International Journal of Computer Applications Technology and Research
Volume 9–Issue 06, 186-193, 2020, ISSN:-2319–8656

Denoising Algorithm for Medical Ultrasound Image Based on 2D-VDM and PM

Manyu Yan
School of Communication Engineering
Chengdu University of Information Technology
Chengdu, China

Chengyu Wen
School of Communication Engineering
Chengdu University of Information Technology
Chengdu, China

Abstract: In order to solve the problem of several common methods in medical ultrasound image processing which includes Poor retention of detailed information and insignificant denoising effect, therefore a new method of ultrasonic image denoising combining two-dimensional variational mode decomposition (abbreviated as 2D-VDM) and anisotropic diffusion (abbreviated as PM) is proposed. This method firstly decomposes the image into a series of modal component (IMF) images through two-dimensional variational mode decomposition (2D-VDM),and then uses the peak signal-to-noise ratio and the normalized mean square error to filter out the effective modal components, finally, the effective modal components are subjected to anisotropic diffusion (PM) filter processing and reconstruct the processed effective components to remove image noise.The evaluation of image quality indicators from peak signal-to-noise ratio and root mean square error shows that this method is superior to other commonly used methods in removing noise and protecting detailed information in the image.

Keywords: Ultrasound image; two-dimensional variational modal decomposition; anisotropic diffusion; modal component

1. INTRODUCTION

At present, ultrasound imaging equipment is becoming more and more popular, such as portable remote medical ultrasound scanning instruments. In the ultrasound imaging process, speckle noise will be formed due to the strength and difference of the signal,which causes the image will be unclear. Therefore, the image denoising technology plays a vital role ,which can obtain higher quality and clearer images from the ultrasound equipment, so that the tissues inside the human body can be observed more clearly[1].The current image denoising methods are generally divided into two categories: time domain and frequency domain denoising methods.The denoising process in the time domain mainly uses the corresponding processing function to convolve with the image containing noise to process the pixels in the image.The denoising process in frequency domain mainly decomposes the image into multiple frequencies in different ranges to obtain the corresponding frequency characteriscs.Time domain denoising methods include linear filtering methods (first-order statistical filtering, higher-order statistical filtering), non-linear filtering methods (median filtering, geometric filtering, uniform geometric filtering, etc.). Although these methods have a good effect of removing noise, there is a phenomenon that the edges are over-smoothed, causing loss of detailed information [2].In order to solve the problem that the edges and details of images are easily blurred, researchers have applied partial differential equations to image processing. Therefore Perona and Malik proposed to use the P-M diffusion equation (PM model) as a filtering method for ultrasound images. The algorithm mainly uses a diffusion function to achieve different diffusion strengths in different directions [3].The denoising methods in the frequency domain mainly include wavelet transform, curved wave transform, etc.These methods have the same problem that the effect of removing noise depends largely on the choice of the fundamental wave.Therefore, American NASA scientist HuangNE.Proposed the Empirical Mode Decomposition (EMD) algorithm in 1998, which mainly decomposes one-dimensional signals into different intrinsic modal components (IMF) [4].In 2001,Song Pingjian and others extended the empirical mode decomposition method to the two-dimensional range, that is the two-dimensional empirical mode decomposition (BEMD),because the empirical mode decomposition (EMD) algorithm is prone to aliasing and other problems [5] .In 2014,for the problems of this algorithm,research scholars proposed an improved variational mode decomposition algorithm (VMD) based on empirical mode decomposition (EMD).The algorithm mainly seeks the optimal value of each natural modal component (frequency,bandwidth) by variational method,so as to overcome the shortcomings of empirical mode decomposition (EMD) [6].In addition,VMD has also been extended to a two-dimensional range, that is, two-dimensional variational mode decomposition (2D-VMD).

Two-dimensional variational mode decomposition can decompose the image into low-frequency components and high-frequency components.In addition, anisotropic diffusion filtering is better than other filtering methods in dealing with ultrasonic images in terms of denoising effect and retaining detailed information.This article mainly combines the characteristics of the two methods, and proposes a new method of ultrasonic image denoising that combines two-dimensional variational mode decomposition and anisotropic diffusion.Firstly,the image is decomposed into low-frequency components and high-frequency components through two-dimensional variational mode decomposition, and then anisotropic diffusion filtering is used to filter the low-frequency components, and finally the filtered components are reconstructed.

2. TWO-DIMENSIONAL VARIATIONAL MODE DECOMPOSITION（2D-VMD）

Variational modal decomposition decomposes the signal into a series of IMF components,and iterates to determine the IMF

through the entire decomposition process.Firstly,setting the number of modal components (IMF) to be decomposed in advance and performing Hilbert transform on each model function $u_k(t)$ to obtain its analytical function, and then add an exponential term e^{-jwt} to adjust its center frequency. Secondly, in order to make the decomposed IMF frequency band as close as possible to the center frequency, thus establish a constraint minimization model.Finally,a quadratic penalty term and a Lagrangian multiplier are added to this model to transform its model into an unconstrained model,in addition, iterate its model continuously to find the optimal solution to determine each modal component (IMF).

2.1 Two-dimensional analytical signal

Two-dimensional variational mode decomposition is extended in one-dimensional range.In the one-dimensional variational mode decomposition,the Hilbert transform is mainly performed on each mode,which is used as the imaginary part to obtain its one-dimensional analytical signal $u_k^A(t)$ [7]as in Eq.1.

$$\begin{aligned} u_k^A(t) &= u_k(t) + jH(u_k(t)) \\ &= (\sigma(t) + j/\pi t) * u_k(t) \end{aligned} \quad (1)$$

Fourier transform Eq.1 to obtain one-dimensional analytical signal in frequency domain as in Eq.2.

$$\begin{aligned} u_k^A(w) &= (1+\mathrm{sgn}(w)) \cdot u_k(w) \\ &= \begin{cases} 2u_k(w) & w > 0 \\ u_k(w) & w = 0 \\ 0 & w < 0 \end{cases} \end{aligned} \quad (2)$$

From the above Eq.2.In the one-dimensional domain,the spectrum of the analytic signal has a unilateral characteristic. The existence of this characteristic can adjust the modal spectrum to the corresponding estimated center frequency. For the analytical signal has this characteristic in the two-dimensional range,thus an interface is selected as the center, and a half plane of the interface must be set to 0,where the interface is equivalent to a vector w_k .two-dimensional analytical signal $u_k^{AS}(x)$ as in Eq.3.

$$u_k^{AS}(x) = u_k(x) * (\sigma(\langle x, w_k \rangle) + j/\pi \cdot \langle x, w_k \rangle) \cdot \sigma(w)) \quad (3)$$

Fourier transform Eq.3 to obtain a two-dimensional analytical signal in the frequency domain $u_k^{AS}(w)$ as in Eq.4.

$$\begin{aligned} u_k^{AS}(w) &= (1+\mathrm{sgn}(< w, w_k >)) \cdot u_k(w) \\ &= \begin{cases} 2u_k(w) & \langle w, w_k \rangle > 0 \\ u_k(w) & \langle w, w_k \rangle = 0 \\ 0 & \langle w, w_k \rangle < 0 \end{cases} \end{aligned} \quad (4)$$

2.2 2D-VMD function and solving u_k and w_k

2.2.1 2D-VMD function

The one-dimensional variational mode decomposition properties are used to define the two-dimensional VMD function. Firstly, the two-dimensional analytical signal of each mode is multiplied by the complex exponential term, then, the bandwidth of each mode is calculated as the gradient mode.Specific formula as in Eq.5.

$$\sum_k \left\| \nabla \left[u_k^{AS}(x) e^{-j\langle w_k, x \rangle} \right] \right\|_2^2 \quad (5)$$

The purpose of variational mode decomposition is to make the frequency bandwidth of each mode decomposed as close as possible to its center frequency.Therefore, the above formula is converted into a constrained minimum model as an objective function to evaluate the modal bandwidth,which is two-dimensional VMD function as in Eq.6.

$$\begin{aligned} &\min_{\{u_k\},\{w_k\}} \{ \sum_k \left\| \nabla \left[u_k^{AS}(x) e^{-j\langle w_k, x \rangle} \right] \right\|_2^2 \} \\ &s.t. \quad \forall x : \sum_k u_k(x) = f(x) \end{aligned} \quad (6)$$

In Eq.6: $\{u_k\} = \{u_1, u_2, u_3, \cdots, u_k\}, \{w_k\} = \{w_1, w_2, w_3, \cdots, w_k\}$

In order to solve the problem of transforming constrained into unconstrained models, thus using an operator which is multiplication operator alternating direction method (ADMM).The quadratic penalty term has good convergence in the case of finite weight coefficients, and Lagrange multipliers can achieve good constraints[8].The unconstrained two-dimensional VDM function is obtained as in Eq.7.

$$\begin{aligned} L(\{u_k\},\{w_k\},\lambda) &= \sum_k \alpha_k \left\| \nabla \left[u_k^{AS}(x) e^{-j\langle w_k, x \rangle} \right] \right\|_2^2 + \left\| f(x) - \sum_k u_k(x) \right\|_2^2 \\ &+ \langle \lambda(x), f(x) - \sum_k u_k(x) \rangle \end{aligned} \quad (7)$$

The ADMM is used to convert the constrained model into an unconstrained model, therefore,the solution of the constrained model is converted into a saddle point for solving the unconstrained model, which is the maximum or minimum value in a certain direction.

2.2.2 Solving and Update u_k and w_k

2.2.2.1 Update u_k

For the problem of solving and updating ,which is converted to find the minimum value of the function of Eq.7.The expression is as in Eq.8.

$$\begin{aligned} u_k^{n+1} = \arg\min_{w_k, u_k \in R} \{ &\alpha_k \left\| \nabla [u_k^{AS}(x) e^{-j\langle w_k, x \rangle}] \right\|_2^2 + \\ &\left\| f(x) - \sum_i u_i(x) + \frac{\lambda(x)}{2} \right\|_2^2 \} \end{aligned} \quad (8)$$

In Eq.8: w_k is equal to w_k^{n+1} , $\sum_i u_i(x)$ is equal to $\sum_{i \neq k} u_i^{n+1}(x)$.

According to the convolution properties and Fourier transform, the above Eq.8 is further transformed into the frequency domain as in Eq.9.

$$\hat{u}_k^{n+1} = \arg\min_{\hat{u}_k, w_k \in R}\{\alpha_k \left\| \mathrm{j}w[(1+\mathrm{sgn}(w+w_k))\hat{\mathrm{u}}_k(\mathrm{w}+\mathrm{w}_k)] \right\|_2^2 + \left\| \hat{\mathrm{f}}(\mathrm{w}) - \sum_i^k \hat{\mathrm{u}}_i(w) + \frac{\lambda(w)}{2} \right\|_2^2\} \quad (9)$$

In Eq.9: $\forall w \in \Omega_k$, $\Omega_k = \{w | < w, w_k >\}$ and adjust the first item to $w \to w - w_k$,the above Eq.9 is transformed into the following formula as in Eq.10.

$$\hat{u}_k^{n+1} = \arg\min_{\hat{u}_k, w_k \in R}\{\alpha_k \left\| \mathrm{j}(\mathrm{w}-\mathrm{w}_k)[(1+\mathrm{sgn}(w))\hat{\mathrm{u}}_k(\mathrm{w})] \right\|_2^2 + \left\| \hat{\mathrm{f}}(\mathrm{w}) - \sum_i^k \hat{\mathrm{u}}_i(w) + \frac{\lambda(w)}{2} \right\|_2^2\} \quad (10)$$

To distribute the bandwidth of the analytical signal of the real signal in the positive half plane $w \in \{\langle w, w_k \rangle > 0\}$ .According to Eq.4,adjust the above Eq.10 to the following Eq.11.

$$\hat{u}_k^{n+1} = \arg\min_{\hat{u}_k, \hat{w}_k \in \Omega}\{\int_0^{+\infty} 4\alpha_k (w-w_k)^2 \left|\hat{u}_k(w)\right|^2 + 2\left| \hat{f}(w) - \sum_i^k \hat{u}_i(w) + \frac{\hat{\lambda}(w)}{2} \right|^2 dw\} \quad (11)$$

The meaning of the mathematical form of the above Eq.11 is to find the saddle point $\hat{u}_k$ where the integral in Eq.11 reaches the minimum value in the positive half plane, $\hat{u}_k$ is the result of the (n + 1th) update.The integral term in the above Eq.11 is denoted as P.To make P the minimum value, find the partial derivative of $\hat{u}_k$ in the formula P. When the partial derivative is equal to 0, the value of $\hat{u}_k$ is the result.as in Eq.12.

$$\frac{\partial p}{\partial \hat{u}_k(w)} = 8\alpha_k (w-w_k)^2 \hat{u}_k(w) - 4\left| \hat{f}(w) - \sum_i^k \hat{u}_i(w) + \frac{\hat{\lambda}(w)}{2} \right| = 0 \quad (12)$$

According to Eq.12,the result of $\hat{u}_k(w)$ is $\hat{\mathrm{u}}_k^{n+1}(w)$,as in Eq.13.

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w) + \frac{\hat{\lambda}(w)}{2}}{1 + 2\alpha_k (w-w_k)^2} \quad (13)$$

The real part after inverse Fourier transform of Eq.13 is the corresponding IMF component u_k .

2.2.2.2 Update w_k

Regarding how to update and solve w_k , make the main center frequency appear in the bandwidth. the objective function is as in Eq.14.

$$w_k^{n+1} = \arg\min_{w_k, u_k \in R}\{\alpha_k \left\| \nabla[u_k^{AS}(x) e^{-j<w_k, x>}] \right\|_2^2\} \quad (14)$$

According to the convolution and Fourier transform and Eq.8,to make the frequency band of each IMF as close as possible to the center frequency, the above Eq.14 is transformed into the frequency domain and the expression is as in Eq.15.

$$w_k^{n+1} = \arg\min_{w_k}\{\int_\Omega (w-w_k)^2 \left|\hat{u}_k(w)\right|^2 dw\} \quad (15)$$

Let the partial derivative of the integral part of Eq.15 for w_k equal to zero,get the value of w_k as the result of its (n + 1th) update .as in Eq.16

$$\frac{\partial\{\int_\Omega (w-w_k)^2 \left|\hat{u}_k(w)\right|^2 dw\}}{\partial w_k} = -\int_\Omega 2(w-w_k)\left|\hat{u}_k(w)\right|^2 dw = 0 \quad (16)$$

According to Eq.16, w_k^{n+1} can be obtained as in Eq.17.

$$w_k^{n+1} = \frac{\int_\Omega w\left|\hat{u}_k(w)\right|^2 dw}{\int_\Omega \left|\hat{u}_k(w)\right|^2 dw} \quad (17)$$

2.3 The specific process of 2D-VMD algorithm

Step 1: Initialized the values of $\{\hat{\mathrm{u}}_\mathrm{k}^0\}, \{\hat{\mathrm{w}}_\mathrm{k}^0\}, \hat{\lambda}_k^0, n,\ k$.

Step2:According to Eq.13, update u_k within a certain range.

Step3:According to Eq.17, update w_k within a certain range.

Step4:Update λ , which is $\hat{\lambda}^{n+1}(w) = \hat{\lambda}^n(w) + \tau(\hat{f}(w) - \sum_k \hat{u}_k^{n+1}(w))$.

Step 5: Repeat the above steps continuously, until the number of cycles is greater than n and

$\sum_k \frac{\left\|\hat{u}_k^{n+1} - \hat{u}_k^n\right\|_2^2}{\left\|\hat{u}_k^n\right\|_2^2} < \kappa e$,then stop iterating.

