measured medical data consists of medical records of different samples, including each patient's physiological attributes, original SPG waveforms, HDPs, and doctor's clinical diagnostic results. Here, the medical symptom space is denoted by $MSS \in R^N$, where $N = 38$, including 6 physiological attributes and 32 HDPs. Totally, 2267 medical records of 165 patients were acquired from two hospitals in China and 850 healthy records were randomly collected. Patient's SPG waveforms and HDPs were measured 12 to 15 times at different time interval within 5 weeks, and their physiological attributes such as age, gender, height, weight, and so forth were also recorded.

Moreover, a medical knowledge base was developed by acquiring information from various medical sources to analyze the relation between the derived HDPs and various pathological conditions of heart. Such medical knowledge base was then verified by doctors from two hospitals in China.

3.2. Define Function Formula in First Order Predicate Logic Form. The key step in constructing inference nets is to define the function formulas in first order predicate logic form using the developed medical knowledge base.

Following equation shows an example of such defined function formula:

$$\forall p \cdot (SP_a(p) \wedge MDP_a(p) \wedge MAP_a(p) \wedge DP_a(p) \rightarrow HT(p)), \quad (13)$$

where p represents patient. SP, MDP, MAP, and DP are the symptoms and HT is the diagnosed hypothesis. Expansion of acronym/abbreviation of the HDPs and CVDs used in function formulas are presented in Figure 3. The suffix "a" represents the value of individual specific condition. The condition values of symptoms for indicating different pathological condition of heart are presented in Table 1.

Deriving from the medical knowledge base, the first order predicate logic formulas for diagnosing various other heart pathologies can be defined as follows:

$$\begin{aligned} & \forall p \cdot (BV_a(p) \rightarrow Low_BV(p)), \\ & \forall p \cdot (MAP_b(p) \wedge SP_b(p) \wedge Low_BV(p) \rightarrow HPT(p)), \\ & \forall p \cdot (PR_a(p) \rightarrow TC(p)), \\ & \forall p \cdot (PR_b(p) \rightarrow BC(p)), \\ & \forall p \cdot (SV_a(p) \wedge SI_a(p) \rightarrow Low_BE(p)), \\ & \forall p \cdot (SV_b(p) \wedge SI_b(p) \rightarrow High_BE(p)), \\ & \forall p \cdot (VPE_a(p) \wedge CI_a(p) \rightarrow Low_CPP(p)), \\ & \forall p \cdot (VPE_b(p) \rightarrow High_CPP(p)), \\ & \forall p \cdot (Y_a(p) \wedge Yr_a(p) \rightarrow HV(p)), \\ & \forall p \cdot (Y_b(p) \wedge Yr_b(p) \rightarrow HPV(p)), \\ & \forall p \cdot (AC_a(p) \wedge FEK_a(p) \wedge Wt_a(p) \wedge BLK_a(p) \\ & \quad \wedge SV_b(p) \wedge HV(p) \rightarrow Dyn_HV(p)). \end{aligned} \quad (14)$$

The defined function formulas can then be used to guide in constructing the hierarchical multistage inference nets to diagnose various CVDs.

3.3. Quantification of Inference Nodes with Dynamic Statistical Parameters. Based on the data distribution, various types of function such as Gaussian, triangle, high-order polynomial, S-type, and so forth can be employed to define the FL membership functions (MF). In the proposed approach, based on the histogram obtained from the frequency plot of observed medical records, high-order polynomial, S-type, or quasi-Gaussian functions are adapted to define the individual MF.

The general formula of i th MF $f_i(s_j^{(0)})$ versus j th symptom in 0th stage $s_j^{(0)}$ expressed in high-order polynomial, S-type, or quasi-Gaussian functions are sequentially listed below:

$$f_i(s_j^{(0)}) = \lambda_0 + \lambda_1 s_j^{(0)} + \lambda_2 (s_j^{(0)})^2 + \dots + \lambda_t (s_j^{(0)})^t, \quad (15)$$

where t is the order of polynomial, λ_0 is random error or noise component, and $\lambda_1, \lambda_2, \dots, \lambda_t$ are coefficients. Consider

$$f_i(s_j^{(0)}) = \frac{1}{1 + e^{-a(s_j^{(0)} - b)}}, \quad (16)$$

where b is the turning point of curve and a is the slope of function. Consider

$$f_i(s_j^{(0)}) = e^{-((s_j^{(0)} - a_i(s_j^{(0)}))/2d_i(s_j^{(0)}))^2}, \quad (17)$$

where $a_i(s_j^{(0)})$ is the maximum membership grade and $2d_i(s_j^{(0)})$ is the bandwidth of that function. Here, the word

TABLE 1: Condition values of symptoms (HDPs) for indicating pathological condition of heart.

