Al-Ahgaff University Journal of Computer Science and Mathematics
مجلة جامعة الأحقاف لعلوم الحاسوب والرياضيات

Vol. 1, Issue. 1, jan 2023, pp. 1~7

# A Numerical Method Based on the NB1-Ball Polynomial for Solving a Class of Linear and Nonlinear Differential Equations

**Ahmed Kherd**
Department of Mathematics, Al-Ahgaff University, Mukalla, Yemen
ahmedkherd@ahgaff.edu

**Article Info**

***Article history:***

Received des, 12, 2022
Revised jan, 10, 2023
Accepted jan, 25, 2023

***Keywords:***

NB1-Ball Polynomial
Operational Matrix
Differential Equation

**ABSTRACT**

In this article, NB1-Ball polynomials method for solving first and second orderordinary differential equation is proposed. Dealing with nonlinear and linear equations generated through matrix operation by simple form is the advantage of the suggested method. In order to show the performance of the proposed method, some real-life problems which include linear and nonlinear form of first and second order ordinary differential equations are introduced. The generated results confirm that the developed method outperform the existing method in terms of error.



**الخلاصة**

في هذه المقالة ، تم اقتراح طريقة NB1-Ball متعددة الحدود لحل المعادلة التفاضلية العادية من الرتبتين الأولى والثانية. يعتبر التعامل مع المعادلات الخطية وغير الخطية الناتجة عن عملية المصفوفة بشكل بسيط ميزة الطريقة المقترحة. من أجل إظهار أداء الطريقة المقترحة ، تم تقديم بعض مشاكل الحياة الواقعية التي تشمل الشكل الخطي وغير الخطي للمعادلات التفاضلية العادية من الدرجة الأولى والثانية. تؤكد النتائج المتولدة أن الطريقة التي تم تطويرها تتفوق على الطريقة الحالية من حيث الخطأ.

## 1. INTRODUCTION

In this paper, we will present an efficient method for computing the numerical solution of differential equations (DEs). Problems of the type (9) and (10) have been considered by a vast number of scientific research fields, spanning from the chemical to the physical sciences and their applications to geophysics, reaction diffusion processes, and gas equilibrium, amongst a great number of other topics. As a result of the broad range of applications for problems of the kind under discussion, it is preferable to find a precise or approximate solution for the problem, which has been investigated by a large number of researchers. Nasab and Kilicman [1] used the wavelet analysis approach in order to solve linear and nonlinear initial (boundary) value problems. The Legendre operational matrix was used by Bataineh and Ishak Hashim [2] in order to come up with an approximation of the solution to two-point point boundary value issues. Bhatti [3] made use of the well-known Bernstein polynomial basis in order to find an approximate solution to the differential equation. Youseffi provided an approximate solution to the Bessel differential equation, and Yuzbasi also solved the fractional riccati type (DEs) [5], following the work of Bhatti, Pandey, and Kumar [4] and Isik and Sezer, who were able to get an analytic solution to the Lane-Emden type equations. In a recent paper, Yiming Chen used Bernstein polynomials in a similar way to turn up at the numerical solution to the variable order linear cable equation [6]. Rostamy also used a similar strategy to solve the backward inverse heat conduction problems [7], but he employed a modified operational matrix approach. Similar to the previous article, This paper likewise adopted the use of the NB1-Ball operational matrix to find solutions to issues involving linear and nonlinear starting (boundary) values. From the numerical answers produced, it is evident that there is commendable precision and reduced computing weight, as compared to the precise solution within a range of no more than 10 digits, just a few NB1-Ball polynomial basis functions are needed to get this approximative solution. This article is organised as follows: Section 2 discusses a review of Ball polynomials and NB1-Ball polynomials, as well as the conventional derivation of NB1-Ball polynomials and the differentiation of its operational matrix, while

Section 3 discusses applications of the operational matrix of the derivative. The numerical results are presented in Section 4, together with the precise solution, and the operational matrices validity, precision, and application are ultimately justified. Section 5 offers a succinct overview and conclusion.

## 2. REVIEW ON BALL POLYNOMIAL

The Ball polynomial was declared by A. A. Ball in his well-known aircraft design system CONSURF in [1]. It is described as a cubic polynomial and explained mathematically as:

$$(1-z)^2, 2z(1-z)^2, 2z^2(1-z), z^2, \quad 0 \le z \le 1. \tag{1}$$

In further research, several studies have discussed about Ball polynomial's high generalization and its properties. For instance, in the 1980s there were two different Ball polynomials of arbitrary degree are called Said-Ball and Wang-Ball [2, 3] and in 2003 there was another generalization of Ball polynomial called DP-Ball [4].

### 2.1. Nb1-Ball Polynomial Representation

Definition:

For any integer $n \geq 3$, the NB1 basis of degree $n$ is defined as [5]

$$NB1_{i=0}^{n} = \begin{cases} \binom{\left\lfloor \frac{n}{2} \right\rfloor - 1 + i}{i} z^i (1-z)^{\left\lfloor \frac{n}{2} \right\rfloor}, & 0 \le i \le \left\lfloor \frac{n}{2} \right\rfloor - 2, \\ \binom{2\left\lfloor \frac{n}{2} \right\rfloor - 2}{\left\lfloor \frac{n}{2} \right\rfloor - 1} z^{\left\lfloor \frac{n}{2} \right\rfloor - 1} (1-z)^{\left\lfloor \frac{n}{2} \right\rfloor + 1}, & i = \left\lfloor \frac{n}{2} \right\rfloor - 1, \\ 2\binom{2\left\lfloor \frac{n}{2} \right\rfloor - 2}{\left\lfloor \frac{n}{2} \right\rfloor - 1} z^{\left\lceil \frac{n}{2} \right\rceil} (1-z)^{\left\lceil \frac{n}{2} \right\rceil}, & i = \left\lfloor \frac{n}{2} \right\rfloor, \\ NB1_{n-i}^{n}(1-z), & \left\lfloor \frac{n}{2} \right\rfloor + 1 \le i \le n. \end{cases} \tag{2}$$

Definition:

The NB1 basis function can be formulated in power basis form by [5]

$$\mathcal{C}^n(z) = \sum_{i=0}^{n} \sum_{j=0}^{n} b_{ij} z^j \tag{3}$$

where

$$b_{ij} = \begin{cases} (-1)^{(j-i)} \binom{\left\lfloor \frac{n}{2} \right\rfloor - 1 + i}{i} \binom{\left\lfloor \frac{n}{2} \right\rfloor}{j-i}, & \text{for } 0 \le i \le \left\lfloor \frac{n}{2} \right\rfloor - 2, \\ (-1)^{(j-i)} \binom{2i}{i} \binom{i+2}{j-i}, & \text{for } i = \left\lfloor \frac{n}{2} \right\rfloor - 1, \\ (-1)^{(j-i)} 2 \binom{2i-2}{i-1} \binom{n-i}{j-i}, & \text{for } i = \left\lfloor \frac{n}{2} \right\rfloor, \\ (-1)^{(j-i)} 2 \binom{2(n-i-1)}{n-i-1} \binom{n-i}{j-i}, & \text{for } i = \left\lceil \frac{n}{2} \right\rceil, \\ (-1)^{(j-n+i)} \binom{2(n-i)}{n-i} \binom{n-i}{j-n+i-2}, & \text{for } i = \left\lceil \frac{n}{2} \right\rceil + 1, \\ (-1)^{(j-\left\lfloor \frac{n}{2} \right\rfloor)} \binom{\left\lfloor \frac{n}{2} \right\rfloor - 1 + n - i}{n-i} \binom{n-i}{j-\left\lfloor \frac{n}{2} \right\rfloor}, & \text{for } \left\lceil \frac{n}{2} \right\rceil + 2 \le i \le n. \end{cases} \tag{4}$$

## Definition

The monomial matrix form for NB1-Ball can be specified as [6]

$$\mathcal{N} = \begin{bmatrix} g_{00} & g_{01} & \cdots & g_{0n} \\ g_{10} & g_{11} & \cdots & g_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{n0} & g_{n1} & \cdots & g_{nn} \end{bmatrix}_{(n+1)\times(n+1)} \tag{5}$$

where $g_{ij}, \quad i,j = 0,1,\cdots,n$ are given as (2).