3. ANISOTROPIC DIFFUSION ALGORITHM (PM ALGORITHM)

In this paper, the PM model (Perona-Malik model) is used to remove the noise of the effective component IMF image decomposed by the 2D-VDM algorithm.The model mainly proposes to replace the diffusion coefficient d with the functional diffusion coefficient on the classical anisotropic diffusion equation[9].The PM model is as in Eq.18.

$$\begin{cases} \frac{\partial g_{i,j,t}}{\partial i} = div[f(|\nabla g_{i,j,t}|)\nabla g_{i,j,t}] \\ \frac{\partial g_{i,j,t}}{\partial i} = [\frac{\partial f(|\nabla g_{i,j,t}|)}{\partial i}\frac{\partial g_{i,j,t}}{\partial i}] + [\frac{\partial f(|\nabla g_{i,j,t}|)}{\partial j}\frac{\partial g_{i,j,t}}{\partial j}] \\ g(i,j,0) = g_0(i,j) \end{cases} \quad (18)$$

In Eq.18, $f(|\nabla g_{i,j,t}|)$ is diffusion coefficient function, ∇ is gradient , div is divergence, $g_0(i,j)$ is a starting pixel, (i,j) is pixel position in the image, t is Time interval. In the PM model, Perona and Malik proposed two functions that meet the following conditions as in Eq.19,Eq.20.

$$f_1(|\nabla g_{i,j,t}|) = \exp(-(|\nabla g_{i,j,t}| / k)^2) \quad (19)$$

$$f_2(|\nabla g_{i,j,t}|) = \frac{1}{1+(\nabla g_{i,j,t} / k)^2} \quad (20)$$

The above medium diffusion coefficient function $f(|\nabla g_{i,j,t}|)$ should have the following characteristics:

(1)The function $f(|\nabla g_{i,j,t}|)$ takes $|\nabla g_{i,j,t}|$ as a variable and decreases as the variable rises.

(2)when $|\nabla g_{i,j,t}| \to 0$, $f(|\nabla g_{i,j,t}|) = 1$.

(3)When $|\nabla g_{i,j,t}| \to 1$, $f(|\nabla g_{i,j,t}|) = 0$.

In Eq.19,Eq.20,Constant k is the gradient threshold.Judging the size relationship between k and $|\nabla g_{i,j,t}|$ is mainly used to diffuse a certain range in the image.Discrete Eq.18 as in Eq.21.

$$\frac{\partial g_{i,j,t}}{\partial t} = \lambda\{d_{i+1,j,t}\nabla_N g_{i,j} + d_{i-1,j,t}\nabla_S g_{i,j} + d_{i,j+1,t}\nabla_W g_{i,j} + d_{i,j-1,t}\nabla_E g_{i,j}\} \quad (21)$$

In Eq.21, $d_{i+1,j,t}, d_{i-1,j,t}, d_{i,j+1,t}, d_{i,j-1,t}$ are the diffusion coefficient values of north, south, east and west respectively and $\nabla_N g_{i,j}, \nabla_S g_{i,j}, \nabla_W g_{i,j}, \nabla_E g_{i,j}$ represent the upward gradient of north, south, east and west respectively, λ is a measure of the rate of diffusion[10].Then the new pixel value $f_{i,j}$ after processing is as in Eq.22.

$$f_{i,j} = g_{i,j} + \frac{\partial g_{i,j,t}}{\partial t} \quad (22)$$

In summary, the PM diffusion model can be obtained as in Fig.1

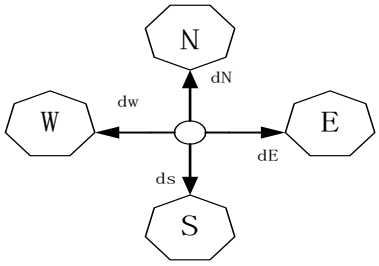

Figure 1. PM diffusion model

4. SPECIFIC STEPS OF MEDICAL IMAGE DENOISING ALGORITHM BASED ON 2D-VDM AND PM

To remove noise and retain the original image features and details to the greatest extent, therefore,this paper proposes an ultrasonic image denoising algorithm based on a combination of two-dimensional variational mode decomposition (2D-VMD) and anisotropic diffusion algorithm (PM).This method firstly decomposes the image into a series of modal component (IMF) images through two-dimensional variational mode decomposition (2D-VDM),and then uses the peak signal-to-noise ratio and the normalized mean square error to filter out the effective modal components, finally, the effective modal components are subjected to anisotropic diffusion (PM) filter processing and reconstruct the processed effective components to remove image noise.Specific steps are as follows:

Step 1:Firstly read the Gaussian speckle noise image and grayscale it to $g(x,y)$, and then initialize the values of $\{\hat{u}_k^0\}, \{\hat{w}_k^0\}, \hat{\lambda}_k^0, n,\ k$ respectively.

Step 2:Through the 2D-VDM algorithm, the processed grayscale image $g(x,y)$ is decomposed into k modal component IMF images $img_k(x,y)$.

Step3:By evaluating the index coefficients, the k-model component IMF images $img_k(x,y)$ after decomposition are screened to select effective model image I components.

Step 4: The effective IMF image components are subjected to anisotropic diffusion (PM) processing to obtain the effective modal components after denoising.

Step 5: Reconstruct the filtered effective IMF component image, and the reconstructed image is the denoised image $\hat{g}(x,y)$.

5. EXPERIMENTAL EVALUATION METHOD AND RESULT ANALYSIS

5.1 Experimental evaluation method

For evaluating the effectiveness and performance of the filtering algorithm, it is mainly measured by the two aspects of denoising and protection of detailed information.Therefore, the peak signal-to-noise ratio (PSNR) and root-mean-square error (RMSE) are used to determine the ability to remove noise and retain detailed information respectively[11].

(1) peak signal-to-noise ratio (PSNR)

The peak signal-to-noise ratio is the main evaluation index to measure the filtering algorithm's ability to remove image noise.This indicator counts the changes in the image signal-to-noise ratio, so the value of this indicator (PSNR)is the larger,the filtering algorithm's ability to remove noise is the stronger [12].Its definition is as in Eq.23.

$$PSNR = 10\log\frac{M_{\max}(i,j)\times M\times N}{\sum_{i=1}^{M}\sum_{j=1}^{N}[g(i,j)-f(i,j)]^2} \qquad (23)$$

In Eq.23, $M_{\max}(i,j)$ is The maximum pixel value of the image after grayscale, M,N are rows and columns respectively, $g(i,j)$ is the pixel value of the image after noise removal, $f(i,j)$ is the pixel value of the initial image, $M\times N$ is the sum of image pixels.

(2) root-mean-square error (RMSE)

Decide the quality of the image filtering algorithm, an aspect of detail information that needs to be considered.therefore,the root mean square error (RMSE) is selected as the main index for evaluating the image quality.This index mainly describes the similarity of the original image and the denoised image in pixels.If the value is smaller, the higher the degree of similarity is the higher and the ability to maintain details is the stronger[13].Its definition is as in Eq.24.

$$RMSE = \sqrt{\frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}[g(i,j)-f(i,j)]^2} \qquad (24)$$

In Eq.24, $g(i,j)$ Is the pixel value of the image after noise removal, $M\times N$ is the sum of image pixels, $f(i,j)$ is the pixel value of the initial image.

5.2 Result analysis

To verify whether the method in this paper can better remove noise while retaining detailed information,three groups of experiments are mainly carried out,which selects two groups of ultrasound fetal and ultrasound kidney images by Field tool simulation and a set of standard images [14].In the actual ultrasonic reflected wave imaging process, it will be interfered by various waves,which results in speckle noise on the ultrasound image with different brightness levels.In addition,the noise is random and its phase distribution and probability density function obey Gaussian distribution.Due to the uncertainty of noise in the ultrasound imaging process, the Gaussian noise variances of 0.02 and 0.05 were added to the three groups of images, respectively, in order to verify the effectiveness of the proposed algorithm in removing the different noise intensity.In addition, Wiener filtering, median filtering, PM filtering, 2D-VDM combined with median filtering and the method in this paper are applied to ultrasound images,finally,the denoising performance of various methods is evaluated by comparing the two coefficients of peak signal-to-noise ratio and root mean square error.

For better filtering processing later, the peak signal-to-noise ratio and the root mean square error are used as evaluation indicators to filter out the effective IMF image components in the ultrasound image after two-dimensional variational mode decomposition.Therefore,the noise variance of the three groups of images is selected as 0.05, and the effective IMF components are selected by comparing the evaluation indexes of the modal components after decomposition.In the two-dimensional variational mode decomposition, the penalty parameter is $\alpha = 5000$ and the mode number is $k = 4$ .Fig.2、Fig.3 and Fig.4 respectively show the decomposed model components of the ultrasound kidney image, ultrasound fetal image, and standard image.

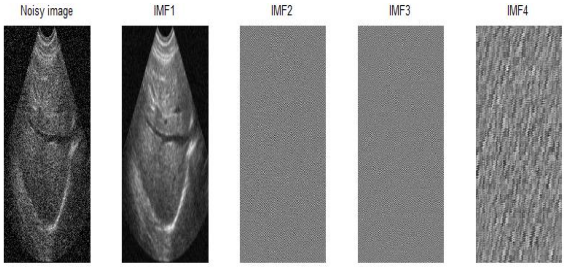

Figure 2. Decomposition picture of ultrasound kidney

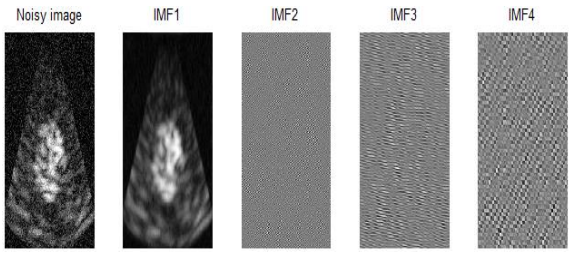

Figure 3. Decomposition picture of ultrasound fetus

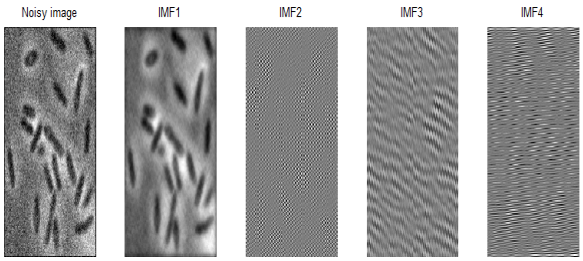

Figure 4. Decomposition picture of standard image

Fig.2、Fig.3and Fig.4 respectively show that three groups of images are decomposed to obtain four different modal components, and each component represents the information characteristics of different frequency bands in the original image.From the three groups of image decomposition diagrams,the first model component IMF1 can well retain the information in the original image and has less noise components,the IMF2 component has partial contour details of the original image but more noise components,IMF3 and IMF4 contain almost no information features of the original image and the noise component is quite large.In order to filter out the effective components, the peak signal-to-noise ratio and the root mean square error of the model components in the three groups of images are compared, as in Table.1.

Table.1 Comparison of IMF component indexes

Image	index	IMF1	IMF2	IMF3	IMF4
Ultras-ound kidney	PSNR	28.413	10.666	10.666	10.572
	RMSE	10.714	62.669	61.510	62.740
Ultras-ound fetus	PSNR	25.483	11.971	12.023	11.516
	RMSE	18.165	66.126	64.110	68.130
Standa-rd image	PSNR	23.041	6.1540	5.8820	5.9676
	RMSE	21.656	116.59	118.88	117.72

According to the data in Table 1, the peak signal-to-noise ratio and root mean square error of the modal components in the three groups of ultrasound kidney, standard image, and ultrasound fetal images are relatively different.In terms of peak signal-to-noise ratio (PSNR), the difference between the first modal component IMF1 and other modal components is about 10 to 20,in addition,the difference between the first modal component IMF1 and other modal components is about 50 ~ 70 in root mean square error (RMSE).More importantly, the peak signal-to-noise ratio of the first modal component IMF1 is the highest and the root mean square error is the smallest,in addition,the peak signal-to-noise of other modal components IMF2, IMF3, and IMF4 is relatively low, the root mean square error is relatively high, and the difference in values between them is small.Therefore, the decomposed first modal components of the three groups of images can better retain the information characteristics of the original image and have less noise components, so the first model component is determined as an effective component.The noise components of other model components IMF2, IMF3, and IMF4 are higher, which leads to the loss of the information characteristics of the original image, and then the other modal components are determined as noise components.By comparing the modal components of the three groups of ultrasound kidney, standard image and ultrasound fetal image,the first component IMF1 of the three groups of images is used as an effective modal component but its modal component still contains less noise components.Thus, anisotropic diffusion filtering is performed on its effective components,finally,reconstructing the filtered effective component is the denoised image.

In the case of noise with different variances (0.02、0.05),the effective modal components of the three groups of images decomposed by 2D-VDM are filtered by five methods.The two indicators of peak signal-to-noise ratio and root mean square error are used to measure the denoising performance of the five methods. The comparison of the PSNR values of the five methods is as in Table.2.

Table.2 PSNR Compare

Image	PSNR					
	variance	Wiener filter	Median Filter	2D-VDM+Median Filter	PM Filter	article method
Ultrasound kidney	0.02	26.85	27.89	28.96	22.49	28.53
	0.05	23.45	23.82	27.44	22.46	28.83
Ultrasound fetus	0.02	28.15	28.29	27.48	23.80	28.72
	0.05	24.77	24.31	26.30	22.56	27.99
Standard image	0.02	25.61	25.40	27.64	24.75	28.70
	0.05	24.52	24.64	26.82	25.00	27.99

According to the data in Table 2,Compared with other methods, this method has the highest peak signal-to-noise ratio and is about 1 ~ 0.2 higher than other methods, indicating that this method is superior to other methods in removing noise.From different noise densities, as the noise density increases, the method in this paper has less volatility in peak signal-to-noise ratio than other methods. In other words, the noise removal performance of other methods is greatly affected by the image noise density.In order to

measure the performance of each method in protecting details, the RMSE values of the five methods are as in Table 3.

Table.3 RMSE Compare

Image		RMSE				
	variance	Wiener filter	MedianFilter	2D-VDM+ Median Filter	PM Filter	Article method
Ultrasound kidney	0.02	13.46	12.02	10.90	14.11	10.56
	0.05	17.96	17.01	16.90	17.54	16.01
Ultrasound fetus	0.02	13.936	12.6182	12.9081	13.2700	12.5247
	0.05	18.405	18.425	18.3858	18.8538	18.1048
Standard image	0.02	16.956	16.598	13.1058	15.6441	12.0709
	0.05	17.738	15.5005	14.8314	15.3772	13.539

According to the data in Table 3,In terms of root mean square error (RMSE),compared with other methods, this method has the smallest root-mean-square error and the difference between other methods is about 1.1 ~ 0.2,which shows that this method is superior to other methods in protecting detailed information. According to the analysis of Table 2 and Table 3,in the case of different variance noise, the method in this paper is superior to other methods in terms of peak signal-to-noise ratio and root mean square error, indicating that the method in this paper can improve the denoising effect while retaining detailed information.

In order to better compare the visual effects of each method, Fig.5、Fig.6 and Fig.7 respectively show the effect of each method on denoising the ultrasound kidney, ultrasound fetus and standard image under the noise variance of 0.02.