Symptoms (units)	Conditions		
	<i>a</i>	<i>b</i>	<i>c</i>
SP (mmHg)	≥ 160	< 90	$= 110 \sim 130$
DP (mmHg)	≥ 95		$= 80 \sim 90$
MAP (mmHg)	> 115	< 65	$= 70 \sim 100$
MDP (mmHg)	> 105		$= 66 \sim 96$
BV (L)	$\leq \{0.75 * Wt * 0.075\}$		$= \{0.75 * Wt * 0.075\} \sim \{1.25 * Wt * 0.075\}$
PR (mmHg)	≥ 104	< 50	$= 60 \sim 100$
Wt (kg)	> 20		$= 50 \sim 80$
SV (mL/stroke)	$\leq \{0.8 * (1 + Q) * 20 * Q\}$	$\geq \{1.3 * 1.2 * (1 + Q) * 20 * Q\}$	$\approx \{(1 + Q) * 20 * Q\}$
SI (mL/stroke/m ²)	$\leq 0.8 * (1 + Q) * 20$	$\geq \{1.3 * 1.2 * (1 + Q) * 20\}$	$\approx (1 + Q) * 20$
VPE (kg/stroke)	$\leq \{0.8 * 2 * (Wt + 45) * 0.0112\}$	$\geq \{1.2 * 2 * (Wt + 45) * 0.0112\}$	$\approx (2 * Wt + 45) * 0.0112$
CI (mL/stroke/m ²)	≥ 2.2		$= \{(1 + Q) * 1.2\} \sim \{(1 + Q) * 2\}$
Y (mpa-s)	$\geq \{1.1 * 4\}$	$\leq \{0.85 * 3\}$	$= 3 \sim 4$
Yr (mpa-s)	$\geq \{1.1 * 4\}$	$\leq \{0.85 * 3\}$	$= 3 \sim 4$
AC ($\mu\text{m}/\text{mmHg}$)	≥ 1.2		≥ 1.2
FEK	$\geq \{0.9 * 0.25\}$		$= 0.35 \sim 0.55$
BLK	$< \{0.85 * 0.22\}$		$= 0.22 \sim 0.26$

*Wt: patient's weight in kg.

*Q: $0.0061 * L \text{ (cm)} + 0.0128 * Wt \text{ (kg)} - 0.1592$.

“quasi” is expressed to indicate that the MF plot would appear as Gaussian distribution shape, but will not cover to the extent of positive and negative infinity.

Based on statistical analysis of site-measured records, the MF for each plot of pair (symptom (HDP) versus membership grade of having specific CVD) is predefined. This therefore fixes all the parameters of (15), (16), or (17) for that particular MF. With an example, Figure 2 illustrates how the MF can be defined from the statistical analysis of site-measured records. Thus, by using such predefined MF, whenever a new patient is tested in the constructed BIN, the statistical parameters can be automatically computed and assigned to relevant symptom node in the inference nets using (15), (16), or (17). It is noteworthy that in this paper a hold-out validation was adopted, whereby for each CVD condition 75% of the acquired data set is used for defining the MF, and the remaining samples are used for validating the constructed BIN.

Now, in HBFIN, when testing data $s_j^{(0)}$ is recorded, it will be substituted in (15), (16), or (17) to compute the relevant $f_i(s_j^{(0)})$. This membership grade dynamically varies according to the value of each symptom, and it approximately reflects the effect of that symptom in diagnosing the CVD. With $f_i(s_j^{(0)})$, dynamic values of statistical parameters are then defined by following rules 1~2 and assigned to the symptom node automatically.

Rule 1. IF $f_i(s_j^{(0)}) \geq 0.2$, then $P_i(s_j^{(0)}) = f_i(s_j^{(0)})$ and $LS_i(s_j^{(0)}) = \alpha * f_i(s_j^{(0)})$.

Rule 2. IF $f_i(s_j^{(0)}) < 0.2$, then $P_i(s_j^{(0)}) = f_i(s_j^{(0)})$ and $LN_i(s_j^{(0)}) = \beta * f_i(s_j^{(0)})$.