In general, we approximate any function u(t) with the first m+1 NB1-Ball polynomials as:

$$y(z) \approx \sum_{i=0}^{m} c_i \mathcal{N}_i^m(z) = C^T \phi(z) = C^T \mathcal{N} T(z). \tag{6}$$

where $C^T = [c_0\ c_1\ c_2 \cdots c_m], H(z) = [1\ \mathrm{z}\ \mathrm{z}^2 \cdots z^m]^T$ and $\mathcal{N}$ is the monomial matrix form was given in (5). The operational matrix of derivative of the NB1- Ball polynomials set $\psi(z)$ is given by

$\frac{d\psi(z)}{dz} = D^{'(1)}\psi(z)$ is the $m+1$by $m+1$operational matrix of derivative define as

$$D^{'(1)} = \mathcal{N}\Lambda\mathcal{N}^{-1} \tag{7}$$

where $\mathcal{N}$is NB1-Ball monomial matrix form given in (5), and

$$\Lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & m & 0 \end{bmatrix} \tag{8}$$

We can generalized Equation (8) as

$$D^{'(n)}\psi(z) = D^{'(n-1)}\left(D^{'(1)}\psi(z)\right) = \cdots = (D^{'(1)})^n\psi(z) = D^{'(n)}\psi(z), m = 1,2,\ldots$$

## 3. Applications of the Operational Matrix of Derivative

We present in this section the derivation of the method for solving differential equation of the form

$$q_0(z)u''(z) + q_1(z)u'(z) + q_2(z)(u(z))^n = g(z) \tag{9}$$

with initial conditions (ICs) or boundary conditions (BCs)

$$\begin{cases} u(0) = 1, \quad y'(0) = 0, \quad \text{or} \\ u(0) = \alpha_1, \qquad u(1) = \alpha_2. \end{cases} \tag{10}$$

where $q_j(z), \quad j = 0,1,2$ and $g(t)$ are given, while $u(z)$ is unknown. We can write the residual $\square_n(z)$ as

$$\begin{aligned} \square(z) \quad = \quad & q_0(z)C^T D^{'(2)}\psi(z) + q_1(t)C^T D^{'(1)}\psi(z) \\ & + q_2(z)(C^T\psi(z))^n - G^T\psi(z) \end{aligned} \tag{11}$$

where $G^T = [g_0, g_1, \cdots, g_m]$, To find the solution of $u(z)$ given in (10), we first collocate (12) at $m-1$ points. For suitable collection points, we use

$$z_i = \frac{1}{2} + \frac{1}{2}cos\left(\frac{(2i+1)\pi}{n}\right) \quad , i = 0,1,\cdots,m-1. \tag{12}$$

Theses equations together with (11) generate $m+1$ nonlinear equations which can be solved using Newton's iteration method. Consequently, $u(z)$ can be calculated.

## 4. NUMERICAL EXAMPLES

### 4.1. Example. 1

At the first we consider the example given in [6]

$$u''(z) + \frac{1}{z}u'(z) + u(z) = 4 - 9z + z^2 - z^3, \tag{13}$$

with BCs

$$u(0) = 0 \quad u(1) = 0. \tag{14}$$

Which has the exact solution is $u(z) = z^2 - z^3$.

To solve (13) and (14) we use our purposed method with $m = 3$. we apply (8) we have,

$$D^{'(1)} = \begin{bmatrix} -2 & -1 & -1 & 0 \\ 2 & -2 & -2 & 0 \\ 0 & 2 & 2 & -2 \\ 0 & 1 & 1 & 2 \end{bmatrix}, D^{'(2)} = \begin{bmatrix} 2 & 2 & 2 & 2 \\ -8 & -2 & -2 & 4 \\ 4 & -2 & -2 & -8 \\ 2 & 2 & 2 & 2 \end{bmatrix}. \tag{15}$$

Therefore, using (13) for (14), we obtain

$$-\frac{55}{64}c_0 - \frac{103}{128}c_1 + \frac{115}{128}c_2 + \frac{65}{64}c_3 - \frac{115}{256} \tag{16}$$

$$\frac{67}{64}c_0 + \frac{25}{128}c_1 - \frac{501}{128}c_2 + \frac{219}{64}c_3 + \frac{501}{256} \tag{17}$$

Now we use the (BCs) we have

$$c_0 = 0, \quad c_3 = 0. \tag{18}$$

Solve Equations (17), (18) and (19) we get $c_0 = 0, \quad c_1 = 0, \quad c_2 = \frac{1}{2}$ and $c_3 = 0$. Thus

$$[u_3(z)] = c_0\mathcal{N}_0^3(z) + c_1\mathcal{N}_1^3(z) + c_2\mathcal{N}_2^3(z) + c_3\mathcal{N}_3^3(z)$$

$$= \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} (z-1)^2 \\ 2z(z-1)^2 \\ -2z^2(z-1) \\ z^2 \end{bmatrix}$$

$$= [z^2 - z^3].$$

Which is the exact solution.

### 4.2. Example. 2

Consider the Bessel differential equation of order zero given in [7-10]

$$zu''(z) + u'(z) + zu(z) = 0, \tag{19}$$

with the ICs

$$u(0) = 1, \quad u'(0) = 0. \tag{20}$$

The exact solution of this example is

$$J_0(z) = \sum_{q=0}^{\infty} \frac{(-1)^q}{(q!)^2}\left(\frac{z}{2}\right)^{2q}. \tag{21}$$

Here we see that $g(t) = 0$ By the suggest method, we obtain the proximate solution when $m = 12$ is

$$u_{12} = 1.0 - 0.24999969t^2 - 0.00000712t^3 + 0.0156966t^4 - 0.0004143t^5 + 0.0010672128t^6$$

$$-0.003570235t^7 + 0.00565212t^8 - 0.00588660t^9 + 0.0038875t^{10} - 0.001473t^{11} + 0.0002435t^{12}.$$

The numerical results of our scheme together with two other [8, 9] are provided in Table 1

Table 1 Errors of the present method compared with results in ref [8, 9] for the Example 2.

| *t* | PM m=12 | Method of [9] for k=2, m=3, | Method of [8] for k=2, m=3 |
|---|---|---|---|
| 0.2 | 0 | 9.36e-05 | 6.01e-05 |
| 0.4 | 7.50e-11 | 2.78e-05 | 1.636e-04 |
| 0.6 | 3.24e-10 | 3.60e-05 | 1.636e-04 |
| 0.8 | 6.66e-09 | 2.695e-04 | 1.636e-04 |
| 1.0 | 1.66e-06 | 2.689e-04 | 1.636e-04 |

### 4.3. Example. 3

Consider the following ordinary differential equation [11]

$$u''(z) + zu'(z) + z^2u^3(z) = (2 + 6z^2)e^{(z^2)} + z^2e^{(3z^2)}, \quad (22)$$

Subject to IC

$$u(0) = 1, u'(0) = 0. \quad (23)$$

with the exact solution $u(t) = e^{t^2}$. We apply the above method when $m = 12$. Table 2 show the absolute error for Example 3.

Table 2 Errors of the present method compared with results in ref [11] for the Example 3.

| t | Ref [17] | PM |
|---|---|---|
| 0.000 | 0 | 0 |
| 0.010 | 0.2000000E-10 | 0.1683250E-10 |
| 0.020 | 0.2900000E-09 | 0.5781250E-10 |
| 0.030 | 0.2900000 E-09 | 0.1120996E-09 |
| 0.040 | 0.4450000 E-08 | 0.1725064E-09 |
| 0.050 | 0.1074000 E-07 | 0.2345360E-09 |
| 0.060 | 0.2207000 E-07 | 0.2956183E-09 |
| 0.070 | 0.4057000 E-07 | 0.3545185E-09 |
| 0.080 | 0.6872000 E-07 | 0.4108806E-09 |
| 0.090 | 0.1093000 E-06 | 0.4648881E-09 |
| 0.100 | 0.1654900 E-06 | 0.5170152E-09 |

### 4.4. Example. 4:

Consider the ordinary differential equation [11]

$$u''(z) + u(z)u'(z) = tsin(2z^2) - 4z^2sin(z^2) + 2cos(z^2), \text{z} \in [0,1], \quad (24)$$

with ICs $u(0) = 0, u'(0) = 0$.

Where the exact solution is $u(z) = sin(z^2)$. Table. 3 show the comparison the absolute error of our method with ref [11]

Table 3 Errors of the present method compared with results in ref [11] for the Example 4 with $m$ = 12

| $t$ | Ref [11] | PM |
|---|---|---|
| 0.0 | 0 | 0 |
| 0.1 | 3.074560E-7 | 7.249816E-9 |
| 0.2 | 1.058636E-5 | 1.483888E-8 |
| 0.3 | 5.114716E-5 | 2.254473E-8 |
| 0.4 | 1.331415E-4 | 3.054861E-8 |
| 0.5 | 2.420463E-4 | 4.135017E-8 |
| 0.6 | 3.299021E-4 | 5.825938E-8 |
| 0.7 | 3.231831E-4 | 9.301242E-8 |
| 0.8 | 1.540876E-4 | 1.668932E-7 |
| 0.9 | 1.870564E-4 | 3.322331E-7 |
| 1.9 | 6.088701E-4 | 3.322331E-7 |

### 4.5. Example. 5:

Consider the first order ode [11]

$$u'(z) - zu(z) + u^2(z) = e^{z^2} \quad (25)$$

subject to IC

$$u(0) = 1. \quad (26)$$

with the exact solution $u(z) = e^{\frac{z^2}{2}}$. The absolute error of Example .5 is presented in Table 4.