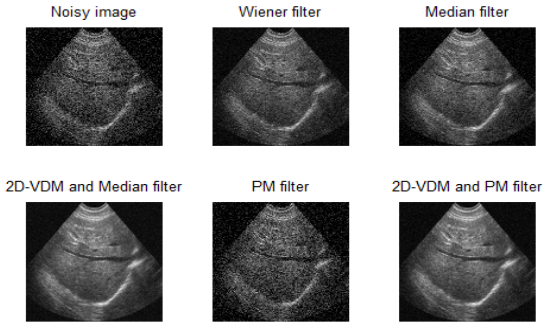

Figure 5. Comparison results of various methods

of ultrasound kidney

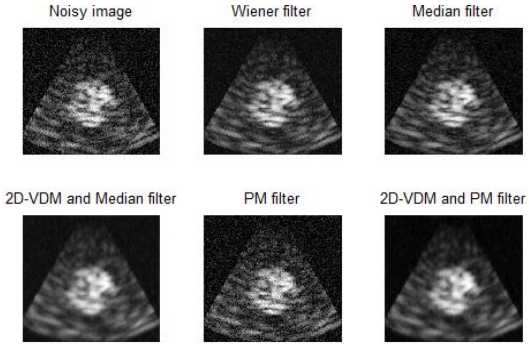

Figure 6. Comparison results of various methods

of ultrasound fetus

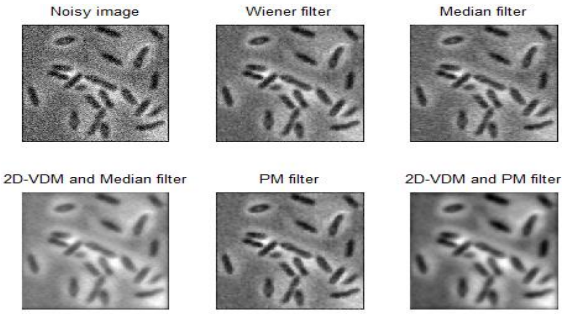

Figure 7. Comparison results of various methods

of standard charts

According to Fig. 5、Fig.6 and Fig.7,this paper applies a combination of two-dimensional variational mode decomposition and anisotropic diffusion to ultrasound images, and the results show that its denoising effect is better than other methods.Two-dimensional variational mode decomposition can decompose its image into low-frequency components and high-frequency components.in addition,for processing ultrasound images,anisotropic diffusion filtering is better than other filtering methods in terms of denoising effect and retaining detailed information.Therefore,this paper combines the characteristics of its two methods.

6. ACKNOWLEDGMENTS

.For the problem of how to suppress the noise of the ultrasound image and protect its detailed information, this paper mainly applies the combination of two-dimensional variational mode decomposition and anisotropic diffusion to the ultrasound image.Two-dimensional variational mode decomposition can decompose its image into low-frequency components and high-frequency components.in addition,for processing ultrasound images,anisotropic diffusion filtering is better than other filtering methods in terms of denoising effect and retaining detailed information.Therefore,this paper combines the characteristics of its two methods.In addition, the peak signal-to-noise ratio and root mean square error are mainly selected as the performance indicators of the algorithm.From the experimental results, the algorithm in this paper is superior to other algorithms in removing noise in the ultrasound image and retaining the original details.The

method in this paper works well in removing noise, but there are some shortcomings, such as the slower decomposition speed of 2D modal decomposition and the selection of the appropriate number of modal decomposition.

7. REFERENCES

[1] Lu Zhaoling.Development trend of medical ultrasound diagnosis [J] .Applied Acoustics,2008,27(4):245─24

[2] Christos P. Loizou, Constantinos S. Pattichis.Despeckle Filtering Algorithms and Software for Ultrasound Imaging [M]. Morgan & Claypool Publishers series,2008:45-74.

[3] Wang Yaqiang, Chen Bo. An improved anisotropic diffusion ultrasonic image denoising algorithm [J]. Liquid Crystal and Display,2015,30(2):30-36

[4] He Peipei. Image Denoising Based on BEMD [J]. Computer Simulation,2009, 26(1):216-218.

[5] DragomiretskiyK,ZossoD. Variational Mode Decomposition[J]. IEEE Transactions on Signal Processing,2014,62(3):531-544.

[6] Dragomiretskiy K，Zosso D. Two-dimensional variational mode de-composition［J]. Springer International Publishing，2015,65(3): 197-208.

[7] Li Xiang, Wang Wenbo. De-noising method of hyperspectral image based on two-dimensional empirical mode decomposition [J]. Laser and infrared,2013,43(11):1311-1315.

[8] Leng Chengcai, Zhao Fengqun, Dai Fang.Improvement of P-M model diffusion function [J]. Computer and Modernization,2008,40(1):19-20.

[9] Wang Yaqiang, Chen Bo et al. An improved anisotropic diffusion ultrasonic image denoising algorithm [J]. Liquid Crystal and Display,2015,30(2):30-36.

[10] Pang Jianxin. Research on objective evaluation of image quality [D]. PhD thesis, University of Science and Technology of China, 2008.

[11] M. Szczypinski, M. Strzelecki, A. Materka, A. Klepaczko,MaZda . a software package for image texture analysis,Comp[J]. Methods Prog. Biomed,2008 ,94 (1):66–76

[12] Christos P. Loizou，Constantinos S. Pattichis. Comparative Evaluation of Despeckle Filtering In Ultrasound Imaging of the Carotid Artery[J]. IEEE transactions on ultrasonics, ferroelectrics, and frequency control,2005,52(10):53- 64

[13] Zhou Guirong, Li Delai. Design Simulation of Ultrasonic Beam Formation Using Field-II [J]. China Medical Device Information Press,2015.36(3):31-38.

[14] Shi Yazheng. Simulation and Realization of Ultrasound Image Processing [D]. Master Thesis of University of Electronic Science and Technology of China,2010.

[15] Liu Jiamin, Peng Ling, Yuan Jiacheng, Liu Junwei. Image denoising method based on two-dimensional variational mode decomposition and adaptive median filtering [J]. Computer Application Research,2017.36(10):32-37.

[16] Chang Qiuhan.Application of Variational Modal Decomposition in Medical Images [D] .Master Thesis of Northeast Petroleum University,2018.

International Journal of Computer Applications Technology and Research
Volume 9–Issue 06, 194-199, 2020, ISSN:-2319–8656

Use of Hybrid Data Mining in Identification of Crime Patterns and Trends in the Matatu Industry in Kenya

Duncan Nyale
Faculty of
Information Science
and Technology
Kisii University
Kisii, Kenya

Samuel Liyala
Faculty of
Information Science
and Technology
Kisii University
Kisii, Kenya

James Ogalo
Faculty of
Information Science
and Technology
Kisii University
Kisii, Kenya

Michael Kangethe
School of Computing
and Informatics
Gretsa Univesity
Thika, Kenya

Abstract: The aim of this study was to propose an automatable technological framework that identifies crime and misconduct patterns and trends in the matatu industry using data mining techniques for intelligence led policing in Kenya. The objectives of the study include to propose a framework for intelligent transport management system with patterns and trends identification capabilities, enhance formulation of policy developments, implementations and government regulations for the transport sector in Kenya, design model system for testing the framework to ascertain its practicability and effectiveness and identify challenges of the transport sector in Kenya. This was an application research which made use of dummy data. The study established that it is possible to use artificial intelligence to manage the transport sector by use of a system that will not only help identify the patterns and trends of matatus' on Kenyan roads but to answer the why's associated with the trends to help come up with meaningful applicable practical solutions to enhance security and integrity in the transport sector in general. The study also unearthed challenges in relation to the implementation of the above. Combination of classification and association rules based data mining approach was utilized for this study due to its effectiveness in bringing out patterns and trends that are interlinked and related to each other.

Keywords: WHO: - World Health Organization, OB: - Occurrence Book, GUI: - Graphical User Interface, SQL: - Structured Query Language, CCTV: - Closed-Circuit Television, CBD: - Central Business District

1. INTRODUCTION

The matatu industry in Kenya has been a regulation nightmare for all the stakeholders despite all the government regulations that have been put in place. This has been a long standing problem as current solutions and efforts to curb this menace have not yet provided effective and long lasting results to tackle this issue. Different criterions have exhibited different problems thus annulling the "one shoe fits all" existing solutions that have currently been proposed and implemented. Problems exhibited in the matatu industry range from the irritant to the diabolical. The issues range from overlapping, over speeding, careless driving, indiscipline, misconduct, theft robberies, accidents, corruption and even worse. This has made it difficult for the authorities to effectively regulate, monitor and maintain this industry as it has been influenced by several factors which include geographic locations, Sacco's, population, income levels (financial class), proximity to the city and matatu physical attributes.

2. BACKGROUND

The public transport industry is one of the most vital development industries in any country. It is the backbone of transportation where those who do not have private transportation can easily access public service vehicles for a fair price. Governments usually ensure that its population has adequate transport fertilities for its masses. In Kenya this industry is usually dominated by the private matatu industry as majority of the people in both urban and rural areas use this network of transport thus proper policies and regulations should be introduced that protect both the consumer and the operators at equal measure. (Nantulya and Muli, 2009).

Crime and criminality in the matatu industry has been a constant thorn in the flesh for both the citizens and government. This has been a constant undesirable feature of the matatu industry and the situation convoluted due to the blatant and constant corrupt activities involving the police and by a large margin bribes given to the police (WHO, 2012). This makes it difficult for the government to arrest offenders and control crime as the complexities involved in both prevention and detection measures are hampered by complicity by the police in the crime activities.

Misconduct and crime reporting can be done using several different channels such as Calling the police and reporting of incidents that occur. The reported incidents are usually recorded in the police occurrence book also known as the OB. Physical reporting directly at the police station where an individual or group physically makes the complaint directly to the police at the station or reporting to the traffic officers on duty. This form has proven less effective in having the offenders apprehended (Ogendi *et al*, 2013).

2.1 Surveillance

Surveillance is the monitoring of the behavior, activities, or other changing information, usually of people for the purpose of influencing, managing, directing, or protecting. It most usually refers to observation of individuals or groups by government organizations; in this case, surveillance is monitoring and recording data about traffic offenses in various forms. This includes:

- Incidences reports
- Location route information
- Physical attributes of matatu's
- Sacco
- Matatu personnel information

2.2 Data Mining and Profiling

For the purpose of this research we focus on the application of a hybrid data mining model combining Classification based and Association based algorithms to discover patterns and trends within matatu industry data. Data profiling in this context is the process of assembling information about a particular individual or group in order to generate a profile - that is, a picture of their patterns and behavior based on the observable event features.

3. METHODOLOGY

3.1 Main Framework Model

The main framework was based on a hybrid model derived partially from the combination of the two case based reasoning techniques which is the Bayesian Belief Networks from the Bayesian Based learning class of algorithms, the main concept of the Bayesian Based Learning which in principle is as below.

Features of Bayesian learning methods:

i. Each observed training example can incrementally decrease or increase the estimated probability that a hypothesis is correct, unlike algorithms which completely eliminate a hypothesis if it is inconsistent with any single example

ii. Prior knowledge can be combined with observed data to determine the final probability of a hypothesis. Prior probability is provided by asserting a prior probability for each candidate hypothesis and also a probability distribution over observed data for each possible hypothesis

iii. Bayesian learning can accommodate hypotheses which make probabilistic predictions e.g. this patient has a 93 % chance of recovery

3.2 Mathematical Model Design

This involves "mapping" the attributes to their occurrences and identification of their associations to the overall result outcomes. Each attribute will be analyzed separately and together with all others to identify the magnitude effect on the overall outcome they have. Once the attributes, their relationships and effects to the final outcome have been discretely identified, they will be combined with respect to their magnitude and effect to the overall outcome to provide the actual mathematical formula that will be mapped on to the framework to provide the complete theoretical blueprint of the system that can be implemented by a variety of different programming technologies and approaches that observe the algorithmic process. This process will involve several mathematical processes. The processes will focus on two main things; attribute incident association calculation and attribute effect calculation.

In the attribute incident association calculation each associated attribute impacts the overall outcome differently as some have a greater impact on the outcome than others. This will be achieved by computing the magnitude effect of the attribute individually and by combination with other attributes that result to the same outcome.

The general mathematical rule to be observed will be as below:

$$w_{c_a} = \sum_{1}^{n} a_{c_i}$$

Equation 1: Proposed Attribute incident Equation

Where:

i. w_{c_a} Is the overall weight of the attribute in terms of how it affects the outcome
ii. n Is the number of possible combinations of the attributes
iii. a_{c_i} Is the combination instance of the attribute combination

This initial model is subject to alteration and improvement during the research process. The result of this model will be the mathematical determination of the impact of specific attributes or conditions to an outcome.

In addition, the attribute effect calculation model will be used to identify the most dominant attribute that leads to an occurrence or incident. The reverse association rule will be as below:

$$a_{Dominant} = a_{i_{MAX}}$$

Where:

i. $a_{Dominant}$ is the dominant attribute
ii. $a_{i_{MAX}}$ is the particular attribute instance *i* of all the available attributes

Equation 2: Attribute Effect Magnitude Equation

Attribute effect magnitude equation means the most dominant attribute that contributes to an occurrence is that attribute that has the highest number of correlation values in the said instance of computation.

This first main stage will involve the most research effort and time of implementation as the whole concept and purpose of the research is based on this stage.

4. PROTOTYPE IMPLEMENTATION

This model will be used to compute and not only identify the attribute that mostly affects the outcome in a particular way but the value that is most prevalent in an event of such a specific outcome.

The attributes that have been considered as necessary variables for the purpose of the framework model include:

i. **Time** in which the incident has been reported by the observer
ii. **Location** where the incident occurred. This also helps in providing a more accurate picture of what happened by where it happened.
iii. **Registration number** which would have initially been entered into the system for accountability purposes.
iv. **Incident type** of the specific crime or misconduct that the vehicle personnel will have been reported to have committed at any one particular instance.
v. **Route plying** of the specific route that the Vehicle and subsequently the crew will have been registered.
vi. **Economic levels** of the specific route the registered

vehicle is plying.

vii. **Security levels** in the area the matatu is plying.

viii. **Proximity to any major city** the matatu route is within.

The data generated includes the complaints sent by the passengers and public against specific matatu/personnel. This has proved important as it acts as a factor for determining the level of association between the matatus/routes of interest and the incidence trends.

$$R_i = \sum_{1}^{n} W_j(A_i)$$

Which is achieved by summing up the collective weight and attribute scores as below:

$$R_i = [W(A_1) + W(A_2) + W(A_3) + ,,, + W(A_{n-1}) + W(A_n)]$$

where:

R_i is the overall Incidence Score Rank ***e.g. Over speeding =123***

W_j is the attribute weight which represents the attribute impact to the incidence ***e.g. Age has a greater influence to offenses such as over speeding than location.***

This will be achieved by computing it as below.

$$\frac{\sum_{1}^{n} C_{a_i}}{\sum_{1}^{n} \theta C_{a_i}}$$ Where $\sum_{1}^{n} C_{a_i}$ is the total Number of times The Attribute *I* will be considered based on a unique case type ***e.g. the value will be 15 as there are 15 reported cases of over speeding thus the total sum of times the attribute*** a_i ***appears in over speeding.***

$$\sum_{1}^{n} \theta C_{a_i}$$

Is the total sum of all the unique non-repeating Attribute values e.g. the sum will be 4 since there are seven different attribute (Route) values associated to over speeding {44, 19, 9, 237} this might be because the over speeding offense has only been reported on those routes.

The above associations ensure that the less the number of deviations or rather distributions the greater the overall impact weight of the attribute to the offense. This ensures that whatever attributes combinations will be done. Also the weights significance of the attributes will be taken into consideration. This will eliminate guess estimation as the final ranking will be purely based on the data and frequency and distributions of the values and attributes.

After computing the weights of the different attributes to each output case the final ranking will be achieved by combining and ordering the different Case scenarios in terms of the frequencies of reported cases and in combination of different multiple attribute search to identify the most probable offense that might be committed based on the attributes and with that the system will produce the most prevalent value of offense based on a specific case or attribute combination.

The final output is a list of numbers in descending order where the Indecent or offense with the combined most likely association value to a particular event is displayed as the first item on the list of entities and the Indecent or offense with the overall least possible likelihood to be associated to the incident of interest is at the bottom of the list.

The validity of this framework is based on the observation of its ranking order of the Offenses that have been given to it based on all attributes individually and combined. This will measure the validity of the output based on all cases and their recorded deviations. There will be dummy records combined specifically for the system testing phase that will contain specific attributes and the system will have to either reach the same conclusion or fail with a certain degree of error margin.