The coefficients α and β , in rules 1~2, can be experimentally assigned as zero or positive integer values. While a bigger α would increase the probability of the hypothesis to be true in presence of the evidence, a smaller β would increase the probability of the hypothesis to be false in absence of the evidence. In this paper, the values $\alpha = 100$ and $\beta = 10$ are experimentally chosen. Also, in order to avoid outliers, a threshold of 0.2 is chosen. Then, the statistical parameters for the intermediate hypothesis nodes are defined and assigned using the principle of indifference (PoI) [34], which states that “each member of a set of propositions could be assigned the same probability of truth in the absence of any reason to assign them different probabilities.”

4. Hierarchical Bayesian Fuzzy Inference Nets to Diagnose Cardiovascular Diseases

4.1. Construction of Hierarchical Bayesian Fuzzy Inference Nets. With the function formulas defined in (13)–(14), the symptoms and intermediate and final hypotheses nodes can be identified. Subsequently, the inference nodes can be generated step-by-step and appropriately linked to construct hierarchical multistage inference nets. Figure 3 shows the partially constructed HBFIN to diagnose heart pathologies based on HDPs. In order to clearly illustrate the construction of BIN, the inference nets for healthy condition have not been presented in Figure 3.

It is worth emphasizing here that the partially constructed HBFIN in Figure 3 can be further developed for diagnosing various CVDs.