Table 4. Errors of the present method compared with results in ref [11] for the Example 5 with $m = 12$.

| $t$ | Ref [11] | PM |
|---|---|---|
| 0.00 | 0 | 0 |
| 0.01 | 1.750000 E-7 | 2.610000E-11 |
| 0.02 | 6.400000 E-7 | 3.195700E-10 |
| 0.03 | 1.314000 E-6 | 5.506400E-10 |
| 0.04 | 2.123000 E-6 | 5.936400E-10 |

| | | |
|---|---|---|
| 0.05 | 2.999000 E-6 | 4.440000E-10 |
| 0.06 | 3.883000 E-6 | 1.595400E-10 |
| 0.07 | 4.720000 E-6 | 1.791700E-10 |
| 0.08 | 5.463000 E-6 | 4.941400E-10 |
| 0.09 | 6.069000 E-6 | 7.240900E-10 |
| 0.1 | 6.501000 E-6 | 8.310500E-10 |

### 4.6. Example 6:

Finally Consider the following form of a singular Dirichlet-type boundary value problem on the interval [0, 1] [12]

$$u''(z) - \frac{1}{z}u'(z) + \frac{1}{z(1+z)}u(z) = -z^3, \tag{27}$$

with BCs

$$u(0) = 0, u(1) = 0. \tag{28}$$

where the exact Solution is

$$u(z) = \frac{1}{144(-1+2ln(2))}(14ln(z+1)t + 14ln(z+1) - 14z + 6z^2 - 12z^2ln(2) - 2z^3 + 4z^3ln(2) + z^4 - 2z^4ln(2) + 9z^5 - 18z^5ln(2)).$$

The absolute error in Table. 5 and in Figure. 1

Table 5 Errors of the present method compared with results in ref [12] for the Example 6 with $m = 11$.

| $t$ | Ref [12] | PM |
|---|---|---|
| 0.2 | 1.88415721000000E–10 | 1.37812040000000E-10 |
| 0.4 | 7.13501861405898E–10 | 1.03644900000000E-10 |
| 0.6 | 8.20803253396388E–10 | 5.55406000000000E-11 |
| 0.8 | 5.53448662985227E–10 | 3.68906000000000E-11 |

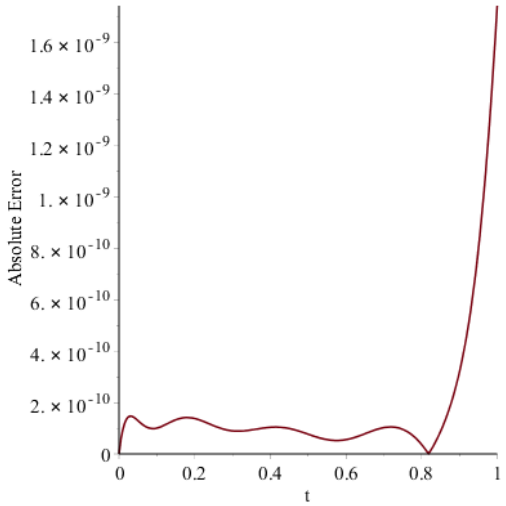


Figure 1. The absolute error for Example 6

## 5. CONCLUSION

In this work, the derivation of the new NB1-Ball polynomials method for solving first and second orders ODE is carried out. This new approach's capacity to solve second orders ODE is its most significant advantage over those that have been previously proposed. The ability of the method is shown in its application to non-linear and linear first and second orders IVP and ICs of ODEs. The generated results approve the supremacy of new Said-Ball polynomials method over existing methods in terms of error as offered in tables 1-5.

**ACKNOWLEDGEMENTS**

Author thanks Al-Ahgaff University.In most cases, sponsor and financial support acknowledgments.

**REFERENCES**

[1] A. K. Nasab, A. Kılıçman, E. Babolian, and Z. P. Atabakan, "Wavelet analysis method for solving linear and nonlinear singular boundary value problems," *Applied Mathematical Modelling*, vol. 37, pp. 5876-5886, 2013.

[2] A. S. Bataineh, A. Alomari, and I. Hashim, "Approximate solutions of singular two-point BVPs using Legendre operational matrix of differentiation," *Journal of Applied Mathematics*, vol. 2013, 2013.

[3] M. I. Bhatti and P. Bracken, "Solutions of differential equations in a Bernstein polynomial basis," *Journal of Computational and Applied mathematics*, vol. 205, pp. 272-280, 2007.

[4] S. A. Yousefi and M. Behroozifar, "Operational matrices of Bernstein polynomials and their applications," *International Journal of Systems Science*, vol. 41, pp. 709-716, 2010.

[5] Ş. Yüzbaşı, "Numerical solutions of fractional Riccati type differential equations by means of the Bernstein polynomials," *Applied Mathematics and Computation*, vol. 219, pp. 6328-6343, 2013.

[6] Y. Chen, L. Liu, B. Li, and Y. Sun, "Numerical solution for the variable order linear cable equation with Bernstein polynomials," *Applied Mathematics and Computation*, vol. 238, pp. 329-341, 2014.

[7] D. Rostamy and K. Karimi, "A new operational matrix method based on the Bernstein polynomials for solving the backward inverse heat conduction problems," *International Journal of Numerical Methods for Heat & Fluid Flow*, vol. 24, pp. 669-678, 2014.

[8] A. A. Ball, "CONSURF. Part one: introduction of the conic lofting tile," *Computer-Aided Design*, vol. 6, pp. 243-249, 1974.

[9] J. Delgado and J. M. Pena, "A linear complexity algorithm for the bernstein basis," presented at 2003 International Conference on Geometric Modeling and Graphics, 2003. Proceedings, 2003.

[10] S.-M. Hu, G.-Z. Wang, and T.-G. Jin, "Properties of two types of generalized Ball curves," *Computer-Aided Design*, vol. 28, pp. 125-133, 1996.

[11] N. Dejdumrong, "A new bivariate basis representation for Bézier-based triangular patches with quadratic complexity," *Computers & Mathematics with Applications*, vol. 61, pp. 2292-2295, 2011.

[12] C. Aphirukmatakun and N. Dejdumrong, "Monomial forms for curves in CAGD with their applications," presented at 2009 Sixth International Conference on Computer Graphics, Imaging and Visualization, 2009.

[13] Y. Ordokhani, "Approximate solutions of differential equations by using the Bernstein polynomials," *International Scholarly Research Notices*, vol. 2011, 2011.

[14] E. Babolian and F. Fattahzadeh, "Numerical solution of differential equations by using Chebyshev wavelet operational matrix of integration," *Applied Mathematics and computation*, vol. 188, pp. 417-426, 2007.

[15] R. K. Pandey and N. Kumar, "Solution of Lane–Emden type equations using Bernstein operational matrix of differentiation," *New Astronomy*, vol. 17, pp. 303-308, 2012.

[16] K. Parand, M. Dehghan, A. Rezaei, and S. Ghaderi, "An approximation algorithm for the solution of the nonlinear Lane–Emden type equations arising in astrophysics using Hermite functions collocation method," *Computer Physics Communications*, vol. 181, pp. 1096-1108, 2010.

[17] A. F. Qasim and E. S. Al-Rawi, "Adomian decomposition method with modified Bernstein polynomials for solving ordinary and partial differential equations," *Journal of Applied Mathematics*, vol. 2018, pp. 1-9, 2018.

[18] A. Secer and M. Kurulay, "The sinc-Galerkin method and its applications on singular Dirichlet-type boundary value problems," *Boundary Value Problems*, vol. 2012, pp. 1-14, 2012.

Al-Ahgaff University Journal of Computer Science and Mathematics
مجلة جامعة الأحقاف لعلوم الحاسوب والرياضيات

Vol. 1, Issue. 1, jan 2023, pp. 8~16

# An Empirical Evaluation for the Improved Model of Agile Kanban: Case Study Results

**Hamzah Alaidaros**
Faculty of Computer Science and Engineering, Al-Ahgaff University, Mukalla, Hadramaout, Yemen
hamzah@ahgaff.edu
**Mazni Omar**
Institute for Advanced and Smart Digital Opportunities (IASDO), School of Computing, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
mazni@uum.edu.my
**Rohaida Romli**
Human-Centred Computing Research Lab, School of Computing, UUM College of Arts and Sciences, Universiti Utara Malaysia 06010 UUM Sintok, Kedah, Malaysia
aida@uum.edu.my

**Article Info**

***Article history:***

Received des, 27, 2022
Revised jan, 20, 2023
Accepted jan, 25, 2023

***Keywords:***

Kanban Method
Monitoring Task
Case Study
Evaluation
Applicability
i-KAM

**ABSTRACT**

This paper aims to empirically evaluate the applicability of an improved model for the Agile Kanban, which is called i-KAM. The i-KAM was proposed to improve the software project monitoring task of Agile Kanban method, which has critical challenges negatively impact the process of Software Project Management (SPM). To achieve this aim, we conducted a case study in a software house to implement i-KAM in actual projecs within a real environment. In this case, a prototype tool designed based on i-KAM, has been used by team members during their daily work in developing new system for three months. Ultimately, data were collected via an evaluation form designed to measure the applicability of i-KAM based on five factors, which are (1) gain satisfaction, (2) interface satisfaction, (3) task support satisfaction, (4) perceived usefulness, and (5) perceived ease of use. Based on the analysis, the findings indicate that the implementation of i-KAM is not only applicable in the real software organizations but also help in delivering required projects within the prescribed cost and time.