4.1 Framework Validity Computation

Percentage of viability is measured as the accuracy of the system producing the same results and in the same order to the generated case scenarios validation data as described below.

i. By Individual attribute: this will be achieved by computing first the deviations of the attributes and average them to get the mid-point. This will then be used to compute the deviation of the midpoint of the computed data on the existing test records to measure weather it falls within a certain range from the midpoint using the Euclidean distance.

ii. The Second measurements will be by testing the overall system performance by summing all the test results accuracy/Deviation values and averaging them all to give the cumulative framework validity value that will in effect be used as the overall system measure of “correctness and reliability”

The main achievement of this project framework is the separation of the stages of development into two subsections.

i. The development of an association model that is used to compute the value of the weights of the attributes, based on each attribute value of the unique indecent type.

ii. The development of an offense ranking model that will compute the commonalities in order of frequency of appearance to any Offenses.

5. PROOF OF CONCEPT

This is an intelligent transport management system intended to mine data and bring out certain crime patterns and trends influencing road etiquette on Kenya roads. The sytem has the following features and structure:

i. Saves and retrieves data on Matatus, Matatu personnel, Matatu owners, Matatu sacco details, Route details and Incidence reports.

ii. Mines and displays dominant values for each of the above attributes for any selected crime.

iii. Predicts likelihood of a certain offence occurring using the attributes above.

iv. It is web based system

v. Uses a GUI (graphical user interface) user access mode

vi. Needs SQL server e.g. Apache to run and connect to its database

Analysis Module

This is accessible via the Analysis menu and allows one to choose a particular offence and analyze it as shown below.

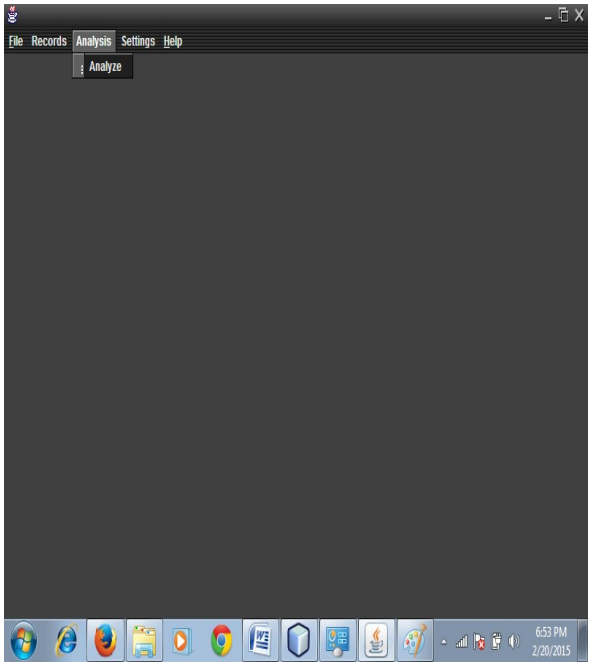

Figure 1. Dashboard Screen

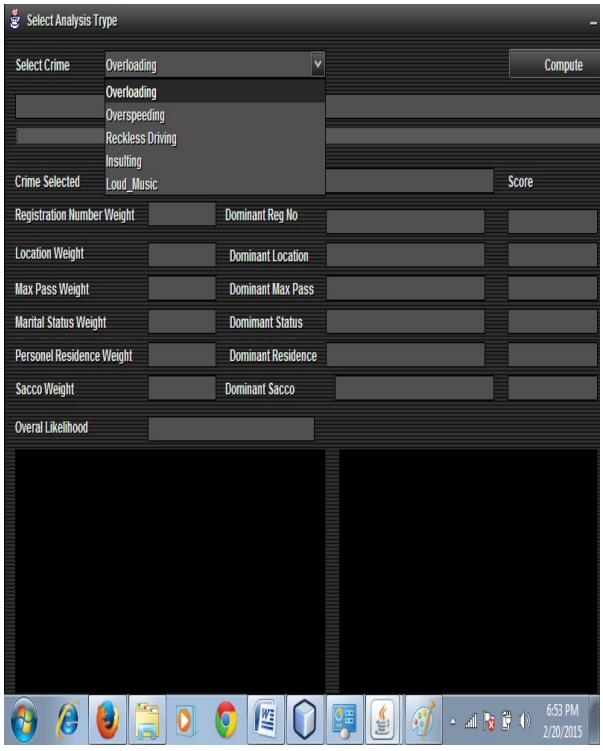

Figure 2. Incidences Analysis Screen

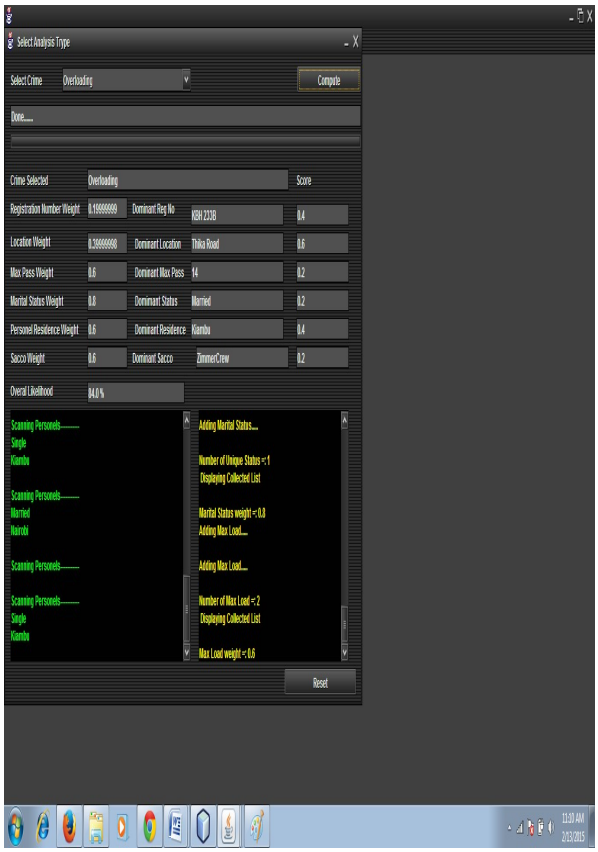

Figure 3. Analysis Output Screen 1

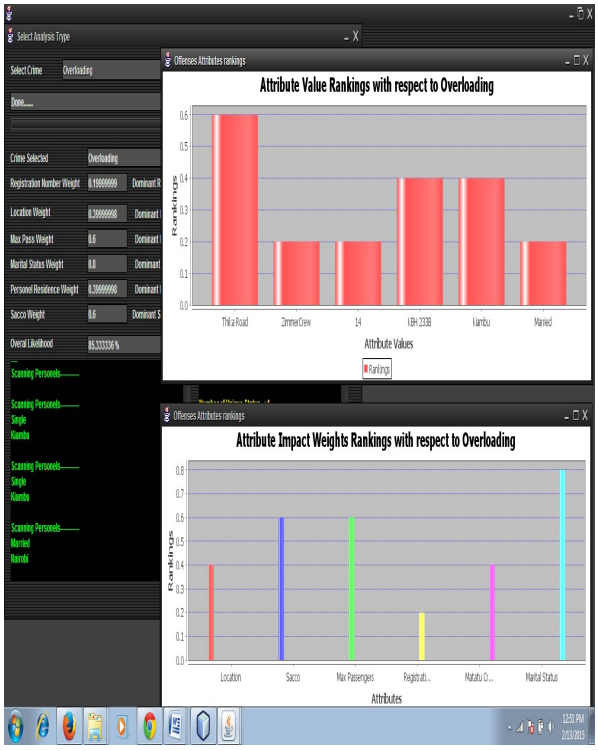

Figure 4. Analysis Output Screen 2

6. KEY FINDINGS

Some key issues have become apparent during the research based on surveillance and below are the noting's.

i. Framework for Intelligent Transport Management System

The first objective was to propose a framework for intelligent transport management system with patterns and trends identification capabilities. The findings were that it is possible to use artificial intelligence to manage the transport sector. This was achieved as illustrated by the and screen shots of the live version of the framework shown earlier.

ii. Enhancement of Policy Formulation and Development

The second objective was to enhance formulation of policy developments, implementations and government regulations for the transport sector in Kenya Events are as a result of active human activity over a period of time. This was shown possible by the capacity of the system to generate and mine vast data generated by the framework which can prove useful decision making.

iii. Testing Model for the Intelligent Prototype

The third objective was to design model system for testing the framework to ascertain its practicability and effectiveness. This was achieved through generation of appropriate dummy data conforming to the information requirements of the model which was then used as a proof of concept of the model proposed by the research.

iv. Identify Challenges of Transport Sector in Kenya

The fourth objective was to identify challenges of the transport sector in Kenya. This was achieved during data generation and extraction which mainly focused on secondary data sources since a lot of research has been conducted previously in this area. Also, through the patterns and trends clearly brought out by the framework

7. KEY ACHIEVEMENTS

The research has resulted to some noted achievements which include:

i. The development of a quantifiable mathematical Model that can be used with desirable confidence to compute the possible relationships between specific attributes both individually and collectively to certain specific crime/Misconduct.
ii. An Automatable framework has been developed that can externalize the human computation process in using the mathematical Model to compute the Associations and identify trends.

8. KEY CHALLENGES

This research like any other has not been without its own challenges, both technical and resource wise as they include.

i. Unavailability of actual real world data to compute real world events thus limiting the models accuracy.
ii. Limited Research time as this research has proved to be wide and has a lot of factors that still need both interpretation and analysis.

9. ASSUMPTIONS AND LIMITATIONS OF SCOPE

The completion of this research and development of this framework and its prototype has come with its own challenges that have been overcome and some ignored due to the near impossibility to tackle natures which are described below.

i. The unavailability of the actual crime data from the police due to bureaucracy and confidential nature of the data.
ii. Also it is assumed that the data collected has been generated from the previous collection of records from offenders.
iii. The above provide limitations to the accuracy of the framework model suggested results as not the actual data has been used but a generated and less accurate data have been used.
iv. The nature of this research is not without its controversial aspect of invasion of privacy to which it is illegal by some government laws. Thus the adoption of this framework means policies within the law will have to be altered to allow the successful implementation of this system.

10. RECOMMENDATIONS

i. Possible Use of Computation

Further Research on the possible use of computations should be considered as it will provide better and more accurate results of the rankings.

ii. Possible Surveillance-Intelligent System Linkage

Actual integration methods that use this framework to various other forms of surveillance such as CCTV and online activity patterns should be considered.

iii. Possible Increase of Attributes Utilized By the System

The use of more attributes should be factored in during the computing process due to the fact that some individuals might exhibit more traits than are currently captured.

11. SUGGESTED AREAS FOR FURTHER RESEARCH

i. Possible Use of Enhanced Security

A study should be conducted on the most effective security mechanisms that can be employed to secure the critical data the system holds and uses.

ii. Attribute - Effect Comparative Study

A comparative study on the effects of each attribute on overall outcome should be conducted to ascertain ranking of attribute – effect relationship.

iii. Broad Study on Challenges of Transport Sector in Kenya

A study on the challenges faced in policing the whole road transport sector in Kenya should be conducted.

iv. Study Expansion to Areas Outside Nairobi CBD

The same research should be carried out in other areas outside Nairobi CBD

12. REFERENCES

[1] Aduwo, I. G. (1990), The Role, Efficiency and Quality of Service of the Matatu Mode of Public Transport in Nairobi: A Geographical Analysis, M.A. Thesis, University of Nairobi.

[2] Asingo, P.O. (2004) The institutional and organizational structure of public road transport in Kenya. IPAR Discussion Paper No.50

[3] Asingo, P. and W. Mitullah (2007). Implementing Road Transport Safety Measures in

[4] Kenya. Working Paper 545. IDS: University of Nairobi

[5] Bruno Agard (2007) Mining Public transport user behavior from smart card data Bruno Agard, Catherine Morency, Martin Trépanier.

[6] Chitere, P. 2006. Public Service Vehicles in Kenya: Their characteristics and Compliance

[7] with Traffic Regulations and Prospects for the Future DP no. 081. Nairobi: IPAR.

[8] Chitere, P., and T. Kibua, 2004. Efforts to Improve Road Safety in Kenya: Achievements

[9] and Limitations of Reforms in the Matatu Industry. Nairobi: IPAR.

[10] Chitere P. O. (2004), Matatu Industry in Kenya. A Study of the performance of their Owners, Workers and theirs Associations and Potentials for Improvement .Institute of Policy Analysis and Research Discussion Paper Series. Discussion Paper No. 055/2004.

[11] Chitere P.O and Kibua T. N. (2001). Efforts to improve road safety in Kenya: Achievements and limitations of reforms in the Matatu industry. Institute of Policy Analysis and Research (IPAR).

[12] Daily Nation Newspapers See Issues of 7th November, 2003 and 3rdFebruary, 4th, 6th, 17th, 19th and 20th, 28th and 30th August, 11th 12th and 14th Sept., 2004

[13] Dale Dzemydiene, (2010),Pattern recognition based on statistics and structural equation models in multi-dimensional data warehouses of social behavioral data Dale Dzemydiene,Raimundas Vaitkevicius, Ignas Dzemyda pg 4 -10

[14] East African Standard Newspaper See Issue of 14th Sept, 2004

[15] East African Newspaper Issue of November 24-30, 2003

[16] Information on transport action plan from Central Connecticut Regional planning agency retrieved from http://www.ccrpa.org/transportation/LRTP_action_plan/ACTION%20PLAN.pdf

[17] Khayesi, M (1999), The struggle for regulatory and economic sphere of influence in the matatu means of transport in Kenya – a stakeholder analysis at Sixth International Conference on Competition and Ownership in Land Passenger Transport

[18] Odero, W., Khayesi, M. and Heda, P.M. (2003) Road traffic injuries in Kenya: Magnitude, causes and status. In: Injury Control and Safety Promotion, Vol.10 (1-2)

[19] Ogonda, R. T. (1992), Post-Independence Trends in Development of Transport and Communications, in Ochieng' W. R. and R. M. Maxon, eds., An Economic History of Kenya, Nairobi: East African Publishers, pp. 313-326.

[20] Republic of Kenya (2003 and 2004) Legal Notices Nos. 161, 83 and 97

[21] Republic of Kenya (2004) Transformation of Road Transport Report, MOTC, Nairobi

[22] Republic of Kenya (2010). Kenya Police Report 2010.

[23] Republic of Kenya (2010). The Constitution of Kenya (27th August (2010). Nairobi: Government Printers

[24] Republic of Kenya (2012). Kenya National Highway Authority (KeNHA): Quality Highways, Better Connections. Feasibility studies and detailed engineering design of the multinational Arusha – Holili/Taveta – Voi road. RESETTLEMENT ACTION PLAN (RAP) FINAL REPORT, October 2012.

[25] Tom M. Mitchell (March 1, 1997), Machine Learning

[26] The World Bank (2007a), India: World Bank Supports National Highway Systems Improvements in Uttar Pradesh and Bihar, Press Release No: 2005/251/SAR Washington, D.C., December 21, 2004, available from: http://www.worldbank.org/ in website publication

[27] The World Bank (2007b), Uganda Signs US$ 107 million Agreement for Road Sector Development, Press Release No: 2005/354/AFR Kampala dated February 23, 2005 available at http://web.worldbank.org/afr/ug

[28] Van der Veer (2009), Data mining for intelligence led policing, van der Veer, H.T. Roos, A. van der Zanden

[29] Vikas.Grover (2009), Review of Current Crime Prediction Techniques, Vikas Grover, Richard Adderley, Max Bramer.

[30] Noemi Dreyer Galvão; Heimar de Fátima Marin (2009), Data mining: a literature review, Técnica de mineración de datos: una revisión de la literature

International Journal of Computer Applications Technology and Research
Volume 9–Issue 06, 200-210, 2020, ISSN:-2319–8656

Web-Based Application for Alcohol Behavioral Change in Nigeria

Mgbike, Charity Ngozi
Department of Computer Science
Chukwuemeka Odumegwu Ojukwu University
Uli. Anambra State. Nigeria.