4.2. Quantifying Inference Nodes of HBFIN with Statistical Parameters. Generally, the inference nodes are quantified

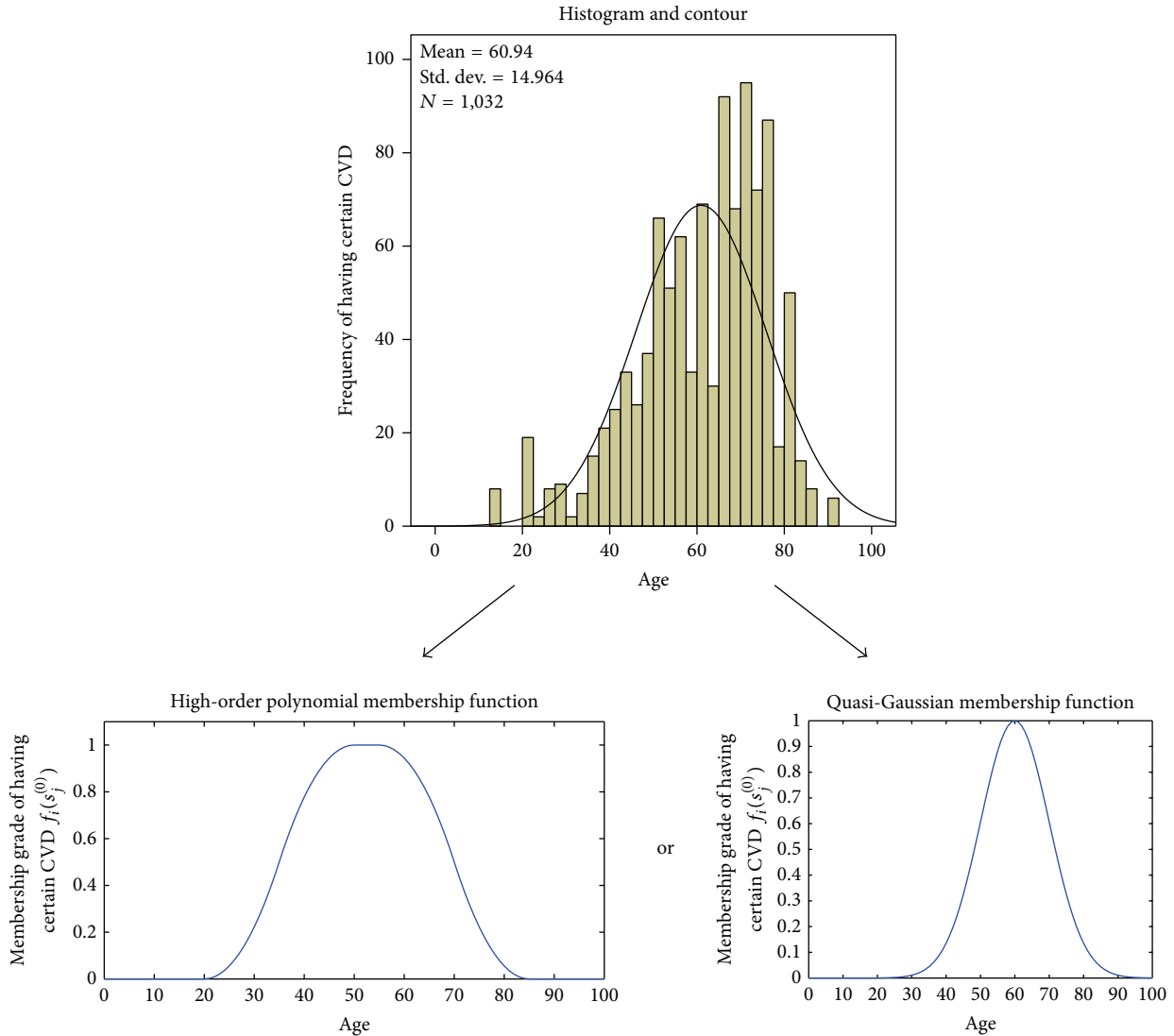


FIGURE 2: Generation of high-order polynomial or quasi-Gaussian membership function for symptom Age versus certain CVD.

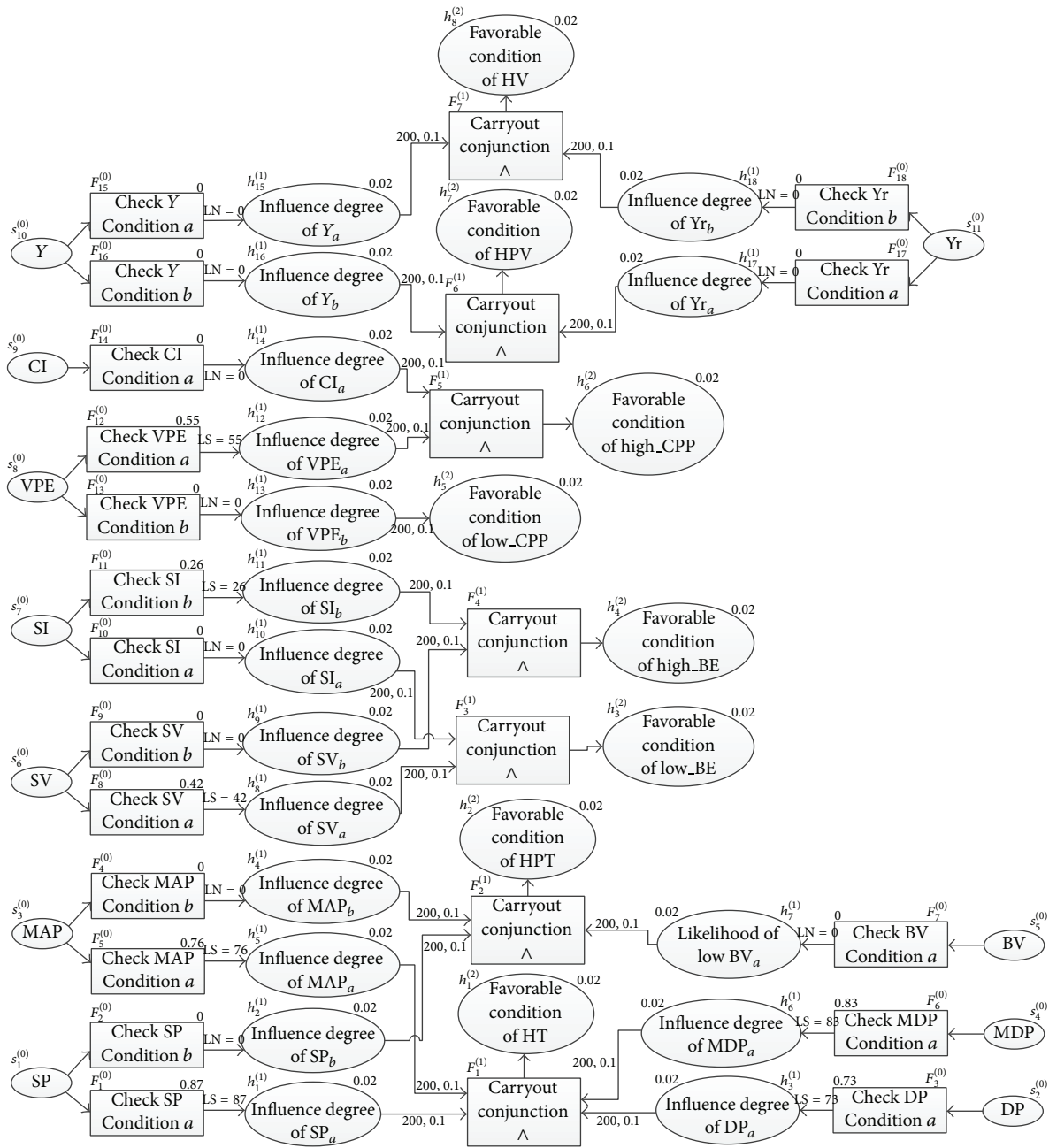
by static values of statistical parameters using subjective (experts' estimation) approach. However, since confliction exists among experts' opinions, defining appropriate static values of statistical parameters to inference nodes has always been a challenge. But, with the proposed methodology using FL technology, dynamic values of statistical parameters can be defined and assigned to inference nodes automatically.