**الخلاصة**

تهدف هذه الورقة إلى تقييم تطبيق النموذج المحسن لأجايل كانبان تجريبيا ، والذي يسمى i-KAM. النموذج i-KAM تم اقتراحه لتحسين مهمة مراقبة تطوير مشاريع البرمجيات باستخدام طريقة أجايل كانبان ، حيث أن هذه الطريقة لها تحديات حرجة تؤثر سلبًا على عملية إدارة مشاريع البرمجيات (SPM). ولتحقيق هذا الهدف ، أجرى الباحثون دراسة حالة في مؤسسة برمجيات من أجل تطبيق i-KAM في مشروع فعلي ضمن بيئة حقيقية. في هذه الدراسة ، استخدام أعضاء فريق عمل المؤسسة أداة النموذج الأولي لرصد التقدم (PM-PT) – التي تم تصميمها اعتمادا على i-KAM - أثناء عملهم اليومي في تطوير نظام جديد ولمدة ثلاثة أشهر. في نهاية دراسة الحالة، تم جمع البيانات عبر نموذج تقييم مصمم لقياس مدى قابلية تطبيق i-KAM استنادًا إلى خمسة عوامل ، وهي (1) اكتساب الرضا ، (2) رضا الواجهات ، (3) الرضا عن دعم المهام ، (4) الفائدة المتصورة ، و (5) السهولة في الاستخدام. تشير نتائج التحليل إلى أن i-KAM يمكن تطبيقه فعليا في مؤسسات البرمجيات ، بل إنه يساعد أيضًا على تسليم المشاريع المطلوبة ضمن التكلفة والوقت المحددين.

## 1. INTRODUCTION

Project Management (PM) traditionally defined as the process of planning, organizing, motivating and controlling resources, actions and rules in order to complete successfully specific goals and objectives within a specified period of time [1, 2]. Typically, this process is used in whole project as a series of activities in order to produce an exclusive deliverable. Meanwhile, it uses deadlines, starting and ending points to reach the target. In this context, the Project Management Institute [3] defines the project management as “the application of knowledge, skills, tools, and techniques to project activities to meet the project requirements”. Hence, Rose [4] claimed that meeting the requirements, on deadlines and within the budget, of project needs an effective management.

Agile Project Management (APM) is an innovative modern approach for managing software development projects [5]. APM allows the organizations to adjust their plans in line with changes in the project

environment, and delivers a number of novelties and benefits for the projects' team and client as well [6]. Kanban is an APM method which is being widely adopted in several fields and numerous settings for the purpose of managing the development of projects [7, 8]. Although this method was initially used in manufacturing domain; however, its adoption in other domains is continuously growing due to its flexibility, reliability, and proven successfulness [9, 10]. On top of that, Kanban method has an effective mechanism for managing projects and visualizing their workflow [11, 12]. Accordingly, the investigation of Kanban method is being a worthy study area, and recently attracted researchers from different perspectives and fields [7].

Among these researches, [11] conducted a study explored challenges of Kanban method in the software development. The results revealed that Kanban has major problem related to the task of monitoring. Thus, the study was concluded by developing a new model based on Agile Kanban in order to improve the software project monitoring task, which so called i-KAM. Though the effectiveness of i-KAM was verified by 11 experts [9] and also re-verified by seven software practitioners [5] [13] ; however, there is a high need to evaluate its applicability of in actual project. Therefore, this study aims to empirically evaluate the i-KAM applicability and ultimately answer the research question: Is i-KAM applicable in the real world environments?

The rest of this paper is structured as follows. Section 2 reviews the related work to this study, while section 3 describes the research method employed to achieve the study objective. Then, the evaluation results are demonstrated in section 4, followed by a discussion of the study findings in section 5. Final section summarizes the current study and provides some recommendations for forthcoming works.

## 2. RELATED WORK

A study conducted by Panigrahi and Behera [14] confirm that Kanban method has a very good system of production control and industries improvement. In addition, a system for production departments was developed based on Kanban to demonstrate the efficiency of the pull system. The developed system has been employed as a signal in order to automatically renew the parts existed in the warehouse. After implementing this system, the communication flow has been improved and the overproduction has been minimized, whilst the efficiency of the production has been increased notably [15]. Indeed, this success was not only in the manufacturing field, but also in other fields such as sports science, learning, Internet of Things (IoT), student projects, and software development.

Based on Common KADS and Kanban method, a Sports Science Knowledge Management (SSKM) system has been developed by Santirojanakul [16]. SSKM was aimed at improving the performance of the sports scientist's reporting system. The author affirmed that Kanban board supports the collaboration and communication between the sports scientists, executive, staff, and sport association. It also displays different types of sport competitions, sports associations, sports scientists, and athletic evaluations.

Moreover, Fitriawati and Lestari [10] designed an information system based on Kanban method for kindergarten learning evaluation. The Kanban method was used as management as well as control guidelines for all evaluation stages. Results indicated that using Kanban method helped the programmers to easily design the required system according to the system users' needs. This study concluded that Kanban method permits end-users to share the activities' flow and to set the activities' level in line with to their demands.

A study conducted by [8] proposed a new system for managing the IoT sensors via using Kanban method. The proposed system was developed to decrease the consumption power of the sensors. Therefore, this study pointed out that using Kanban method was an efficient management approach that increased the sensor lifetime and decreased the communication traffic. Furthermore, Saltz and Heckman [17] conducted a study compared various methods on how to guide students through computing projects. Their study was aimed to understand if one method is better than the other methods for student teams. The findings confirmed that Kanban is the effective method for guiding and managing student projects, as it improved student outcomes and minimized instructors' workload.

On the other hand, there are several studies conducted to investigate the Kanban method use in the software development [18, 19]. Particularly, in the Software Development Management (SPM), Kanban method has received a significant reputation specially for managing the process of software development projects [20-22]. Kanban method has a board used to visualize the workflow and control the progress of projects [23, 24]. Besides, [25] affirmed that software practitioners have shifted towards using Kanban method due to its effective communication, transparency, and limited by limiting the Work In Progress (WIP) traits.

However, Kanban method has three key challenges, which are: (1) it needs an effective tracking mechanism, (2) it has a difficulty in controlling WIP limits, and (3) it lacks displaying valuable information about the project progress. In consequence, these challenges have a significant impact on the fail of software delivery on the predefined time and cost [24, 26, 27]. To overcome the crucial challenges of Kanban, [11] carried out a comprehensive review to develop an initial model of i-KAM. As a result, sets of criteria were identified and then aligned with three main components proposed to constitute i-KAM.

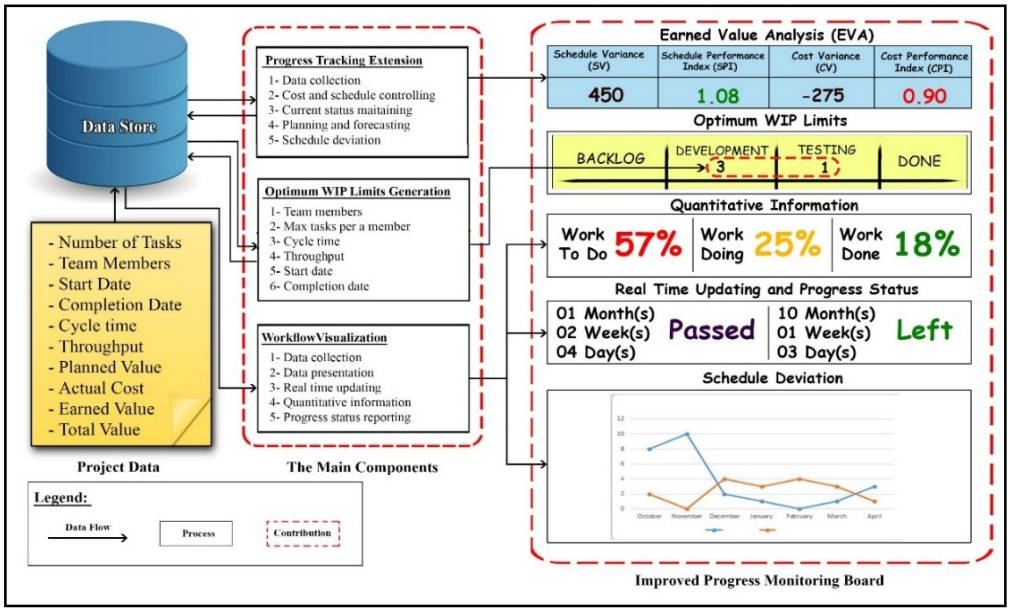

Figure 1. The proposed model (i-KAM)

As depicted in Figure 1, the proposed model consists of three main components, which are (1) extending progress tracking, (2) generating optimum WIP limits, and (3) visualizing useful insights for workflow. Each component has a major influence to solve the current challenges of the monitoring task as confirmed by [20]. In i-KAM, the first component uses the Earned Value Analysis (EVA) method to extend the progress tracking mechanism of Kanban method. The second component finds out the optimum WIP limits for all stages of Kabana board. However, the third component depicts significant information concerning with the project workflow. [28] developed a Progress Monitoring Prototype Tool (PM-PT) grounded on i-KAM to prove its concepts. PM-PT consists of several pages, and each page has different functions represent all tasks and process of i-KAM. For instance, Figure 2 shows the main page of PM-PT that appears after successful login displaying an overview about PM-PT. On the left side, a main menu contains a list of tabs for performing numerous functions executed the functions proposed in i-KAM.