Okeke, Ogochukwu C.
Department of Computer Science
Chukwuemek Odumegwu Ojukwu University
Uli, Anambra State. Nigeria

Abstract: Everyday young people are introduced to alcohol and this early introduction to alcohol consumption eventually leads to abuse. The burden of diseases caused by the abuse of alcohol, the harm done to their social life and the decimation of their finances are some of the motivation for this work. This paper aims at providing an alternative and better approach to solving the problem of alcohol addiction using behavioral change theories and persuasive software. It has been found that many people are extremely open and motivated when talking to a computer about personal problems. The system is designed making with human behaviour theories; Trans theoretical Model (TTM), Goal Setting Theory, Reinforcement Theory, Social Learning Theory and Knowledge-Attitude-Behaviour Model (KAB). This was done bearing in mind that proper understanding of human behaviour theories enables designers to know which theory to apply and how the application of the theory will help the user. In addition to human behaviour theories, persuasive strategies were incorporated in the design of the system. Object-Oriented Analysis and Design Methodology (OOADM) was employed. Implementation was achieved using Hypertext Preprocessor scripting language and MySQL server. The resulting system is a web-based application that provides users with not just the requisite information needed about alcohol addiction but also exercises and activities that persuades them to think deep about their addiction and work hard towards overcoming them.

Keywords: Alcohol, Alcohol Abuse, Behaviour Change, Persuasion, Persuasive Technology

1. INTRODUCTION

Alcohol is one of the most popular psychoactive substances in the world. It can have powerful effects on your mood and mental state. By reducing self-consciousness and shyness, alcohol may encourage people to act without inhibition. At the same time, it impairs judgment and promotes behavior people may end up regretting (Arnarson, 2018). Alcohol is as old as human history and its consumption in different socio-cultural backgrounds extends beyond the last ten thousand years (Smart, 2007). Its consumption has been considered normal especially when drunk without outright intoxication in Africa and other parts of the globe. Wine, beer, spirit and other fermented alcoholic beverages were drunk in traditional societies and some of these beverages are still used in this modern era for different purposes (Dumbili, 2013). In Africa, these and other alcoholic beverages such as *palm wine, burukutu*, etc. were consumed for pleasure soon after brewing or tapping and were rarely traded in the market, [World Health Organization (WHO), 2002]. Though alcoholic beverages have been consumed for hundreds of years, the pattern and purpose of consumption vary considerably among and even within communities. Excess consumption was not widely tolerated in many societies while few communities permitted it (Willis, 2006).

Drinking alcohol in moderation has a lot of health benefits. Moderate amounts of alcohol raise levels of High-Density Lipoprotein (HDL) or 'good' cholesterol and higher HDL levels are associated with greater protection against heart disease. Moderate drinking might actually protect against erectile dysfunction in the same way that drinking red wine might benefit heart disease. It can also help reduce the risk of gallstone, help decrease the chances of dementia and lower the chances of type 2 diabetes (Bachai, 2013).

The impact of alcohol in the body starts from the moment the first sip is taken. A glass a day may do little damage to a person's overall health. But if the habit grows or if you find it difficult to stop after

just one glass, the cumulative effects can add up (Pietrangelo & Holland, 2017). Though alcohol has some benefits, its negative effects far outweighs the benefits. It's safe to say that alcohol is both a tonic and a poison. The difference lies mostly in the dose. Heavy drinking can damage the liver and heart, harm an unborn child, increase the chances of developing breast and some other cancers, contribute to depression and violence, and interfere with relationships. Heavy drinking is a major cause of preventable death in most countries. Alcohol is responsible for about half of fatal traffic accidents. Even moderate drinking carries some risks. Alcohol can disrupt sleep. Its ability to cloud judgment is legendary. Alcohol interacts in potentially dangerous ways with a variety of medications, it is also addictive, especially for people with a family history of alcoholism (Harvard School of Public Health, 2015).

If consumed to excess, alcohol can affect all areas of a person's life, as well as the lives of their family and friends. Personal relationships can be subjected to arguments over drinking which can lead to ongoing conflict and break ups. There are also negative consequences in the workplace arising from poor performance, accidents and absenteeism as a result of alcohol. For some, drinking in excess can lead to legal problems as a result of anti-social and violent behaviour or the loss of their driving license. Alcohol consumption is also associated with noise, violence, offensive behaviour, vandalism, petty crime and motor vehicle accidents; all of which can reduce quality of life for family members, bystanders and the community. These factors are associated with social concern and presumably their presence reduces the quality of life of the affected individuals. In addition, these factors significantly contribute to the costs of welfare and policing services (Social Issues Research Centre, 2009)

In the traditional era, alcohol was consumed by the men served by the youth. Consequently, women and youths were hardly seen consuming alcohol. Today, the reverse has become the case as alcohol is no longer the reserve of men neither is the age of the drinker considered as so important; it is widely accepted in our society as most homes use it as a form of entertainment for their guests. Although the minimum drinking age remains 18 years, young people buy and drink alcohol freely in public places (Dumbili, 2013). It is also one of the mostly anticipated beverages at weddings, parties, child dedication and other functions. Any occasion or function as it were is never complete without the presence of alcoholic beverages. In addition, pubs and clubs serve it freely as a major recipe. Among the youths of today, it is used as a show of maturity or 'senior boy', as a youth who does not drink alcohol is looked down upon by his peers.

The new trend of alcohol consumption among young people contributes to Nigeria's ranking among thirty countries with highest per capita consumption of alcohol globally (WHO, 2004). A major contributor to alcohol abuse is the absence of alcohol policy in Nigeria. Globally, alcohol producers frown at strict measures to regulate the production and marketing of alcohol through legislation due to economic interest (Dumbili, 2013). Although Nigeria and many other African countries contribute to the global burden of alcohol-related problems due to increasing harmful use, only a few countries within Africa have policies to regulate alcohol use and misuse. Even those that have policies, the vested interest, which affects not just Africa, but the other part of the world can render policy ineffective and subject to misuse (Dumbili, 2013).

It is against this background that the researcher seeks to develop an alcohol intervention tool, to deal with this increasing issue of alcohol abuse. This tool shall be developed using an aspect of Computer Science known as Persuasive Technology together with behavior change theories fronm psychology and the main tool of this persuasive technology for will be tailored information.

Scholars do not agree on the precise definition for persuasion, but for the purpose of this work, persuasion is a non-coercive attempt to change attitudes or behaviors. There are some important things to note about persuasion. First, persuasion is non-coercive. Coercion which is the use of force is not persuasion; neither is manipulation or deceit. These methods are shortcuts to changing how people believe or behave (Fogg, Cuellar, & Danielson, 2007).

Persuasion is part and parcel of human interaction. From the serpent in the Garden of Eden to our modern mass-media society, persuasive efforts abound in a continuous attempt to influence our attitudes and behaviors, convincing us to spend money on one product rather than another, to vote for a particular political party, to stop smoking, to quit alcohol consumption, to exercise more, to take the stairs instead of the elevator, to fight for environmental conservation, animal well-being, better schools. (IJsselsteijn. De Kort, Midden, Eggen & Van den Hoven, 2006).

Consequently, Persuasive computers are defined as the computing systems, devices, or applications intentionally designed to change a person's attitudes or behavior in a predetermined way (Cheng, n.d.). Persuasive Technology (PT) is a vibrant interdisciplinary research field, focusing on the design, development and evaluation of interactive technologies aimed at changing users' attitudes or behaviors through persuasion and social influence, but not through coercion or deception. Persuasive technologies are used to change people's behavior in various domains such as healthcare, sustainability, education or marketing.

Persuasive technology, which can take the form of apps or websites, marries traditional modes of persuasion using information, incentives with the new capabilities of devices to change user behavior. Persuasive technology can be found in mobile downloads, or on the digital homes of tech giants like Amazon and Facebook, where behavior-oriented design persuades us to buy more often (one-click checkout) or stay logged in (manipulating social media news feeds). Many mobile apps that try to influence user behavior are either health-oriented – apps that incentivize weight loss, help to manage addictions and other mental health issues, or influence sleep practices or promote environmental awareness. Though it is been around for a while, persuasive technology is becoming increasingly popular and profitable, inviting a deeper look into its ethics and efficacy (Larson, 2014).

This incessant abuse of alcohol has dire consequences on the health, social life and economy of individuals in particular and the society at large. Social effects in the sense that a young man or woman who abuses alcohol and has become an addict will soon find out that it starts affecting his performance at work and his relationship with people. Its health effect is obvious in that a lot of chronic diseases like diabetes have been linked to alcohol as one of its causal elements. The economic effect is not far-fetched because a person who keeps track of his alcohol consumption will discover after a period of time that the amount spent on alcohol would be have enough to make a reasonable investment.

Several persuasive applications have been developed by various stakeholders in order to combat the abuse of alcohol especially among youths, but none of them is in the form of persuasive web-based application for Nigerian youths. Some of these persuasive applications did not make use of behavior change theories in their design, the concentration was mainly on persuasive strategy. The aim of this paper is to apply a web-based effect of alcohol behavioral change in Nigeria; with the following objectives; Develop a web-based application for alcoholic behavior change using behavioral change models, Implement an application that will enable users know the concentration of alcohol in their body, Create an efficient system that will enable users keep track of their consumption in order to make an informed decision, and to evaluate the effectiveness of the effects in reducing alcohol consumption using quantitative analysis.

This paper will be of enormous benefits to the general public, private and community.

Finally, this work is also significant to the government especially the Federal Road Safety Commission who deals with drunk drivers on a daily basis. Research has shown that a lot of the accidents that happen especially on weekends are as a result of drunk driving. This application will help their "Don't drink and drive" campaign and also in the rehabilitation of drunk drivers.

2. REVIEW OF RELATED WORKS

2.1 Persuasive Technology

Persuasion has been part and parcel of human relations. From the days of Eve and the serpent in the Garden of Eden to our present day mass-media society, persuasive efforts continue to thrive in diverse areas of human endeavor, in an attempt to influence our attitudes and/or behaviors. Trying to convince us to spend money one product instead of another, to eat well in order to keep fit, to reduce alcohol consumption, to take the stairs instead of the elevator, to exercise more, to work hard now and rest at old age, etc.(IJsselsteijn *et al*., 2006) .

Persuasive computers are the computing systems, devices, or applications intentionally designed to change a person's attitudes or behavior in a predetermined way (Cheng, n.d.). In the same way, (Khaled, Barr, Noble, Fischer & Biddle, 2006) asserted that Persuasive Technologies are interactive computing systems designed to change attitudes or behaviours, which are utilized in areas as diverse as marketing, health, safety, environmental conservation, politics, religion, gaming, self-efficacy, occupational effectiveness among a list of others. Similarly, Persuasive

Technology is a term used to describe technologies that change behavior and/or attitude in an intended way without the use of deception or coercion (Orji, Vassileva & Mandryk, 2012). Also, Persuasive technologies can be said to be defined as computerized software or information systems designed to reinforce, change or shape attitudes or behaviours or both without using coercion or deception (Lehto & Qinas-Kukkonen, 2009).

From the above definitions of Persuasive Technologies, it can be deduced that the main aim of Persuasive Technologies is to bring about a change in attitude or behavior, which will eventually result in a change of character of its users. Persuasive Technology is fundamentally about inducing behavior and/or attitude change using computers. Developing effective persuasive technology for behavior change requires in-depth and practical understanding of human behavior theories, so that designers of persuasive intervention do not have to guess at which ad-hoc approaches will succeed. The most effective persuasive interventions for behavior change usually occur when the intervention is behaviorally focused and theory driven. Therefore, to design technologies that will successfully motivate behavior change, persuasive researchers rely on behavioral theories from psychology and social science to inform their decisions (Orji et al., 2012).

In all persuasive technologies, motivation is a very important element that leads to initiation of behavior change actions, continuation of the actions and maintenance of the behavior over time (Orji et al., 2012). The researcher will look at two important topics which are important to the design of persuasive technologies for wellbeing namely: Social Influence (social competition, collaboration, and comparison) and Learning and Reflection.

2.1.1 Social Influence

Social influence is defined as change in an individual's thoughts, feelings, attitudes, or behaviors that results from interaction with another individual or a group. It is the process by which individuals make real changes to their feelings and behaviors as a result of interaction with others who are perceived to be similar, desirable, or expert. People adjust their beliefs with respect to others to whom they feel similar in accordance with psychological principles such as balance. Individuals are also influenced by the majority: when a large portion of an individual's social group holds a particular attitude, it is likely that the individual will adopt it as well. Additionally, individuals may change an opinion under the influence of another who is perceived to be an expert in the matter at hand (Rashotte, 2015).

Systems that use social influence as a motivator typically focus on sharing information about one's physical activity with one's social groups such as friends, coworkers, neighbours and family. In these systems, social competition, collaboration, and social comparison are often an explicit design goal or a consequence of how the systems are used. Chick Clique and Houston are examples of systems in this category. Chick Clique uses a mobile phone and pedometer to help teenage girls track and share their step counts with their friends. Houston, similar to Chick Clique, encourages groups of users to track and share their step counts as recorded in a pedometer via their mobile phone (Orji et al., 2012).

2.1.2 Reflection and Learning

The difficulties associated with making people change their behavior using an approach that causes immediate performance of the behavior led to a search for an alternative way of effecting behavior change. For example, persuasive approaches based on tracking of behavior might require that the user continuously uses the application to enact the desired behavior. It remains to be seen for how long the users will use these applications. It will not be realistic to expect that users will use these applications for their life-time (Orji et al., 2012). For instance, in the evaluation of the HealthyEdge persuasive application (Xu, Chen, Uglow, Scott & Montague, 2011), the participants expressed discomfort that they experienced when they attached the device to their body. Therefore, a number of studies have looked at emphasizing reflective thinking about health as an approach to behavior and attitude change that indirectly impacts the behavior. The work based on this approach uses both personal prompt and group discussion as a mechanism to trigger reflective thinking. For example, the participants in the evaluation of a persuasive game (Grimmes et al., 2010) reported how playing the game prompted group discussions (in line with social learning theory) that facilitated reflection about healthy eating. Playing the game increased the consciousness of the players toward what they eat, which also led to increased personal reflection about their diets. The increasing interest in reflective approaches to

behavior change is due to its potential to intrinsically motivate and thereby results in a long-term behavior change (Colineau & Paris, 2010). This approach is supported by a research finding that critical reflection is a key to transformative learning (Taylor, 2000).

2.2 The Functional Triad

Computers play many roles, some of which go unseen and un-noticed. From a user's perspective, computers function in three basic ways: as (a) tools, as (b) media, and as (c) social actors. In the last two decades, researchers and designers have discussed variants of these functions, usually as metaphors for computer use. However, these three categories are more than metaphors; they are basic ways that people view or respond to computing technologies. These categories also represent three basic types of experiences that motivate and influence people (Fogg et al., 2007).

The functional triad is a framework for thinking about the roles that computing products play, from the perspective of the user (Fogg, 2003). Described in more detail, the Functional Triad is a framework that makes explicit these three computer functions – tools, media and social actor. (Fogg et al., 2007).

The Functional Triad is not a theory; it is a framework for analysis and design. In all but the most extreme cases, a single interactive technology is a mix of these three functions, combining them to create an overall user experience. In captology the Functional Triad is useful because it helps to show how computer technologies can employ different techniques for changing attitudes and behaviors. For example, computers as tools persuade differently than computers as social ac- tors. The strategies and theories that apply to each function differ. Below it is explained in detail the roles the computers play (Fogg et al., 2007).

2.2.1 Computers as Persuasive Tools

Computers as persuasive tools affect attitude and behavior changes by increasing a person's abilities or making something easier to do. Although one could propose numerous possibilities for persuasion in this manner, below are four general ways in which computers persuade as tools: by

- **a.** Increasing self-efficacy
- **b.** Providing tailored information -
- **c.** Triggering decision making (Fogg et al., 2007).

2.2.2 Computers as Persuasive Media

The next area of the Functional Triad deals with computers as persuasive media. Although "media" can mean many things, here the focus is on the power of computer simulations. In this role computer technology provides people with experiences, either first-hand or vicarious. By providing simulated experiences, computers can change people's attitudes and behaviors. Outside the world of computing, experiences have a powerful impact on people's attitudes, behaviors, and thoughts. Experiences offered via interactive technology have similar effects. Three types of computer simulations are relevant to persuasive technologies:

- **a.** Simulated cause-and-effect scenarios
- **b.** Simulated environments
- **c.** Simulated objects (Fogg et al., 2007).