Here, with a specific example, by testing a patient's medical record (partially shown in Table 2) in HBFIN to diagnose CVD, we demonstrate how dynamic values of statistical parameters can be defined/assigned automatically. When the testing data is presented into the symptom nodes of HBFIN, it will be automatically substituted in predefined (15), (16), or (17) to calculate the relevant membership grade $f_i(s_j^{(0)})$, which will then be used in rules 1~2 to define/assign statistical parameters. Figure 3 shows the partially constructed HBFIN with statistical parameters assigned for the sampled testing data shown in Table 2. It is important to note that, with

this approach, the statistical parameters assigned to symptom nodes would dynamically change according to the variation of symptoms. The statistical parameters for the intermediate hypothesis or conclusion nodes are defined and assigned using the PoI. As a result, it can be noted that the intermediate hypothesis nodes are assigned with a prior probability = 0.02, LS = 200, and LN = 0.01, respectively, in HBFIN. For inference nodes executing conjunction and disjunction operations, the statistical parameters would be later calculated from the evidences contributing to these nodes according to Bayesian theory.

5. Evaluation of Constructed Bayesian Fuzzy Inference Nets

5.1. *Mathematical Inference Model Using Bayesian Theory.* The Bayesian inference nets generally form a static knowledge structure, in which the probability associated with each



- | | | |
|------------------------------|--|---------------------------------|
| SP: systolic blood pressure | VPE: left ventricular effective pump power | HPT: hypotension |
| DP: diastolic blood pressure | CI: cardiac index | TC: tachycardia |
| MAP: mean arterial pressure | Y: whole blood viscosity | BC: bradycardia |
| MDP: mean diastolic pressure | Yr: restorer whole blood viscosity | BE: blood ejection |
| BV: blood volume | AC: artery compliance | CPP: cardiac pump power |
| PR: pulse rate | FEK: vascular elasticity divergence ratio | HV: hyperviscosity |
| SV: stroke volume | BLK: aorta evacuate coefficient | HPV: hypoviscosity |
| SI: stroke index | HT: hypertension | Dyn_HV: dynamic hyper viscosity |

FIGURE 3: Partially constructed HBFIN for diagnosing heart pathologies with statistical parameters assigned for a sampled medical record.

TABLE 2: A patient's partial medical record.

Symptoms (units)	Patient's partial medical record
SP (mmHg)	168
DP (mmHg)	100
MAP (mmHg)	130.98
MDP (mmHg)	113.09
BV (L)	3.5212
PR (mmHg)	68
Wt (kg)	49
SV (mL/stroke)	63.81
SI (mL/stroke/m ²)	45.54
VPE (kg/stroke)	2.18
CI (mL/stroke/m ²)	2.7
Y (mpa-s)	3
Yr (mpa-s)	3.8
AC (μm/mmHg)	0.66
FEK	0.11
BLK	0.197

*The expansion of symptom acronym is provided in Figure 3.

inference node consequently changes when the evidence is certain or uncertain. This change in probability is propagated up stage by stage through the hierarchical Bayesian inference nets to ultimately support or disprove the top-level hypothesis/conclusion. In this paper, the following inference model is used to compute the inference through the nets. In this model, for addressing uncertainty in evidence, conditional independence of the evidence is assumed. Therefore, for partially known or uncertain evidence, according to its degree of belief, it is categorized as true or false and the inference through the nets is computed accordingly.

(1) Prior odds of k th hypothesis on q th stage are

$$O(h_k^{(q)}) = \frac{P(h_k^{(q)})}{1 - P(h_k^{(q)})}. \quad (18)$$

$P(h_k^{(q)})$ = prior probability.

(2) Posterior odds of k th hypothesis on q th stage are

$$O(h_k^{(q)} | x \text{ (or) } e) = \left[\prod_{i=1}^N L_i \right] O(h_k^{(q)}). \quad (19)$$

x : evidence is certain;

e : evidence is uncertain;

L_i : LS_i {likelihood of sufficiency}

L_i : LN_i {likelihood of necessity}

L_i : 1 {evidence is unknown}.