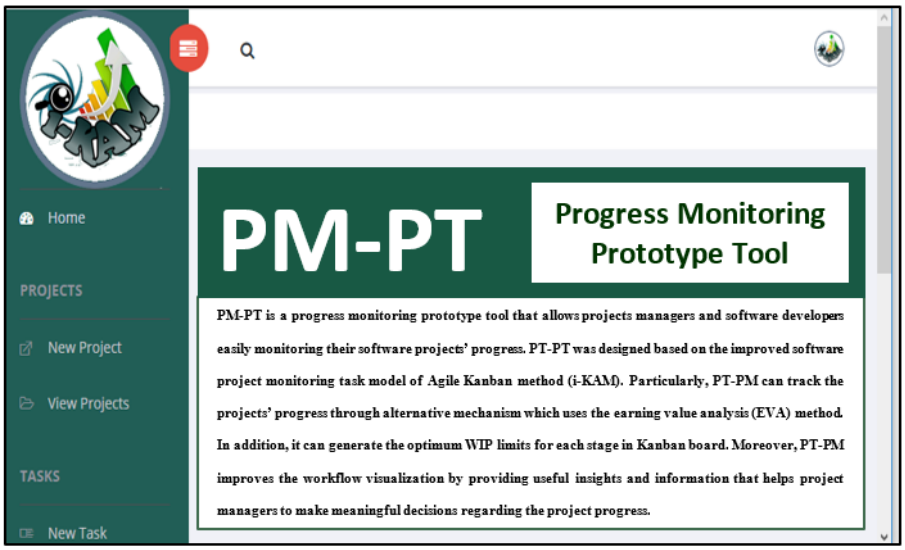

Figure 2. PM-PT main page

Besides that, another key page created to represent the improved progress monitoring board as shown in Figure 3. Mainly, this page contains variety functions and activities that were added to the original Kanban board in order to improve it and address previously recognised flaws and concerns. Notably, the improved board has three stages, which are To Do, Doing, and Done, whereby all projects' tasks are moved through the three stages starting from To Do stage up to Done stage. In addition, the tasks description together with the team member assigned to perform a potential task are visualized. Meanwhile, the improved board visualizes quantitative information regarding the current project status. The displayed insights summarize how much of

work has been done, is doing, and remained. It also visualizes the period passed and period left of project in months, weeks, and days.

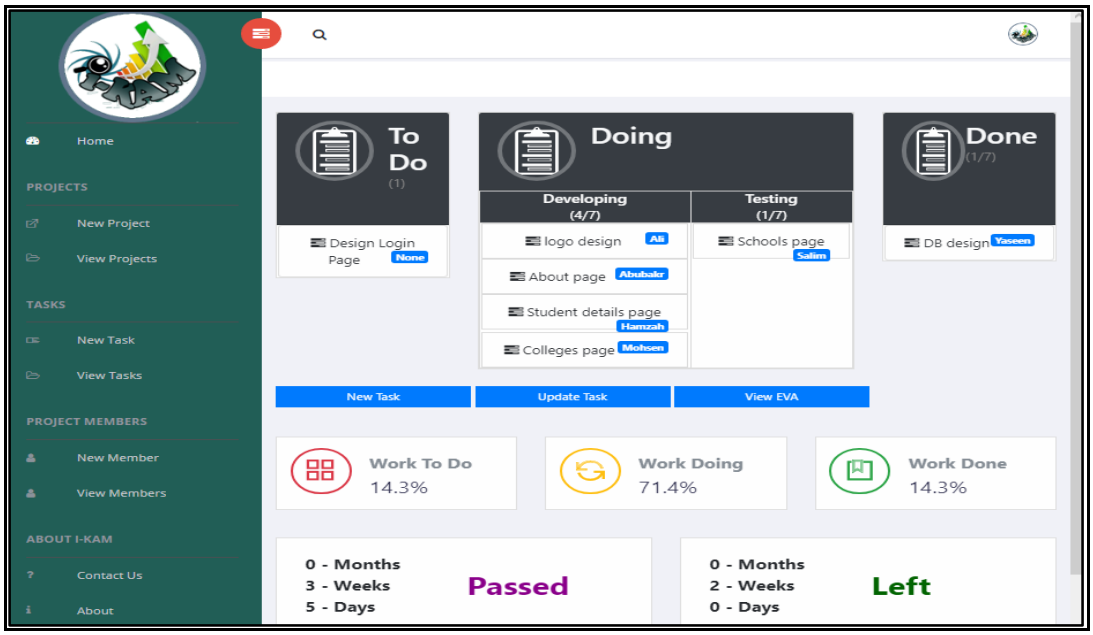


Figure 3. The improved progress monitoring board

In addition, the design of PM-PT has been initially validated through conducting interviews with seven practitioners work in Malaysia software companies [28]. The preliminary outcomes of the study indicated that the PM-PT has got their approval since the tool is an effective solution for tracking software projects. However, some participants suggested implementing PM-PT in actual projects in order to evaluate its applicability in software houses. Therefore, the focus of this study is to evaluate the practical applicability of PM-PT in a real project through conducting a case study.

## 3. Research Method

This study employed the case study method to empirically evaluate the applicability of i-KAM by software practitioners through using PM-PT in a real project within their organization. The case study method was selected as it is common approach used for evaluating the applicability [29] of the new approaches proposed in Software Engineering (SE) domain [30]. To achieve the study objective, three main procedures, which are (1) planning the case study, (2) conducting the case study, and (3) data collection and analysis, were followed. These procedures were adapted from Altarawneh [31] and Runeson and Höst [32], and discussed in the following subsections.

### 3.1 Planning the case study

The first procedure was carried out to plan the case study, in which it started by identifying an organization to implement PM-PT. Accordingly, to identify the potential organization, certain characteristics were defined, in which the identified organization should: (1) have a small size team, (2) develop software projects, and (3) have willingness to cooperate in evaluating i-KAM by implementing PM-PT in one of its projects. In this procedure, a focus was given to select an organization based on how it manages the development process of software projects, regardless of the used method in that organization. Consequently, the Universiti Utara Malaysia Information Technology (UUMIT) was selected as it is an educational organization meets the predefined requirements for conducting the case study. Moreover, one of the key reasons for selecting UUMIT is to response to the suggestion provided by a software practitioner during interviewed him when initially developed PM-PT. He suggested evaluating the applicability of i-KAM in educational setting having a small team for developing simple projects [28]. Furthermore, the universities recently became realised the significance of the Kanban method and their researches focus on its trends and applications in the industry [33].

Initially, a formal letter attached with an e-mail was sent to the UUMIT director requests giving the researchers a permission to conduct a case study by applying the designed tool (PM-PT) in any forthcoming system will be developed by UUMIT. Within the application, the purpose of conducting a case study was explained which is to evaluate the i-KAM applicability within a small sized project in a real environment. In response, the UUMIT director agreed on above mentioned request and sent an acceptance letter to conduct a

case study in UUMIT. The director suggested working on the latest project which is a "System for Applying Business License by Student". This system aims to simplify and speed up the application process for business license of students who have business in UUM campus. This appointed system was requested by the Co-operative and Entrepreneurship Development Institute (CEDI), which is an UUM centre, aims to nurture and develop student entrepreneurs through student business entrepreneurs.

After receiving the acceptance, a first meeting was held in UUMIT with the head of University Information Systems (UIS) and the project manager. The UIS is a UUMIT unit which is liable to develop all systems requested by any UUM department. Hence, the first meeting was aimed to introduce the idea of the proposed model and to present the purpose of conducting the case study. It was also aimed to get an overview on the forthcoming project and understand its requirements. Figure 4 displays a side of the discussion during meeting the UIS administration in UUMIT.

Figure 4. Meeting with UIS administration in UUMIT

In the presentation, the researchers introduced the importance and contributions of i-KAM, and briefly illustrated about the tool (PM-PT). Besides that, the case study procedures were explained, and the factors that would be ultimately used to evaluate PM-PT were presented. In addition, the initial steps to start conducting the case study were highlighted along with identifying the information required before developing the project.

### 3.2 Conducting the case study

The second procedure was carried out to conduct the case study, which has been started by holding the second meeting with the team members of UIS unit who have appointed to develop the nominated project. The purpose of the second meeting was to explain to the team members how to use PM-PT practically. Thus, the researchers illustrated the tasks, which are performed by project manager, and tasks, which are executed by team members during developing the software projects. The explanation was supported by giving examples for creating and updating projects and tasks, along with adding team members to the PM-PT. Moreover, the main functions of PM-PT were also highlighted, which include tracking the progress and utilizing the useful information visualized in an improved progress monitoring board for obtaining an effective monitoring of the project progress. Meanwhile, they were reminded about the factors and items that would be used ultimately to evaluate PM-PT.