2.2.3 Computers as Persuasive Social Actors

The final part of the Functional Triad focuses on computers as "persuasive social actors," a view of computers that has only recently become widely recognized. Past empirical research has shown that individuals form social relationships with technology, even when the stimulus is rather impoverished. For example, individuals share reciprocal relationships with computers and are polite to computers. In general, it has been found that, computers as social actors can persuade people to change their attitudes and behaviors by

- **a.** Providing social support
- **b.** Modeling attitudes or behaviors
- **c.** Leveraging social rules and dynamics (Fogg et al., 2007).

2.3 Overview of Behaviour Change Theories

Since Persuasive Technologies are designed with the intent of deliberately bringing about behavior change, designers of persuasive technologies must have an in-depth and practical understanding of behavior change theories. This is to prevent designers from guessing at which ad-hoc approach will succeed. Research has shown that the most effective persuasive interventions for behavior change is when the intervention is behaviourally focused and theory driven (Orji et al., 2012). Therefore, in order to design persuasive applications that will motivate behavior change, persuasive researchers make use of behavior theories from

psychology and social sciences. Some of the theories that will be studied in this work include; Trans theoretical Model (TTM), Goal Setting Theory, Reinforcement Theory, Social Learning Theory and Knowledge-Attitude-Behaviour Model (KAB).

2.3.1 Trans theoretical Model

The Trans theoretical Model is one of the most used model for health-related behavior change interventions (Orji et al., 2012). Stages of Change lie at the heart of the Trans theoretical model. The model believes that a change in behaviour follows five distinct stages and that while the time a person can stay in each stage is variable, the tasks required to move to the next stage are not. It recognizes that change is a process that unfolds over time, involving progress through a series of stages (Noar, Benac, & Harris, 2007). According to TTM the five (5) stages that an individual will undergo in order to attain behavior change includes: Pre contemplation (Not Ready), Contemplation (Getting Ready), Preparation (Ready), Action and Maintenance.

2.3.2 Goal Setting Theory

Goal Setting Theory is one of the theories of motivation and also a behaviorism theory. Based on hundreds of studies, the major finding of goal setting theory is that individuals who are provided with specific, difficult but attainable goals perform better than those given easy, nonspecific or no goals at all. At the same time, however, the individual must have sufficient ability, accept the goals and receive feedback related to performance (Lunenburg, 2011). A goal is the aim of an action or task that a person consciously desires to achieve or obtain (Locke and Lotham, 2002). Furthermore, Goal Setting involves the conscious process of establishing levels of performance in order to obtain desirable outcomes (Locke & Latham, 2006).

2.3.3 Reinforcement Theory

Reinforcement theory is a psychological principle maintaining that behaviors are shaped by their consequences and that, accordingly, individual behaviors can be changed through rewards and punishments (Wigmore and Rouse, 2016). Generally speaking, there are two types of reinforcement: positive and negative (Barnett and Simmering, 2016).

Positive reinforcement results when the occurrence of a valued behavioral consequence has the effect of strengthening the probability of the behavior being repeated. The specific behavioral consequence is called a reinforcer. An example of positive reinforcement might be a salesperson that exerts extra effort to meet a sales quota (behavior) and is then rewarded with a bonus (positive reinforcer). The administration of the positive reinforcer should make it more likely that the salesperson will continue to exert the necessary effort in the future (Barnett and Simmering, 2016).

Negative reinforcement results when an undesirable behavioral consequence is withheld, with the effect of strengthening the probability of the behavior being repeated. Negative reinforcement is often confused with punishment, but they are not the same. Punishment attempts to decrease the probability of specific behaviors; negative reinforcement attempts to increase desired behavior.

2.3.4 Social Learning Theory

Social Learning Theory posits that people learn from one another, via observation, imitation, and modeling. Most human behavior is learned observationally through modeling: from observing others, one forms an idea of how new behaviors are performed, and on later occasions this coded information serves as a guide for action. Social learning theory explains human behavior in terms of continuous reciprocal interaction between cognitive, behavioral, and environmental influences.

2.3.5 Knowledge-Attitude-Behaviour Model (KAB)

The Knowledge-Attitude-Behavior model emphasizes on the importance of knowledge as a prerequisite for intentional performance of health-related behavior. An acquisition of new knowledge leads to changes in attitude, which in turn leads to change in lifestyle (Orji et al., 2012). Knowledge at some level, in some form, is a logical prerequisite to the intentional performance of health-related behaviors (Baranowski et al., 2003).

The Knowledge-Attitude-Behavior (KAB) model has been proposed as a way of explaining the role of knowledge. The KAB model proposes that behavior changes gradually. As knowledge

accumulates in a health behavior domain, changes in attitude are initiated. Over some period of time, changes in attitude accumulate, resulting in behavioral change (Baranowski et al., 2003).

The primary resource in this model seems to be the accumulation of knowledge. At some point, this accumulation cascades into changes in attitudes, behaviors, or both (Baranowski et al., 2003).

2.4 Knowledge Gap

After a careful study works related to this study, it was observed that previous systems used behavior change model sparingly. Secondly, they also made use of the one-size-fits-all approach, that is, they used the same solution for different alcohol users. Nevertheless, research has shown that there are three group of drinkers

1. hazardous or risk drinkers who consume alcohol above recommended limits without noticeable harm.
2. harmful drinkers who have experienced some physical, social or psychological harm without meeting the criteria of dependence, and
3. alcohol abusing or dependent drinkers who meet these criteria and sustain consuming alcohol regardless of substantial negative consequences.

This new solution makes a provision for different solutions for these three group of drinkers, which is; it provides different menus for the three groups of alcohol users. Finally, the system incorporated the used of behavior change theories in the design of persuasive technologies.

3. METHODOLOGY AND SYSTEM ANALYSIS

3.1 Methodology Adopted

The methodology adopted for this research work is the Object Oriented Analysis and Design Methodology (OOADM) with the Unified Modeling Language (UML) as the modeling language for the research work. OOADM is a structured method for analyzing and designing a system by applying the object-orientated concepts, and developing a set of graphical system models during the development life cycle of the software to foster better stakeholder communication and product quality.

3.2 Analysis of the Existing System

In the existing system, people who are trying to quit alcohol or people who for one reason or the other want to reduce their alcohol consumption used this website to achieve a great result. Users sign up with an alias and are given a toolbox which contains four tools. The function of these tools is to help the users start their journey towards combating alcohol, deal with challenges encountered in the process and maintain the success attained. The same toolbox is given to all the classes of alcohol users and information is presented in a plain manner without much preaching or persuasion.

One of the feature of this application is its educational email program. It sends an email to users over a period of eleven weeks with interesting facts about drinking and health. More so, there is also a provision for a diary, with which users can record their progress. Another feature of this website is the Education section where users are provided educative materials that will get them motivated for the journey towards alcohol cessation/reduction. Finally, the website offers referral services to users who have other health issues asides alcohol like depression, smoking and other problems. The following weaknesses were found in the existing system;

i. Persuasive strategies were sparingly used in the development of the applications.

ii. Behavioural change theories were not used in the design

iii. The existing system uses the one size fits all approach in the design of the system. It does not take into consideration the different types of alcohol abusers.

iv. The tutorial parts of the applications are too lengthy for the average Nigerian lacks a reading culture.

v. The applications were designed mainly for non-Nigerians, this is evident in the setting of the application and the alcoholic beverages used.

3.3 Analysis of the New System

The researcher designed and developed a persuasive web-based application for alcohol intervention. In particular, system is an application for alcoholic behavior change. It is a website where participants in the alcohol intervention programme sign up to the portal and each participant is assigned menus according to their level of addiction. These menus contain interactive activities and educational materials which the new participant uses to start up on the journey. This application employs persuasive principles, bearing in mind the fundamental aim of persuasive technology which is to cause a change of behavior or attitude in an individual without the use of deception or coercion. Thus, in essence the proposed system is a persuasive technology.

First and foremost, the new system employs behavior change theories; Transtheoretical Model (TTM), Goal Setting Theory, Reinforcement Theory, Social Learning Theory and Knowledge-Attitude-Behaviour Model (KAB). The need for behavior change theories is based on findings that developing an effective persuasive technology, requires an in-depth and practical understanding of human behavior theories so that designers of persuasive application do not guess at which ad-hoc approaches will succeed. Thus, the researcher adopted human beahaviour theories from social sciences and psychology in the design and development of the new system.

With the help of the Transtheoretical Model (TTM), once an individual's behavioural stage is determined, interventions are used to help facilitate progression into the next stage of TTM. On the other hand, Goal Setting Theory motivate behavior change, by encouraging individuals to set goal which in turn leads to higher performance. Moreover, the use of Reinforcement helps the researcher employ the use of incentive in reinforcing a desirable behavior. Social Learning Theory is another theory which is very important to this research work. The researcher uses it to show the point that learning a desirable behavior can occur through interactions between individuals, between individuals and groups and between different groups. Finally, Knowledge-Attitude-Behaviour (KAB) Model is used to show that an acquisition of new knowledge leads to a change in behavior.

Secondly, the proposed system accommodated different types of problem drinkers. It was found from research that there are three groups of problematic drinkers; hazardous or risk drinkers, harmful drinkers and alcohol abusing or dependent drinkers. The participants at the point of registration will chooses which category they fall into and start the intervention program from there.

Moreover, the application consists of four major sections namely; Alcohol Overview, Are You Addicted, Cutting Down and Staying on Track. These sections contain series of activities and educational information that a new user makes use of in the course of the intervention program. These activities include; Blood Alcohol Calculator, Do I Drink Too Much, Set Your Goal, Track Your Drinking, Cutting Back, Changing the Rules, The First Two Weeks, Dealing with Desires, Forums, Result Achieved and a lot of other educational information that will help the participants fight and win the battle against alcohol abuse.

The first section which is the Alcohol Overview contain information about alcohol, its effects to the health, socially and economically. Following the first section is the Are You Addicted section which helps users determine whether they are addicted and their level of addiction. Next is the Cutting Back section which helps users with actual activities needed to be performed on a daily basis while trying to quit alcohol. Finally, the researcher has the Staying on Track section which is kind of a maintenance section. When users have recorded a certain level of success, they can use the activities and exercises in this section to maintain their new found behaviour.

Finally, the proposed system adopted persuasion strategies such as framing, salience, commitment, authority, social proof and reciprocation in the design and development of the application.

3.4 Class Diagram of the New System

Figure 3.1 is the class diagram of the proposed system. It shows the building blocks of the new system. Class diagrams depict the static view of the model or part of the model, describing what attributes and behaviour it has rather than detailing the methods for achieving operations. Class diagrams are most useful to illustrate relationships between classes and interfaces.

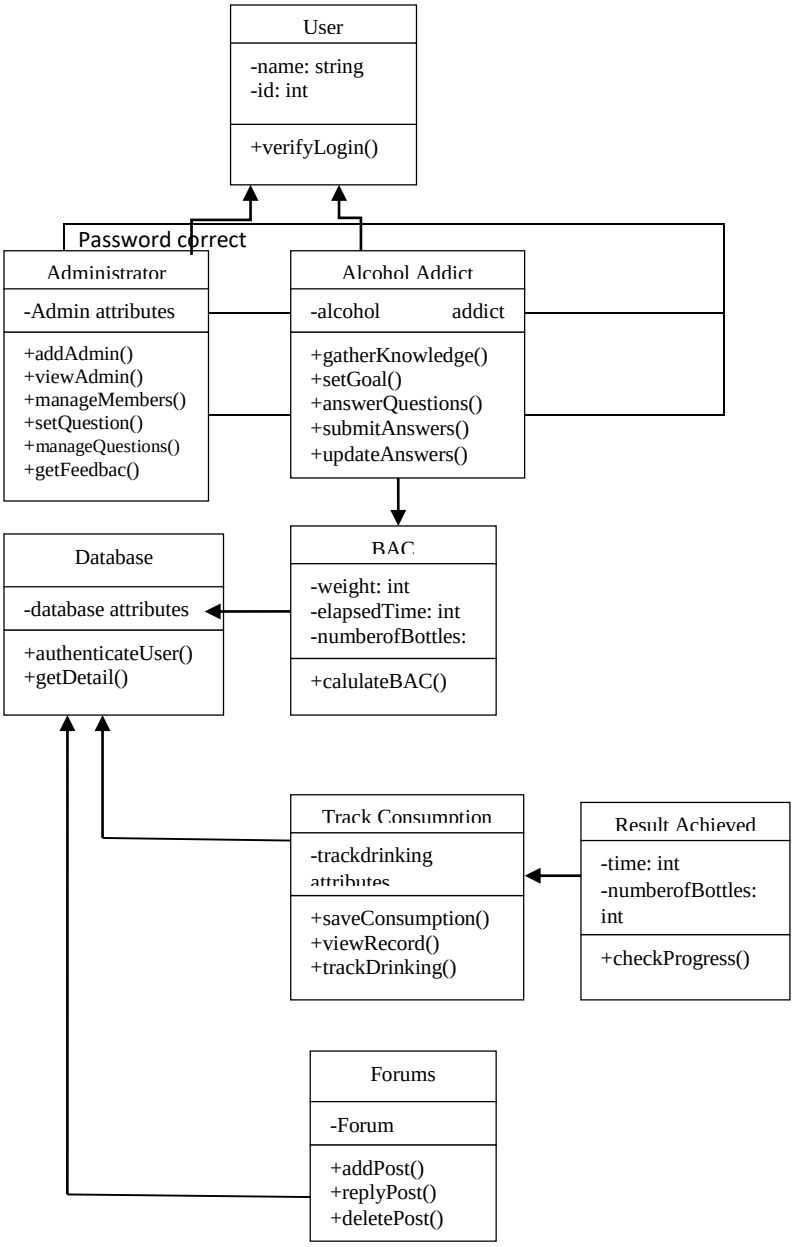

Figure 3.1: Class diagram of the new system

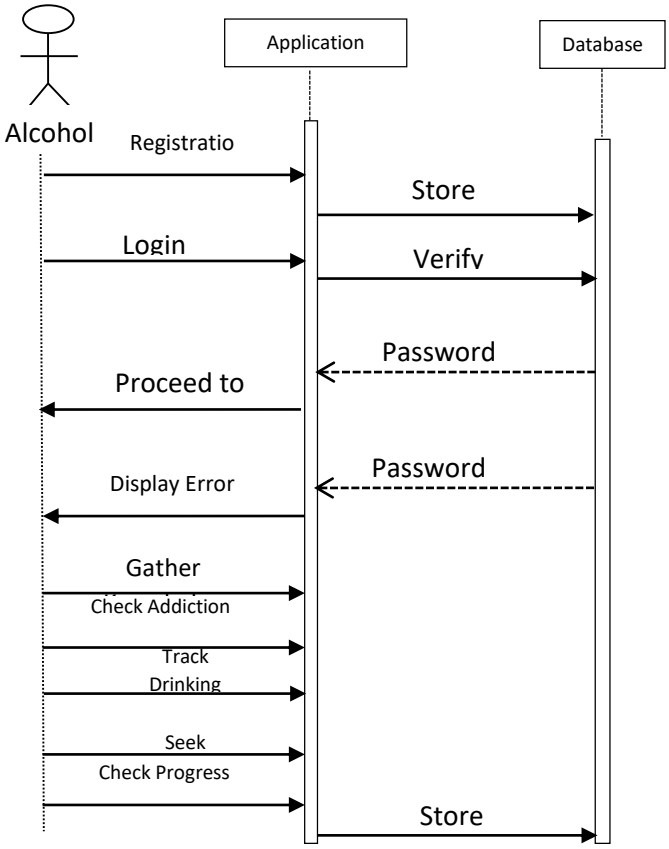

Figure 3.2 Sequence of the new system

3.5 Sequence diagram of the new system

Figure 3.1 shows the sequence diagram for the new system. A sequence diagram shows objects as lifelines running down the page and with their interactions over time represented as messages drawn as arrows from the source lifeline to the target lifeline. Sequence diagrams are good at showing which objects communicate with which other objects and what messages trigger those communications.