(3) Posterior probability of k th hypothesis on q th stage when the evidence is certain is

$$P(h_k^{(q)} | x) = \frac{O(h_k^{(q)} | x)}{1 + O(h_k^{(q)} | x)}. \quad (20)$$

TABLE 3: Diagnostic results of partially constructed HBFIN.

Person's health status	Number of samples	Diagnostic accuracy (%)
HT	53	78
HPT	17	82
Low_BE	13	76
High_BE	17	82
Low_CPP	10	80
High_CPP	18	83
HPV	8	87
HV	13	84

*The expansion of symptom acronym is provided in Figure 3.

(4) Posterior probability of k th hypothesis on q th stage when the evidence is uncertain is as follows:

If $0 \leq P(x | e) < P(x)$, then

$$P(h_k^{(q)} | e) = P(h_k^{(q)} | x') + \frac{P(h_k^{(q)}) - P(h_k^{(q)} | x')}{P(x)} P(x | e); \quad (21)$$

If $P(x) \leq P(x | e) \leq 1$, then

$$P(h_k^{(q)} | e) = P(h_k^{(q)} | x) + \frac{P(h_k^{(q)} | x) - P(h_k^{(q)})}{1 - P(x)} [P(x | e) - P(x)]. \quad (22)$$

For conjunction inference node, $P(x) = \min[P(S_1^{(0)}), \dots, (S_n^{(0)}), P(h_1^{(q-t)}), \dots, P(h_d^{(q-t)}), P(F_k^{(q-t)})]$.

For disjunction inference node, $P(x) = \max[P(S_1^{(0)}), \dots, (S_n^{(0)}), P(h_1^{(q-t)}), \dots, P(h_d^{(q-t)}), P(F_k^{(q-t)})]$.

5.2. *Testing Results.* The function and validity of partially constructed HBFIN are examined using the reserved testing samples. The number of samples used for testing and the obtained diagnostic accuracy are presented in Table 3.

It is noteworthy that the partially constructed HBFIN in Figure 3 was further constructed using function formulas to diagnose various CVDs, such as coronary heart disease (CHD), arrhythmia (AR), pulmonary heart disease (PHD), cerebral infarction (CIN), hyperlipemia (HL), and so forth. Thus, in such inference nets, the hypothesis nodes Low_BE, High_BE, Low_CPP, High_CPP, and so forth became the intermediate hypothesis nodes, which were further linked with other symptoms or hypothesis nodes to diagnose various CVDs. The diagnostic accuracy of the complete HBFIN is presented in Table 4.

Considering that the diagnosis results are derived only from the HDPs and physiologic parameters in the proposed noninvasive approach, the above diagnostic accuracy

TABLE 4: Overall diagnostic accuracy of HBFIN in CVD detection.

Person's health status	Diagnostic accuracy (%)
Healthy	91
HT	78
CHD	68
AR	73
PHD	65
CIN	72
HL	73
Mixed CVD	58

TABLE 5: Diagnostic accuracy of intelligent CVD diagnosis systems using adapted versions of AI technology.

CVD type	Adapted version of AI technology		
	NN	FNN	HBFIN
CHD	78	65	68
HT	70	67	78
HL	64	77	73
Mixed CVD	—	40	58

is highly acceptable and therefore is suitable for ehome healthcare usage.

Furthermore, Table 5 shows the comparison results of intelligent CVD diagnosis systems using adapted versions of AI technologies, including neural networks (NN) [35], fuzzy neural networks (FNN) [36], and the proposed HBFIN. For performing a fair comparison, these methods were developed and validated by our research team with the same medical database used in this paper.

The diagnostic results in Table 5 demonstrate that the proposed HBFIN provides comparable performance *vis-a-vis* other intelligent CVD diagnostic systems. Moreover, it is important to emphasize here that HBFIN could distinctly trace back from the final hypothesis to its initial symptoms and provide reasoned conclusion to the user. Whereas, in using the systems based on NN that are black box in nature, such a feedback cannot be provided.