After finishing the explanation, the researchers shared with the team members the URL of PM-PT because it is a web based tool, and then provided them by login information to start using PM-PT. The team members have regularly used PM-PT to report their progress when releasing a task or changing its status from stage to another. Furthermore, during conducting the case study, any problems faced the team members were discussed and solved by holding instant meetings or via WhatsApp group.

### 3.3 Data collection and analysis

The last procedure in implementing the case study was to collect and analyse data regarding the evaluation of PM-PT. As whole, the team members of UUMIT have used PM-PT during a period of three months, whilst the data collection took around two hours as all team members work in one place. They were met to provide their opinion on PM-PT, which has been implemented in developing the project. In this procedure, an evaluation form was used for data collection to evaluate the applicability of i-KAM after using PM-PT.

The evaluation process was conducted based on five factors, which are (1) gain satisfaction, (2) interface satisfaction, (3) task support satisfaction, (4) perceived usefulness, and (5) perceived ease of use as claimed by [30]. After that, researchers started analysing the data collected by the evaluation form that provided by participants. In this procedure, the content analysis approach was employed because it is the common method used in quality researches. Therefore, the content analysis was used to describe the feedback of team members by identifying, coding, and categorizing the main themes in the collected data [34].

## 4. EVALUATION RESULTS

This section demonstrates the evaluation results obtained from the UIS team members. It starts by providing the team demographic information, followed by discussing each factor in subsection. As required, this project has been assigned to small team involves two members, wherein the development team is typically responsible for developing any projects related to student affairs. The team member's age is 42 and 35 old, and they have 17 and 8 years of experiences in developing software projects respectively. Mainly, the development team with small members enhances interpersonal communication skills [35]. Besides that, Malik [36] confirms that a small team can be highly efficient. Thus, the number of the team members who had participated in the study is sufficient to achieve valid and reliable results.

### 4.1 Gain satisfaction of PM-PT

The gain satisfaction was assessed based on four variables, which are decision support satisfaction, comparing with current method, clarity, and task appropriateness. The participants were satisfied with the PM-PT, whereby they indicated that tool can help the management to take a well-defined decision regarding to software project monitoring task. In addition, all functions provided by PM-PT were clear to the team members, wherein each phase obviously presents the required inputs, processes, and outputs. Although team members have never use any monitoring tool before, they affirm that having PM-PT would effectively track their progress and thus deliver software products within the planned cost and time. In general, the participants agreed that the PM-PT is an appropriate tool for monitoring the development process of software projects.

### 4.2 Interface satisfaction of PM-PT

Among the variables that were assessed for interface satisfaction are internally consistent, organization (well organized), appropriate for audience, and presentation. The participants confirmed that PM-PT is internally consistent due to the good design of interfaces that makes performing the PM-PT functions in accurate manner. Moreover, PM-PT is found to be well organized and structured, thus it makes the project tasks easy to be performed. Particularly, it enables project manager creating new projects by keying in main information, such as project name, customer name, e-mail, etc. Then, project manager can assign tasks to a particular project by keying in task description, along with its required cycle time (in days).

Furthermore, the participants pointed out that PM-PT is appropriate for the project managers and development team. They also affirmed that results presented by PM-PT are in a readable and useful format, as showed in an improved progress monitoring board. This board visualizes sufficient information and critical insights for the project workflow. For example, it reports quantitative information on the current project status - by percentage - that helps project managers to make significant judgements concern with the projects' workflow.

### 4.3 Task support satisfaction of PM-PT

This subsection demonstrates the results provided the participants on the task support satisfaction factor of PM-PT. This factor was measured based on three variables, which are ability to produce expected results, completeness, and ease of implementation. The participants indicated that PM-PT is able to generate expected and valued results. The information visualized in the improved progress monitoring board was very beneficial to the team members. For instance, displaying the project tasks, project duration (passed and left), project status, is a valuable approach to encourage team members performing well. To some extent, the participants were agreed that PM-PT is adequate and sufficient for monitoring the progress task during software development projects. PM-PT can assist to track their tasks and control their schedule during the development of the required system. In addition, participants also asserted that PM-PT is easy to be implemented within actual projects in real world environments.

### 4.4 Perceived usefulness of PM-PT

The perceived usefulness factor of PM-PT was assessed based on five variables, which are: accomplishing more work, work performance, make tasks easier, usefulness, and increasing productivity. Generally, the participants indicated that PM-PT is useful for their working environment, as it can enable them to accomplish their tasks more quickly. They also affirmed that using PM-PT improved their performance and

made them performing their tasks easier. Meanwhile, the participants highlighted that using PM-PT can partly assist in increasing their productivity.

### 4.5 Perceived ease of use of PM-PT

This subsection explains the results of the perceived ease of use factor. The assessment of PM-PT was conducted according to six variables, which are ease of learning, confusing, flexible, understandable, effort to become skilful, and ease to use. According to the participants' responses, the PM-PT was perceived as easy to be used because it has well-defined functions and features. Thus, they indicated that learning to operate PM-PT is easy for them and for other software practitioners as well. Moreover, the participants found that they can easily make PM-PT does exactly what they want. Importantly, PM-PT has deemed as flexible tool as the team members confirmed that their interaction with PM-PT was clear and understandable too. Furthermore, the participants indicated that it is easy to become skilful when using PM-PT.

## 5. DISCUSSION

The current case study has really used PM-PT to evaluate the applicability of i-KAM in developing an actual project. In this context, the required system has been developed by UIS team with employing the PM-PT in monitoring the development process. The findings from the conducted case study confirm that the proposed model (i-KAM) is applicable to be implemented in real projects which are developed by software houses. In addition, the findings indicate that the PM-PT has gained participants' satisfaction, whereby they affirmed that i-KAM deemed as practical model can be implemented in the actual projects. Primarily, the team members emphasize that PM-PT is useful for projects managers to monitor the projects development life cycle.

Moreover, they affirm that tool is beneficial in terms of organizing the project tasks. Thus, developers can know what tasks which need to work on. Particularly, in PM-PT, tasks are assigned to team members clearly, in which they can update their progress easily. Therefore, project manager can straightforwardly monitor all projects tasks. Furthermore, the team members acknowledge that PM-PT enables monitoring numerous tasks because of the work list is more visible and organized too.

In the evaluation form, the participants were asked to provide comments and suggestions towards enhancing PM-PT. In response, they suggested assigning starting/ completion dates for each task of project, thus would help project manager to easily monitor the development team when starting or finishing a task. However, PM-PT provides the similar function which is assigning a cycle time, in days, for each task in a project. Besides that, the participants also suggested to notify the project manager (e.g. by e-mail) when a task has been done by a team member. In addition, the participants recommended adding a button for editing the progress of a developer, in which he/she can search back in the assigned list of project tasks. Nevertheless, this property is already existed in PM-PT wherein team members can update their progress. Moreover, one of the participants argued that PM-PT needs a time to record all project tasks, thus would add more workload to the team members as they are required to update their progress continuously. Conversely, the use of SPM tools has numerous benefits to the team members. For instance, it enable them planning their tasks and collaborating together, as well as monitoring their progress [37]. In addition, using SPM tools in practices is being on the rise within SDOs, because it is one of key factors of successfully managing the development of software projects [26].

Despite of the above mentioned positive findings, one of the main functions of PM-PT was not implemented due to its unsuitable use in the current case. In PM-PT, EVA method is used to track project progress effectively. It mainly requires three inputs values, which are Planned Value (PV), Actual Cost (AC), and Earned Value (EV). In this case, PV is a recognized value; however, the values of AC and EV could not be determined. This is because UUMIT is non-profit in-house organization established to develop systems needed by the UUM departments; accordingly, the customers here are typically under UUM environment. Furthermore, UUMIT development team members only receive their monthly salary, rather than earning any stipend against their efforts in developing systems. In this case, applying EVA method for tracking the progress has not been performed. Therefore, another case study would be conducted in a profit software development house.

## 6. CONCLUSION AND FUTURE WORK

This paper demonstrated the results of evaluating the applicability of i-KAM. A case study has been conducted in UUMIT, in which PM-PT was used by team members during developing a real project. The case study conducted was ideally suited to achieve valid and reliable results due to the small size team as argued by Malik [36]. The results showed that i-KAM is an applicable model as well as it is easy to be implemented in actual environments. Besides that, the results affirmed that the application of i-KAM will practically contribute to different stakeholders, such as software engineers, software practitioners, and software project managers. By implementing the proposed model, an improved Agile Kanban method would be introduced to SE domain, and thus would be useful for software engineers to be adopted or adapted in different countries or studies areas.

In addition, software practitioners can effectively monitor the development process of software projects' progress. Besides that, useful insights and information are visualized to assist project manager in making expressive judgements with regards to project progress.