System Architecture and Implementation

The new system is a combination of windows project and database project hence the program development was carried out using PHP programming language. The new system was developed using windows forms which contains Graphical User Interface. The database aspect of the new system was developed using Microsoft SQL server. The system shall be in three layers architecture namely: the client-tier, the web-tier and the database management system. The client-tier components were implemented using HTML and CSS, the web-tier was implemented using Hypertext Preprocessor (PHP) Scripts version 5.4. While the database management system is a relational database designed using MySQL Database Engine Version 5.6.

CONCLUSION

For so long the abuse of alcohol has been o great worry to all stakeholders concerned. This has led to the design of a persuasive technology for alcohol behaviour in order to alleviate the problem of alcohol abuse. This work has shown that the fight against alcohol abuse can be won with the use of persuasive technology that is; tailoring information, behavior change theories and persuasive strategies. The study therefore concludes that, with persuasive technology for alcohol behaviour change, users are able to increase their knowledge about alcohol and this application has the capability to cause deep reflection on the part of the users which will in turn lead to a change in attitude of the users.

REFERENCES

1. Arnarson, A. (2018). Alcohol and Health: The Good, the Bad and the Ugly. *Healthline.* Retrieved from https://www.healthline.com/nutrition/alcohol-good-or-bad
2. Bachai, S. (2013). *7 Health Benefits of Drinking Alcohol*. Retrieved from .2.
3. Baranowski, T., Cullen, K. W., Nicklas, T., Thompson, D. & Baranowski, J. (2003). Are Current Health Behavioral Change Models Helpful in Guiding Prevention of Weight Gain Efforts? *Obesity Research. 2003*;11:23S-43S.
4. Barnett, T. & Simmering, M. (2016). Motivation and Motivation Theory. Retrieved from http://www.referenceforbusiness.com/management/Mar-No/Motivation-and-Motivation-Theory.html.
5. Colineau, N. & Paris, C. (2010). *Can Beneficial Habits be Induced Through Reflection?* Retrieved from http://www .2shopUM4Motivation.pdf.
6. Cheng, R. (n.d.). *Persuasive Strategies for Computers as Persuasive Technologies.* Retrieved from https://pdfs.semanticscholar.org/bec7/83c4f752ff89b4a5ef863cd66ad76c8f3f51.pdf<<
7. Dumbili, E. (2013). Changing Patterns of Alcohol Consumption in Nigeria: An Exploration of Responsible Factors and Consequences. *A Journal of the BSA MedSoc Group, 7*(1) 20-26.
8. Fogg, B. J. (2003). Persuasive Technology: Using Computer to Change What we Think and Do. Morgan Kaufmann Publishers. 25-29.
9. Fogg, B.J., Cuellar, G., & Danielson, D. (2007). Motivating, Influencing and Persuading Users. In Sears, A. & Jacko, J. A. (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2nd Ed., pp 133-144). Retrieved from https://pdfs.semanticscholar.org/6f20/ab660d5b72f7142a51f3c479809a971b7c5f.pdf
10. Harvard School of Public Health. (2015). *Alcohol: Balancing Risks and Benefits*. Retrieved from http://www.hsph.harvard.edu/nutritionsource/alcohol-full-story/
11. IJsselsteijn, W., de Kort, Y., Midden, C., Eggen, B. & van den Hoven, E. (2006). *Persuasive Technology for Human Wellbeing: Setting the Scene.* Paper presented at the First International Conference on Persuasive Technology for Human Well-Being, PERSUASIVE May 18-19, 2006 Eindhoven, The Netherlands. Retrieved from http://www.elisevandenhoven.com/publications/ijsselsteijn-persuasive06.pdf.
12. Khaled, R., Barr, P., Noble, J., Fischer, R. & Biddle, R. (2006). *Persuasive Technology for Human Wellbeing: Setting the Scene.* Paper presented at the First International Conference on Persuasive Technology for Human Well-Being, PERSUASIVE 2006 Eindhoven, The Netherlands. Retrieved from http://www.elisevandenhoven.com/publications/ijsselsteijn-persuasive06.pdf.
13. Larson, J. (2014). The Invisible Manipulative Power of Persuasive Technology. *Pacific Standard.* Retrieved from https://psmag.com/environment/captology-fogg-invisible-manipulative-power-persuasive-technology-81301
14. Lehto T. & Qinas-Kukkonen H. (2009). *The Persuasiveness of Web-based Alcohol Interventions: International Federation for Information Processing*, 316-327
15. Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task

motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705-717.

16. Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science, 15*(5), 265-268.
17. Lunenburg, F. C. (2011). Goal Setting Theory of Motivation. International Journal of Management, Business and Administration 15(1), 1-5.
18. Noar, S.M., Benac, C.N., & Harris, M.S. (2007) Does Tailoring Matter? Meta-Analytic Review of Tailored Printrint Health Behavior Change Interventions. *Psychological Bulletin, 4*, 673-693
19. Orji, R., Vassileva J., & Mandryk R. L. (2012). *Lunchtime: A Slow Casual Game for Long-Term Dietary Behaviour Change.*
20. Pietrangelo A. & Holland K. (2017). The Effects of Alcohol on Your Body. *Healthline.* Retrieved from https://www.healthline.com/health/alcohol/effects-on-body#1
21. Rashotte, L. (2015). *Social Influence.* Retrieved from http://www.sociologyencyclopedia.com/fragr_image/media/social
22. Smart, L. (2007). *Alcohol and Human Health*. Oxford: Oxford University Press.
23. Social Issues Research Centre (2019) Social and Cultural Aspects of Drinking. Retrieved from https://www.myvmc.com/lifestyles/alcohol-the-health-social-and-financial-burden/
24. Taylor, E. W. (2000). Analyzing Research on Transformative Learning Theory. In: Mezirow, J. (ed), *Learning as Transformation: Critical Perspectives on a Theory in Pogress*. Jossey-Bass, San Francisco, pp 29–310.
25. Wigmore, I. & Rouse, M. (2016). *Reinforcement Theory*. TechTarget. Retrieved from http://whatis.techtarget.com/definition/reinforcement-theory.
26. Willis, J. (2006). Drinking Crisis: Change and Continuity of Drinking in Sub-Saharan Africa. *African Journal of Drugs and Alcohol Studies, 5*, 1-15.
27. Xu, J., Chen, P-y., Uglow, S., Scott. A., & Montague, E. (2011). *A Case Study of the Design and Evaluation of a Persuasive Healthy Lifestyle Assistance Technology:*
28. *allenges and Design Guidelines.* Universe Access Hum 6768:464–471

International Journal of Computer Applications Technology and Research
Volume 9–Issue 06, 211-216, 2020, ISSN:-2319–8656

Multidimensional and Spatio-Temporal Design Pattern for Decision-Support Systems Schema Generation

Apollinaire Bamana Batoure
University Institute of Technology
The University of Ngaoundere
Ngaoundere, Cameroon

Kolyang
Higher Teachers' Training College
The University of Maroua
Maroua, Cameroon

Abstract: Design patterns are frequently used as a generic solution in software development, as well as in information systems, because they bring several improvements. In this work, we propose a design pattern that will be used for the generation of data schemas of spatio-temporal decisional systems. This use is done within the framework of the application of the Multidimensional Canonical Partitioning (MCP) approach, used for the design of decision-making systems.

Keywords: Design Pattern, Multidimensional Canonical Partitioning, Spatio-temporal decision system, UML profile, Archetype

1. INTRODUCTION

Patterns are among the reuse techniques that have been widely used in the design of information systems. They are presented as solutions to recurring problems [1, 2, 3, 4]. They help to reduce the production costs of applications, while ensuring better quality. They are also used as models.

A design pattern is an approved solution to a recurring design problem within a defined context [5, 31]. A Multidimensional Pattern (MP) is a design meta-model, representing a complete and generic star schema, in a domain of the company's activity, built from standard real-world entities [6, 7]. Their construction is based on empirical studies. It facilitates the task of the architect or designer of the system, by providing solutions to similar situations already experienced.

Design patterns are defined as abstract (e.g. specified in UML) and concrete (e.g. implementation archetypes) forms of general use, whose adaptation to a special development context is technically straightforward and cost-effective [8]. The abstract form can also be defined in formal language as B [9].

Design patterns represent expert knowledge, in a form that can be reused by the less expert. Because of their generic nature, they are considered microarchitectures that aim to reduce complexity, promote reuse and provide a common vocabulary for designers [10].

The purpose of using multidimensional patterns is to provide a generic solution for a domain decision problem that allows the design of data stores. It must be understandable and adaptable to the needs of the decision-makers, comprehensive, well-trained and produce the expected results. It thus reduces the time and cost of design, while preparing the way for implementation. The approach is platform-independent because it avoids, as much as possible, technological aspects that are characterized by continuous and rapid evolution [8, 11].

The research activities carried out by Jamel FÉKI's team[1] allow us to understand that the problems that have encouraged the emergence of multidimensional design patterns are:

- the fact that the design always starts from scratch, even in areas where the needs are known, and the solutions are mastered;
- the lack of capitalization of the know-how of data warehousing experts;
- the lack of reusable decision support system design approaches, and;
- the absence of techniques to assist in the specification of OLAP (OnLine Analytical Processing) requirements.

There are three types of patterns:

- creative patterns, which concern the creation of classes or objects;
- structural patterns, are interested in the composition of objects or classes to realize new functionalities;
- and behavioural patterns concern class interactions and the assignment of responsibilities [10, 12, 13].

There are three types of approaches in the use of patterns, similar to those used in the design of decision-making systems [14]. These are the bottom-up, top-down and mixed approaches [1, 8, 11].

Also, three groups of patterns exist [15, 16]. These are the:

1- Decision-making information systems team, Mir@cl Laboratory, FSEG Sfax.

- analysis patterns, which are used in the definition phase of system needs or requirements;
- design patterns, that define the abstract aspects of the system;
- and implementation patterns specified in a programming language for code generation.

In the present work, we specify a bottom-up, structural design pattern to generate multidimensional and spatio-temporal data schemas.

The purpose of this article is to introduce this pattern. For it, after this introduction where we have defined and given the interest of design patterns, we will present some cases of use in decisional systems. Then, we will bring up the MCP approach, which is the application framework of this pattern. Finally, before the conclusion, we present the profile of the archetype of the proposed pattern.

2. USE OF PATTERNS IN THE DESIGN OF DECISION-SUPPORT SYSTEMS

Several works on Decision Support Systems (DSS) take into account design patterns.

In [17] documents circulating within the company are used as Real World Entities (RWE). An association between these RWE and the Multidimensional Pattern (MP) is then made. These documents are classified on the basis of an empirical study. The proposed MP is made up of elements grouped into three categories (important, recommended and optional), distinguished on the basis of a graph. The pre-instantiation and instantiation of the MP entities is done at the logical level. Operators are defined for the instantiation of the MP. After this step, the elements that should actually make up the final data store are selected. At the physical level, relational views defined from heuristics are implemented. In order to be exhaustive, the patterns are overloaded. The application of the approach is made in the commercial domain and a tool that accompanies the approach is proposed.

In [18, 19], Feki J. et al. describe the application called Conception Assistée de Magasins de données en Étoiles (CAME) which implements the approach they propose. Emphasis is placed on the automation of the conceptual phase. The approach is divided into three main steps:

- pre-construction from source data schemas;
- automatic construction of the data store schemas to extract multidimensional concepts;
- and their validation by the designer.

Dimensions, facts and hierarchies are extracted from relational sources using heuristics.

The approach presented by Y. Hachaichi et al. in [6] is also declined in three steps, namely:

- collection and standardization of RWE;
- their classification into factual and basic entities;
- and the of DSS Schema.

The application of the approach is made in the medical field. Patterns are specifically used at the logical level, where the designer adapts the multidimensional pattern to the analytical needs. A correspondence is then made with the real-world entities, in order to propose a data store schema that best meets the business requirements.

Jones M. E. and Song I. Y. in [16] propose to set up a decision-making information system by using patterns, as early as the needs analysis phase. This approach is derived from the Who, What, Where, When, How and Why Quality Principle (QQOQCP). The approach is declined in two steps. In the first step, questions relating to the six domain patterns are asked in order to determine the associated dimensions. In the second step, questions are asked to determine the attributes associated with each dimension.

Annoni's thesis [2] proposes a tool to facilitate rapid development by reusing patterns. Patterns are grouped in a catalogue that capitalizes the development approach. The proposed tool is called electronic Business Intelligence Patterns for Analysis and Design (eBIPAD) and is dedicated to system administrators and business intelligence designers. The approach takes into account the tactical and strategic needs of decision-makers and then, the data sources to represent them intuitively.

In the present work, we use a multidimensional and spatio-temporal design pattern to derive the data schema of a decision system, from the elements of the multidimensional annotation obtained by applying the first five steps of the Multidimensional Canonical Partitioning (MCP) approach. The proposed multidimensional pattern is a continuation of the decision system design approach defined in [20] and will be used within the general framework of a model-driven architecture (MDA). This architecture makes it possible to implement all the phases (conceptual, logical and physical) of decision system modelling. The proposed multidimensional pattern is spatio-temporal in order to better understand, implement and exploit the studied phenomena.

In order to better understand the position of the proposed pattern, the following section provides a brief reminder of the MCP approach for designing and implementing decision-making systems.

3. MULTIDIMENSIONAL CANONICAL PARTITIONING APPROACH (MCP)

In [20], Batoure et al. propose a supply-driven approach, called Multidimensional Canonical Partitioning (MCP). It takes into account, without distinction, Entity/Relationship (ER) data schema. From universal relations assumption, the schema is derived into a universal relation (UR). This step provides a flat schema, where all features are grouped into a single entity. This entity is then partitioned vertically, according to characteristics or attributes semantics. To achieve it, we use a heuristic greedy type algorithm. Resulting partitions are candidates for being dimensions in the future data schema. To do this, we use an algorithm that matches the attributes present in the partitions and those that must actually be in the dimensions. Obtained dimensions are, if necessary, snow-flaked (normalized), using the third normal form algorithm (3NF). Because all the multidimensional elements are obtained, the data schema is generated using model transformations from QVT (Query View Transformation) language and a multidimensional and spatio-temporal design pattern.

The approach is recapped into six steps:

1. Verification of provided ER schema. If it is on universal relation form, we go straight to step 2, otherwise we restructure it according to universal relation assumption;

2. Vertical partitioning by fragmentation and distribution of attributes in obtained partitions;

3. Transformation of partitions into dimensions;

4. Normalization of dimensions;

5. Construction of facts table;

6. Generation of multidimensional schema.

The proposed multidimensional pattern is used in step six as a model, to transform by QVT, multidimensional elements to a data warehouse schema.

In Figure 1 and in step six of the approach, the elements of the multidimensional annotation, the design pattern and the QVT transformations are put together to obtain a specific decision data schema. The model transformations will be the subject of another work.

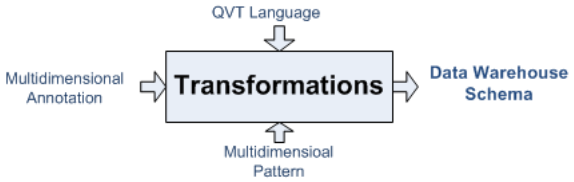

Figure. 1 Design Pattern for Approach Stage Six

We are now going to present the proposed pattern.

4. MULTIDIMENSIONAL AND SPATIO-TEMPORAL DESIGN PATTERN

The goal in this work is to propose a multidimensional and spatio-temporal design pattern, to generate data schemas for decision systems. In order to better exploit this design pattern, it is defined in a formal way by a profile, and in a concrete way by an archetype.