6. Conclusion

An intelligent CVD diagnosis system based on HDPs derived from SPG signal is presented in this paper. By availing the benefit of some unique features of hybrid AI, BIN, and FL technologies, an intelligent CVD diagnosis system is proposed. A new approach for constructing hierarchical multistage BIN guided by function formulas defined in first order predicate logic form is proposed. A mathematical inference model using Bayesian theory is presented, and a method using FL technology to quantify dynamic values of statistical parameters to inference nodes is suggested. With the proposed methodology, HBFIN is constructed to diagnose various CVDs based on HDPs. The site-measured medical records from two hospitals of China have been used to design and validate the proposed HBFIN. For such

a noninvasive diagnostic approach, the testing results with acceptable diagnostic accuracy in diagnosing six important CVDs prove the suitability of HBFIN for home healthcare usage.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by Research Committee of University of Macau under Grant no. MYRG2014-00060-FST and also by the Science and Technology Development Fund (FDCT) of Macau S.A.R with project reference no. 016/2012/A1.

References

- [1] B. D. Sekar, J. Y. Dou, B. B. Fu, X. L. Fei, and M. C. Dong, "Complexity and similarity analysis of heart sound for cardiovascular disease detection," *Biotechnology*, vol. 10, no. 4, pp. 316–322, 2011.
- [2] B. D. Sekar, C. M. Dong, J. Shi, and X. Y. Hu, "Fused hierarchical neural networks for cardiovascular disease diagnosis," *IEEE Sensors Journal*, vol. 12, no. 3, pp. 644–650, 2012.
- [3] B. N. Li, M. Dong, V. Mang I, and M. P. Un, "A novel intelligent sphygmogram analyzer for health monitoring of cardiovascular system," *Expert Systems with Applications*, vol. 28, no. 4, pp. 693–700, 2005.
- [4] B. N. Li, M. C. Dong, and M. I. Vai, "Modelling cardiovascular physiological signals using adaptive Hermite and wavelet basis functions," *IET Signal Processing*, vol. 4, no. 5, pp. 588–597, 2010.
- [5] W. K. Lei, B. N. Li, M. C. Dong, and M. I. Vai, "AFC-ECG: an adaptive Fuzzy ECG classifier," in *Soft Computing in Industrial Applications*, vol. 39 of *Advances in Soft Computing Series*, pp. 189–199, Springer, Berlin, Germany, 2007.
- [6] B. N. Li, B. B. Fu, and M. C. Dong, "Development of a mobile pulse waveform analyzer for cardiovascular health monitoring," *Computers in Biology and Medicine*, vol. 38, no. 4, pp. 438–445, 2008.
- [7] B. N. Li, M. C. Dong, and M. I. Vai, "On an automatic delineator for arterial blood pressure waveforms," *Biomedical Signal Processing and Control*, vol. 5, no. 1, pp. 76–81, 2010.
- [8] A. Avolio, B. E. Westerhof, M. Siebes, and J. V. Tyberg, "Arterial hemodynamics and wave analysis in the frequency and time domains: an evaluation of the paradigms," *Medical and Biological Engineering and Computing*, vol. 47, no. 2, pp. 107–110, 2009.
- [9] J. J. Xia and S. Liao, "Cardiovascular detecting via pulse analysis," *Engineering*, vol. 5, pp. 176–180, 2013.
- [10] M. F. O'Rourke and A. P. X. Jiang, "Pulse wave analysis," *British Journal of Clinical Pharmacology*, vol. 51, no. 6, pp. 507–522, 2001.
- [11] P. Salvi, *Pulse Waves—How Vascular Hemodynamics Affect Blood Pressure*, Spriger, Dordrecht, The Netherlands, 2012.
- [12] E. Olsson, P. Funk, and N. Xiong, "Fault diagnosis in industry using sensor readings and case-based reasoning," *Journal of Intelligent and Fuzzy Systems*, vol. 15, no. 1, pp. 41–46, 2004.