In this vein, software houses can deliver their software projects successfully according to its specifications, within the prescribed period and budget. Therefore, the rate of successful software projects would be increased, and will thus help the developers, software houses, and nation to elevate a more advanced economy. However, this study was limited to obtain results from single case study in non-profit in-house organization. Thus, to generalize the results, future work will focus on conducting more case studies in numerous settings, such as profit houses and/or out-house in other countries as well. Within this in mind, different methods, such as experts' review, and controlled experiments, could be used to evaluate i-KAM applicability. Moreover, it is recommended to investigate alternative dimensions and new directions towards enhancing i-KAM effectiveness.

## ACKNOWLEDGEMENTS

The authors would like to thank the participants in the case study and special thanks go to the Al-Ahgaff University, Yemen and Universiti Utara Malaysia (UUM), Malaysia.

## REFERENCES

[1] K. Schwalbe, *Information technology project management*, 8th ed. United States: John Wiley & Sons, Inc., 2016.
[2] J. R. Turner and R. Müller, "On the nature of the project as a temporary organization," *International Journal of Project Management,* vol. 21, pp. 1-8, 2003.
[3] Project Management Institute, *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)*, Kindle Edition, 6th ed. USA: Project Management Institute, Inc, 2017.
[4] K. H. Rose, "A Guide to the Project Management Body of Knowledge (PMBOK® Guide)—Fifth Edition," *Project Management Journal,* vol. 44, pp. e1-e1, 2013.
[5] H. Alaidaros, M. Omar, R. Romli, and A. Al-Sakkaf, "The Evaluation of an Improved Model of the Agile Kanban Using Focus Group," *Journal of Information Technology Management,* vol. 14, pp. 1-22, 2022.
[6] M. Masood, F. Azam, M. W. Anwar, and A. Amjad, "Defining Meta-Model for Value-Oriented Requirement Prioritization Technique," in *Proceedings of the 2019 7th International Conference on Computer and Communications Management*, 2019, pp. 72-77.
[7] H. Alaidaros, M. Omar, and R. Romli, "The State of the Art of Agile Kanban Method: Challenges and Opportunities," *Independent Journal of Management & Production (IJM&P),* vol. 12, pp. 2535-2550, 2021.
[8] T. Kurita, K. Matsuo, and L. Barolli, "A Wheelchair Management System Using IoT Sensors and Agile-Kanban," presented at the The 11th International Conference on Intelligent Networking and Collaborative Systems (INCoS-2019), Oita University, Oita, Japan, 2020.
[9] H. Alaidaros, M. Omar, and R. Romli, "An improved model of Agile Kanban method: Verification process through experts' review," *Int. J. Agile Systems and Management (IJASM),* vol. 13, pp. 390–416, 2020.
[10] M. Fitriawati and R. Lestari, "Design of the Information System for Kindergarten Learning Evaluation used Kanban Methodology," in *IOP Conference Series: Materials Science and Engineering*, 2019, pp. 1-6.
[11] H. Alaidaros, M. Omar, and R. Romli, "Towards an improved software project monitoring task model of Agile Kanban method," *International Journal of Supply Chain Management (IJSCM),* vol. 7, pp. 118-125, 2018.
[12] H. Lei, F. Ganjeizadeh, P. K. Jayachandran, and P. Ozcan, "A statistical analysis of the effects of Scrum and Kanban on software development projects," *Robotics and Computer-Integrated Manufacturing,* vol. 43, pp. 59-67, 2017.
[13] H. Alaidaros, M. Omar, R. Romli, and A. Hussein, "Re-verification of the Improved Software Project Monitoring Task Model of the Agile Kanban (i-KAM)," in *Lecture Notes on Data Engineering and Communications Technologies*. vol. 72, F. Saeed, F. Mohammed, and A. Al-Nahari, Eds., ed Cham: Spriner, 2021, pp. 1247-1258.
[14] B. Panigrahi and D. K. Behera, "Improvement in productivity using Kanban process based cognitive automation system," *International Journal of Mechanical Engineering and Technology (IJMET),* vol. 9, pp. 122–130, 2018.
[15] M. Jamaliah, A. Mohamad, and I. Azianti, "Improvements of worksite control for pull system," *ARPN Journal of Engineering and Applied Sciences,* vol. 11, pp. 7699-7705, 2016.
[16] S. Santirojanakul, "The development of sports science knowledge management systems through CommonKADS and digital Kanban board," in *2018 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, Penang, Malaysia, 2018, pp. 119-124.
[17] J. S. Saltz and R. R. Heckman, "A scalable methodology to guide student teams executing computing projects," *ACM Transactions on Computing Education (TOCE),* vol. 18, pp. 1-19, 2018.
[18] M. K. Yacoub, M. A. A. Mostafa, and A. B. Farid, "A New Approach for Distributed Software Engineering Teams Based on Kanban Method for Reducing Dependency," *JSW,* vol. 11, pp. 1231-1241, 2016.
[19] M. Maneva, N. Koceska, and S. Koceski, "Introduction of Kanban methodology and its usage in software development," presented at the International Conference on Information Technology and Development of Education Zrenjanin, Republic of Serbia 2016.
[20] H. Alaidaros, M. Omar, and R. Romli, "Improving the progress monitoring task of Agile Kanban method: An enhanced theoretical framework and its implication," *International Journal of Scientific and Technology Research (IJSTR),* vol. 9, pp. 611-615, 2020.
[21] I. Shamshurin and J. S. Saltz, "Using a coach to improve team performance when the team uses a Kanban process methodology," *International Journal of Information Systems And Project Management (IJISPM),* vol. 7, pp. 61-77, 2019.
[22] M. O. Ahmad, D. Dennehy, K. Conboy, and M. Oivo, "Kanban in software engineering: A systematic mapping study," *Journal of Systems and Software,* vol. 137, pp. 96-113, 2018.
[23] D. Dennehy and K. Conboy, "Going with the flow: An activity theory analysis of flow techniques in software development," *Journal of Systems and Software,* vol. 133, pp. 160-173, 2017.

[24] H. Alaidaros and M. Omar, "Software project management approaches for monitoring work-in-progress: A review," *Journal of Engineering and Applied Sciences,* vol. 12, pp. 3851-3857, 2017.
[25] S. Shafiq, I. Inayat, and M. Abbas, "2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)," presented at the Euromicro Conference on Software Engineering and Advanced Applications, Thessaloniki, Greece, 2019.
[26] M. O. Ahmad, J. Markkula, and M. Oivo, "Insights into the perceived benefits of Kanban in software companies: Practitioners' views," presented at the International Conference on Agile Software Development, Cham, 2016.
[27] J. F. Tripp, J. Saltz, and D. Turk, "Thoughts on current and future research on Agile and Lean: Ensuring relevance and rigor," presented at the Proceedings of the 51st Hawaii International Conference on System Sciences, USA, 2018.
[28] H. Alaidaros, M. Omar, R. Romli, and A. Hussein, "The Development and Evaluation of A Progress Monitoring Prototype Tool for Software Project Management," presented at the the 2019 First International Conference of Intelligent Computing and Engineering (ICOICE), Hadramout, Yemen, 2020.
[29] K. Eldrandaly, "Developing a GIS-based MCE site selection tool in ArcGIS using COM technology," *International Arab Journal of Information Technology (IAJIT),* vol. 10, pp. 276-282, 2013.
[30] H. Alaidaros, M. Omar, and R. Romli, "The key factors of evaluating Agile approaches: A systematic literature review," *International Journal of Supply Chain Management (IJSCM),* vol. 8, pp. 954-964, 2019.
[31] M. H. A. Altarawneh, "Monitoring oriented Agile based web applications development methodology for small software firms in Jordan," PhD Thesis, Universiti Utara Malaysia, Malaysia, 2016.
[32] P. Runeson and M. Höst, "Guidelines for conducting and reporting case study research in software engineering," *Empirical software engineering,* vol. 14, p. 131, 2009.
[33] V. Mahnič, "Kanban in software engineering education: an outline of the literature," *World Transactions on Engineering and Technology Education,* vol. 17, pp. 23-28, 2019.
[34] K. Krippendorff, *Content analysis: An introduction to its methodology*. United States: SAGE Publications, Inc, 2018.
[35] R. Hannemann, "Modifying senior design: A design review," in *American Society for Engineering Education Annual Conference and Exposition 2009*, Austin, Texas, USA, 2009, pp. 10560-10568.
[36] F. Malik, "Syntegrating city management," in *Institutional and Social Innovation for Sustainable Urban Development*. vol. 258, ed: ROUTLEDGE in association with GSE Research, 2013, pp. 258-274.
[37] M. Choetkiertikul, H. K. Dam, T. Tran, and A. Ghose, "Predicting the delay of issues with due dates in software projects," *Empirical Software Engineering,* vol. 22, pp. 1223-1263, 2017.