4.1 Design Pattern Profile

For this, we use the Unify Modelling Language (UML), which is one of the Object Management Group (OMG) standards [21]. UML enables the definition of the structural and behavioural aspects of a system via a set of models. These aspects explain the structure of a system thanks to attributes, and their behaviours thanks to operations, through classes [11, 22, 23]. The use of these models reduces the effort of the designer in learning new models and methods. Indeed, class models are unsuitable for multidimensional modelling. To overcome this deficiency and to make the semantics of multidimensional concepts explicit, we propose a UML profile.

A profile is a set of mechanisms and techniques for adapting the UML modelling language to a specific application domain. From a technical point of view, a profile is a set of stereotypes. A stereotype is defined as a domain-specific concept [24, 25]. The profile thus enables the extension of the UML meta-model. A meta-model defines a domain-specific modelling language and a profile defines a dialect or a variant of this language [15]. UML profiles involve the central concept of stereotype. It is described by a meta-class or package and several stereotypes.

In the proposed profile, stereotypes are used to define the overall data pattern, facts and their measures, then dimensions and their hierarchies. A measure in the fact table can be temporal, spatial, spatio-temporal or thematic. On the other hand, a dimension can be spatial, temporal or thematic. The dimension is described by hierarchies that follow the same characteristic as itself. The term thematic is used to describe all objects other than temporal and spatial objects [27, 28]. The spatio-temporal profile in Figure 2 has been designed using the Software Engineering Workshop (SEW) Perceptory pictograph language, also known as the "repository of user perceptions" [28, 29]. This SEW follows the UML formalism.

The proposed meta-model includes, in addition to the meta-class that defines the specific language package, four main stereotypes. These are the stereotypes on:

- the overall data schema of the decision-making system;
- the facts table (central table of the data schema) which groups together the different measures. These measures are the elements of decision-making analysis. To this main table are linked the dimension tables;
- the measures are the attributes of the fact table. They are either thematic, spatial, temporal or spatio-temporal;
- the dimensions are the tables connected to the fact table. They are either thematic, spatial or temporal. The dimensions can have granularities or hierarchies in order to give more details on the entities they describe. In this case, the global data schema is of the snowflake type. These granularities are of the same type as the dimensions they describe.

The spatial and temporal dimensions, as well as their granularity, are automatically added because the phenomena studied are defined in time and space references.

The abstract form of the design pattern allows its concrete form to be derived.

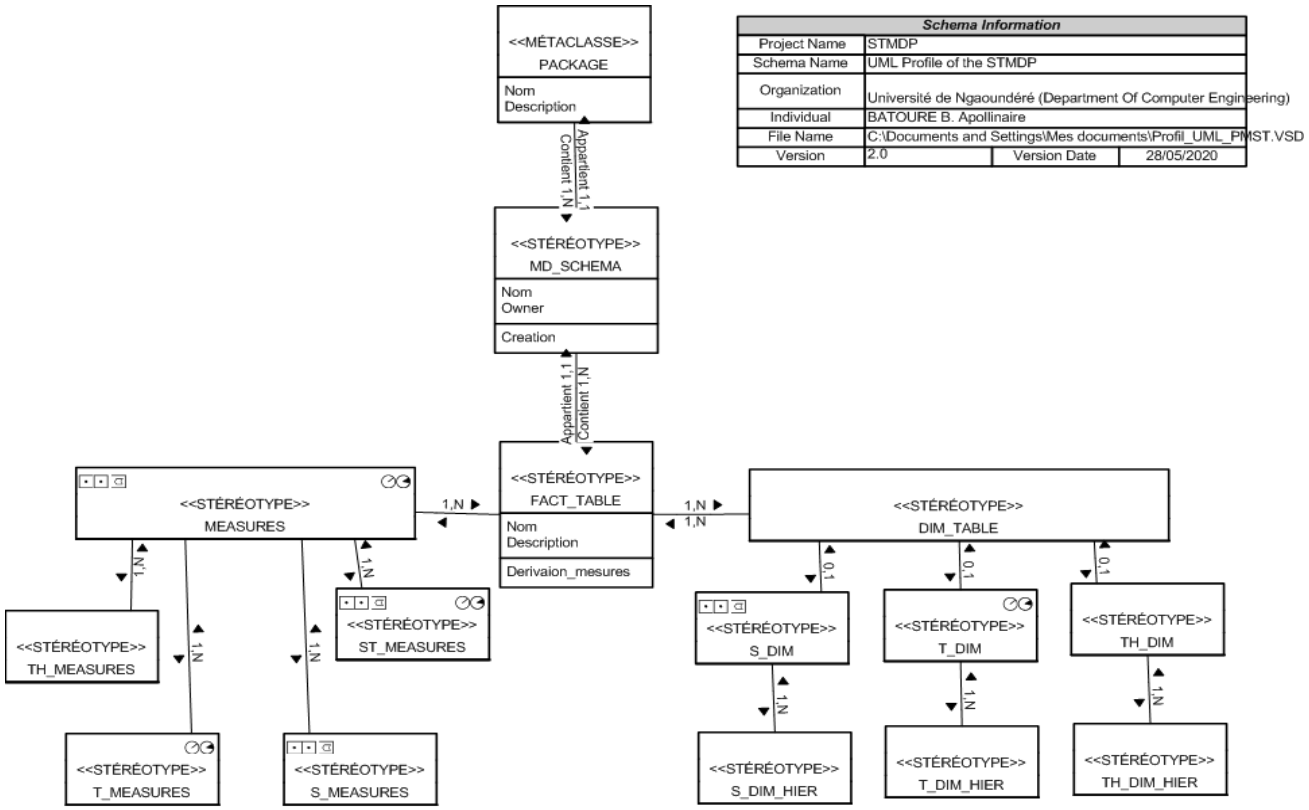

Figure. 2 Design Pattern Profile

4.2 Pattern Archetype

The concrete shape of the profile of the multidimensional design pattern allows us to define an archetype (Figure 3). It is a generic pattern obtained from the meta-model. This makes the definition of the multidimensional pattern concrete as being a standard solution, in a given domain [6, 30].

This archetype integrates the temporal and spatial dimensions as well as their granularities. It also includes thematic dimensions, the facts table and measures. The archetype can therefore be instantiated according to a specific problem from the real world. This instantiation is based on the elements of multidimensional annotation obtained in the previous steps of the MCP approach, and on models' transformation.

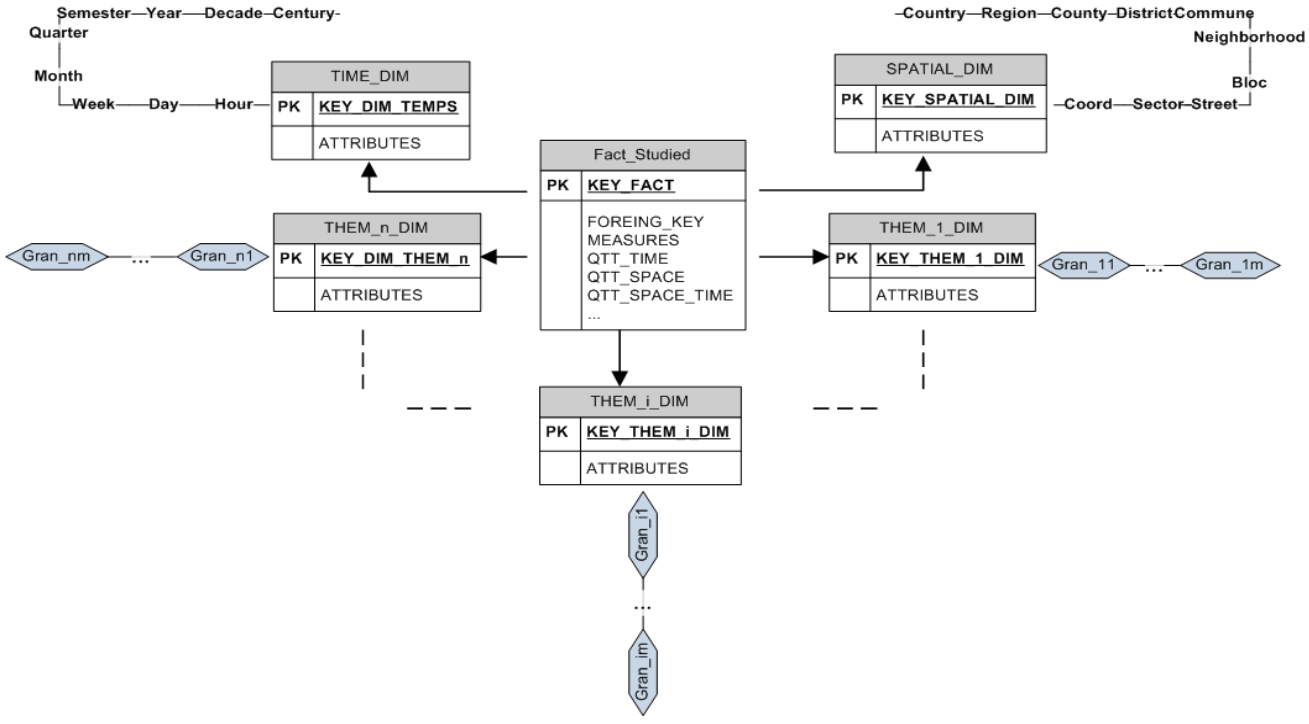

Figure. 3 Generic Data Scheme or Archetype

5. CONCLUSION

In this work, we have presented a multidimensional pattern needed for decision systems design. After defining and giving uses cases of patterns in the design of decision systems, we made a reminder of the MCP approach, which is the framework for the use of this proposal. Finally, we presented the abstract form (UML profile) and the concrete form (archetype) of this design pattern. This result, which is a continuation of the MCP decision systems design approach previously proposed and will be used within the general framework of a model-driven architecture. This architecture will enable the implementation of all modelling phases (conceptual, logical and physical) of decision system.

6. REFERENCES

[1] Lacob, M.-E., Steen, M. W. A. and Heerink, L. 2008. Reusable Model Transformation Patterns, in: Enterprise Distributed Object Computing Workshop, IEEE Computer Society, Munich, Germany, pp. 1–10.

[2] Annoni E. 2007. Éléments méthodologiques pour le développement des systèmes décisionnels dans un contexte de réutilisation, Ph.D. thesis in Computer Science, Université Paul Sabatier, Toulouse III.

[3] Gamma, E., Helm, R., Johnson, J. and Vlissides, J. 2015. Design Patterns : Element of Reusable Object Oriented Software. Addison-Wesley.

[4] Zimmer, W. 1995. Relationships between Design Patterns. In Design, volume 1, pages 345–364. ACM Press/Addison-Wesley Publishing.

[5] Miralles, A. 2006. Model Engineering for Environmental Applications, Thesis for the PhD in Computer Science at the University of Montpellier II.

[6] Hachaichi, Y., Ben-Abdallah, H. and Feki, J. 2008. Towards an MDA approach to data warehouse development, Information Systems and Business Intelligence, 158-170pp.

[7] Feki J. and Hachaichi Y. 2007. MD Assisted Design: An approach and a tool. Journal of Decision Systems vol 16/3, 303- 333.

[8] Cho, H. and Gray, J. 2011. Design Patterns for Metamodels, in: SPLASH ’11 DSM Workshop, Portland, OR, USA, pp. 25–32;

[9] Marcano-Kamenoff R., Levy N. and Losavio F. 2000. " Spécification et spécialisation de patterns en UML et B " Langages et modèles à objets, LMO'2000

[10] El Boussaidi, G. 2009. Développement logiciel par transformation de modèles, Thèse en vue de l'obtention de grade de Philosophiæ Doctor (Ph.D.) en Informatique.

[11] Lano, K., and Kolahdouz R. S. 2014. Model-Transformation Design Patterns, IEEE Transactions on Software Engineering 40 (12) 1224–1259.

[12] Ergin, H., Syriani, E. and Gray, J. 2016. Design Pattern Oriented Development of Model Transformations, Computer Language, Systems and Structures.

[13] Blazy, S., Gervais, F. and Laleau, R. 2003. Une démarche outillée pour spécifier formellement des patrons de conception réutilisable, Workshop OCM-SI.

[14] Batoure, A. B. and Kolyang. 2018. Designing Decision Support Systems Approaches using Entity/Relationship Data Schema : A Survey. International Journal of Scientific & Engineering Research (IJSER), Volume 9, Issue 9.

[15] Sabri, A. and Kjiri, L. 2012. Patterns to analyze requirements of a Decisional Information System, Workshop AIde à la Décision à tous les Étages, AIDE, 12th International Francophone Conference on Knowledge Extraction and Management EGC Bordeaux.

[16] Jones, M. E. and Song, I.-Y. 2005. Dimensional modeling: identifying, classifying & applying patterns. In Song and Trujillo. 29–38.

[17] Feki, J., Ben-Abadallah, H. and Ben Abdallah, M. 2006. Reuse of star patterns. XXIVth INFORSID'06 Congress, pp. 687-701, Hammamet, Tunisia.

[18] Feki, J. and Hachaici, Y. 2007. MD Assisted Design. Journal of Decision Systems, Vol 16 - N° 03, pp 303-333.

[19] Ben Abdallah M., Feki J. and Ben-Abdallah H. 2006. MPI-EDITOR: a tool for specifying OLAP needs by logical reuse of multidimensional patterns. Decision-making Systems Workshop (ASD'06), Morocco.

[20] Batoure, A. B., Kolyang, and Tchotsoua, M. 2017. Using Multidimensional Canonical Partitioning (MCP) as a Supply-Driven Approach for Data Warehouses Design. International Journal of Computer (IJC), 25(1):52–62.

[21] Boulil, K. 2012. Une Approche Automatisée basée sur des Contraintes d'Intégrité définies en UML et OCL pour la Vérification de la Cohérence Logique dans les Systèmes SOLAP - Applications dans le domaine agri-environnemental. Thèse soutenue en vue de l'obtention de l'titre de Docteur en informatique de l'Université de Blaise Pascal - Clermont-Ferrand II.

[22] Atigui, F., Ravat, F., Teste, O. and Zurfluh, G. 2010. Démarche dirigée par les modèles pour la conception d’entrepôts de données multidimensionnelles. In IRIT (SIG/ED).

[23] Laplanche, F. 2002. Conception de projet SIG avec UML. Bulletin de la Société géographique de Liège, 42 :19–25.

[24] Kadima. H. 2005. MDA : Conception Orientée Objet Guidée Par Les Modèles. Dunud, Paris.

[25] Zubcoff J. and J. Trujillo. 2007. A UML 2.0 profile to design Association Rule mining models in the multidimensional conceptual modeling of data warehouses. Data and Knowledge Engineering, 63(1) :44–62.

[26] Aissi S. and Gouider, M. S. 2012. Spatial and Spatio-Temporal Multidimensional Data Modelling : A survey. Internatioanal Journal of Advanced Research in Computer Science and Software (IJARCCE), 1(1) :1–8.

[27] Malinowski E. and Zimanyi E. 2004. Representing spatiality in a conceptual multidimensional model. In Proceedings of the 12th annual ACM international workshop on Geographic information systems - GIS ’04.

[28] Khalfi, B., De Runz, C., Faiz, S., and Akdag, H. 2015. Modélisation conceptuelle d’objets géographiques imprécis et multiples : une approche basée F-Perceptory. In Spatial Analysis and GEOmatics conference, SAGEO, pages 1–15.

[29] Zoghlami, A., De Runz, C., and Akdag, H. 2015. F-Perceptory : an approach for handling fuzziness of spatiotemporal data in geographical databases. International Journal of Spatial, Temporal and Multimedia Information Systems (IJSTMIS), 1(1) :30–62, 2015.

[30] Ben-Abdallah, H., Feki, J., and Ben-Abdallah, M. 2009. A multidimensional pattern based approach for the design of data marts', Progressive Methods in Data Warehousing and Business Intelligence: Concepts and Competitive Analytics, 23 January, pp.172–192.

[31] Sánchez Cuadrado J., Guerra E., and de Lara J. 2011. Generic Model Transformations: Write Once, Reuse Everywhere. In: Cabot J., Visser E. (eds) Theory and Practice of Model Transformations. ICMT 2011.