- [13] R. Chowdhary, J. Zhang, and J. S. Liu, "Bayesian inference of protein-protein interactions from biological literature," *Bioinformatics*, vol. 25, no. 12, pp. 1536–1542, 2009.
- [14] L. Tamine and M. Boughanem, "Inferring document utility via a decision making based retrieval model," *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 14, no. 2, pp. 73–93, 2010.
- [15] A. Djebbar and H. F. Merouani, "Retrieval and adaptation in CBR through Bayesian Network for diagnosis of hepatic pathologies," *International Journal of Hybrid Intelligent Systems*, vol. 9, no. 3, pp. 123–134, 2012.
- [16] D. L. Hudson, M. E. Cohen, and S. E. Hudson, "The use of consequential reasoning in cancer chemotherapy," *Medinfo*, vol. 10, no. 2, pp. 1349–1353, 2001.
- [17] F. Azuaje, W. Dubitzky, P. Lopes et al., "Predicting coronary disease risk based on short-term RR interval measurements: a neural network approach," *Artificial Intelligence in Medicine*, vol. 15, no. 3, pp. 275–297, 1999.
- [18] K. B. Wagholikar and A. W. Deshpande, "Fuzzy relation based modeling for medical diagnostic decision support: case studies," *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 12, no. 5-6, pp. 319–326, 2008.
- [19] P. J. F. Lucas, L. C. Van Der Gaag, and A. Abu-Hanna, "Bayesian networks in biomedicine and health-care," *Artificial Intelligence in Medicine*, vol. 30, no. 3, pp. 201–214, 2004.
- [20] E. H. Shortliffe, *Compute-Based Medical Consultations*, MYCIN, Elsevier, North Holland, New York, NY, USA, 1976.
- [21] E. H. Shortliffe, A. C. Scott, M. B. Bischoff, A. B. Campbell, W. V. Melle, and C. D. Jacobs, "ONCOCIN: an expert system for oncology protocol management," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, pp. 876–881, Vancouver, Canada, 1981.
- [22] S. Andreassen, M. Wolbye, B. Falck, and S. K. Andersen, "MUNIN—A causal probabilistic network for interpretation of electromyographic findings," in *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, J. McDermott, Ed., pp. 366–372, Milan, Italy, August 1987.
- [23] M. A. Shwe, B. Middleton, D. E. Heckerman et al., "Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base. I. The probabilistic model and interference algorithms," *Methods of Information in Medicine*, vol. 30, no. 4, pp. 241–255, 1991.
- [24] W. Long, "Temporal reasoning for diagnosis in a causal probabilistic knowledge base," *Artificial Intelligence in Medicine*, vol. 8, no. 3, pp. 193–215, 1996.
- [25] J. Pearl, "Reverend Bayes on inference engines: a distributed hierarchical approach," in *Proceedings of the 2nd National Conference on Artificial Intelligence*, D. L. Waltz, Ed., pp. 133–136, The AAAI Press, 1982.
- [26] D. M. Chickering, "Optimal structure identification with greedy search," *Journal of Machine Learning Research (JMLR)*, vol. 3, no. 3, pp. 507–554, 2003.
- [27] D. Heckerman, "A Bayesian approach to learning causal networks, Technical report," in *Advances in Decision Analysis: From Foundations to Applications*, W. Edwards, R. F. Milse, and D. Winterfeldt, Eds., pp. 202–220, Cambridge University Press, 2007.
- [28] N. Friedman and D. Koller, "Being Bayesian about network structure. A Bayesian approach to structure discovery in Bayesian networks," *Machine Learning*, vol. 50, no. 1-2, pp. 95–125, 2003.
- [29] L. M. de Campos, "A scoring function for learning Bayesian networks based on mutual information and conditional independence tests," *Journal of Machine Learning Research (JMLR)*, vol. 7, pp. 2149–2187, 2006.
- [30] T. Claassen and T. Heskes, "A logical characterization of constraint-based causal discovery," in *Proceeding of the 27th Conference on Uncertainty in Artificial Intelligence (UAI '11)*, pp. 135–144, July 2011.
- [31] N. J. Cooke, "Knowledge elicitation," in *Handbook of Applied Cognition*, F. T. Duso, Ed., pp. 479–509, John Wiley & Sons, Chichester, UK, 1999.
- [32] V. G. Almeda, J. Vieira, P. Santos et al., "Machine learning techniques for arterial pressure waveform analysis," *Journal of Personalized Medicine*, vol. 3, no. 2, pp. 82–101, 2013.
- [33] Z. R. Liu and X. X. Li, *The Theory and Methods of Hemodynamics*, Press of Fudan University, Shanghai, China, 1997.
- [34] P. Castell, "A consistent restriction of the principle of indifference," *British Journal for the Philosophy of Science*, vol. 49, no. 3, pp. 387–395, 1998.
- [35] J. Shi, B. D. Sekar, M. C. Dong, and W. K. Lei, "Cardiovascular disease risk detection using modified T-S fuzzy neural networks," in *Proceedings of the 2nd International Conference on Computer and Automation Engineering (ICCAE '10)*, pp. 17–20, Singapore, February 2010.
- [36] J. Shi, B. D. Sekar, M. C. Dong, and W. K. Lei, "Fuzzy neural networks to detect cardiovascular diseases hierarchically," in *Proceedings of the 10th IEEE International Conference on Computer and Information Technology (CIT '10)*, pp. 703–708, Bradford West Yorkshire, UK, July 2010.