Al-Ahgaff University Journal of Computer Science and Mathematics
مجلة جامعة الأحقاف لعلوم الحاسوب والرياضيات

Vol. 1, Issue. 1, Jan 2023, pp. 17~24

# Efficient Information Retrieval for bug localization using fewer objective functions

**Ahmed Sheikh Al-Aidaroos**
Information Technology Department, Al-Ahgaff University, Mukalla, Yemen
asalaidaroos@ahgaff.edu
**Sameera Bin Ali Al haj**
Information Technology Department, Al-Ahgaff University, Mukalla, Yemen
Seba.alhaj@gmail.com

**Article Info**

***Article history:***

Received jan 11, 2023
Revised jan 19, 2023
Accepted jan 25, 2023

***Keywords:***

Bug Report
Bug Localization
Information Retrieval based Bug Localization
Multi objective optimization
Automatic Query Reduction

**ABSTRACT**

Bug localization is one of the hardest, most costly, and most time-consuming tasks that faces software developers. Therefore, many Information Retrieval based Bug Localization (IRBL) approaches have been proposed to reduce the time and effort spent in localizing bugs. However, the quality of the queries positively affects the performance of the bug localization. ManQ is a multi-objective optimization based IRBL approach that seeks to improve query quality. It studied a set of IRBL approaches and converted them to a group of 15 objective functions. However, some researchers claimed there are a set of attributes that make a query high quality much less than ManQ's objective functions. Therefore, this study aims to adapt the ManQ approach by reducing the objective functions and keeping only the objective functions that correspond to these attributes. The adapted approach is named R-ManQ. The results show that both R-ManQ and ManQ have similar performances, but R-ManQ is much faster in terms of execution time and has a smaller number of query terms.



**الخلاصة**

تعد توطين الأخطاء من أصعب المهام وأكثرها تكلفة واستهلاكًا للوقت التي تواجه مطوري البرامج. لذلك، تم اقتراح العديد من مناهج توطين الأخطاء المستندة إلى استرداد المعلومات (IRBL) لتقليل الوقت والجهد المبذولين في توطين الأخطاء. ومع ذلك، تؤثر جودة الاستعلامات بشكل إيجابي على أداء توطين الخطأ. ManQ هو منهج IRBL متعدد الأهداف قائم على التحسين والذي سعى إلى تحسين جودة الاستعلام. فقد درس مجموعة من مناهج IRBL وحولتها إلى مجموعة من 15 وظيفة موضوعية. ومع ذلك، ادعى بعض الباحثين أن هناك مجموعة من السمات التي تجعل الاستعلام ذا جودة عالية أقل بكثير من وظائف ManQ الموضوعية. لذلك، تهدف هذه الدراسة إلى تكييف نهج ManQ عن طريق تقليل الوظائف الموضوعية والحفاظ فقط على الوظائف الموضوعية التي تتوافق مع هذه السمات. النهج المعدل سمي R-ManQ. تظهر النتائج أن كلا من R-ManQ وManQ متشابهان في الأداء، لكن R-ManQ أقل بكثير من حيث وقت التنفيذ وعدد مصطلحات الاستعلام.

## 1. INTRODUCTION)

All software goes through the testing and maintenance phases within the development life cycle in order to ensure that it is free from problems and errors. Despite this, many errors and issues appear, which are reported to the development teams during these phases. Those bugs are defined as an existing unexpected behavior or unexpected performance of a predefined functionality in the source code of the software [1, 2]. Those bugs are written by the end users in a text file in a natural language and then delivered to the development team in order to use it to understand the nature of those bugs and the reasons for their occurrence. This file is called a bug report [1, 3-5]. Therefore, a bug report is defined as a document created by end-users that describes this unexpected behavior and errors in the performance of the program when they use it and what the steps that lead to that behavior. A bug report contains several sections, including the bug ID, summary, and description [6].

According to [1, 6], these bugs go through various phases: unconfirmed, new, assigned, resolved, verified, closed, or reopened, as shown in Figure 1. But locating these bugs during the development phase is a difficult, costly, and time-consuming process [2, 3, 5, 7].

In addition, according to [8], it is more costly and complex to deal with bugs after the delivery of software. The process of searching and exploring source code for identifying the locations of buggy files by the developers depends on those reports, known as bug localization [5]. The localization of these bugs can be considered a somewhat simpler task for developers involved in the program development process, but a daunting task for other developers who are not involved in the development process of this software [5, 9]. In addition, manual localization of these bugs is considered difficult and expensive, especially when the software is large and complex [7, 10].

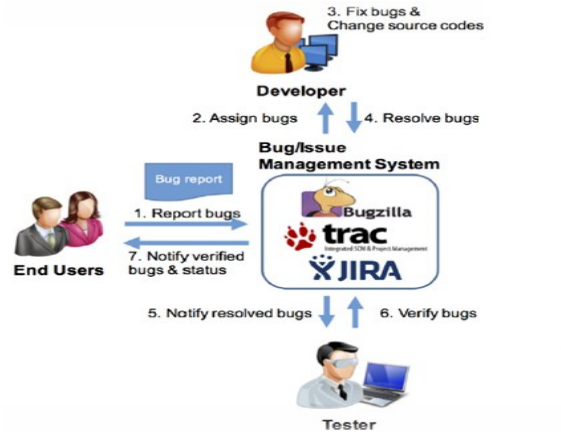


Figure. 1 Management Process of Bug [1]

In general, there are two approaches to localizing bugs: 1) dynamically locating bugs spectrum-based through program execution combined with techniques such as implementation, data monitoring, and breakpoints, and 2) statically locating bugs across various forms of analyses using bug reports along with the code, such a bug localization based on information retrieval. The first type is time-consuming and expensive, while the second type is preferred [11, 12].

Therefore, to reduce the time and effort spent in localizing these bugs, a number of IRBL approaches are proposed. The IRBL approach mechanism can be described by using a bug report as a query and source code elements as a document collection in order to find similarities between them to localize the bugs. The documents are ranked according to their relevance to the bug report, returning a ranked list of candidate source files, and then the developer checks the top-N of them, one at a time, and determines whether or not they contain the bug [7, 12-16].

According to [5], , the IRBL approaches must suggest a suitable search terms automatically from the bug report. The baseline approach applies the IR steps without any optimization. Also, it applies a three-step preprocessing step to a bug report and a buggy program (camel case splitting, stop word removal, and stemming) in order to generate the initial query and search the document corpus to retrieve relevant documents [12, 14]. However, the quality of the query is more important and it determines the performance of IRBL, whereas many IRBL approaches cannot perform well due to their use of low-quality queries [17-19].

The relevant document results in High-quality queries are at the top of the list of results, whereas low-quality queries either retrieve the requested document at the bottom of the list of results or return no result at all. Therefore, using low-quality queries makes bug localization a tedious, time-consuming, and poorly performing process [3, 11, 16, 20].

In addition to that, there are two strategies that produce a new, high-quality query based on a bug report. It involves an expansion strategy that expands the initial query with appropriate keywords and other reduction strategies that discard the noisy words from a bug report and focus on the important keywords only [20-22]. According to [11, 21] , the terms of a query selected from the bug report must be chosen carefully. Mills, et al. [16] show that bug reports alone contain enough keywords to form high-quality queries and, therefore, provide optimal performance for bug localization. Furthermore, even natural language-only bug reports might be a sufficient source of perfect terms for a query [11]. Rahman, et al. [11] define the optimal query as a query that can locate the buggy document at the top of its relevant results list.

Mills, et al. [16] proposed an approach called Query Quality Predictor (Q2P). It predicted the quality of bug report queries in concept location and traceability link recovery for Software Engineering (SE). Q2P employed a set of 28 measures of query properties (21 pre-retrieval properties and 7 post-retrieval properties) and then a machine-learning algorithm (Random Forest) in order to define a set of rules that will identify which of these queries are high-quality and which are low. Q2P was an improvement over a previous approach that relied on using only 21 pre-retrieval properties to predict the quality of the query [20]. By using Q2P, when the developer writes a low-quality query, the developer can directly reformulate it without spending time analyzing potentially useless documents retrieved by the text-relevant engine and reformatting that low-quality query.

Kim and Lee [15] proposed a new approach called ManQ based on a multi-objective genetic algorithm called Non-dominated Sorting Genetic Algorithm III (NSGA-III). ManQ focused on a set of IRBL approaches specialized in Automated Query Reduction (AQR). It investigated finding a high-quality sub-query. ManQ converted those approaches into a group of objectives and combined them into one approach. It extracted 15 objectives from them and created 15 objective functions that include eight objectives to maintain the return values of Query Quality Properties (QQP) assessment metrics, four objectives to maintain the important terms for IRB, two objective functions to maintain initial information, i.e., cosine similarity and term-occurrence in the context of the sentences, and one objective to reduce the query length. ManQ introduced a better performance in IRBL. ManQ went through three steps. 1. calculating objective functions; 2. implementing the NSGA-III algorithm and initializing the final query selection; and 3. implementing bug localization. The first step was calculating 15 objective functions that were divided into four groups, group one of which includes eight objective functions ($f_1(Q) - f_8(Q)$) to maintain the return values of Q2P assessment metrics. They were finding only the pre-retrieval property measures (coherency, similarity, term relatedness, and specificity). Group two included four objective functions ($f_9(Q)$- $f_{12}(Q)$) to maintain the important terms for IRBL. In addition, group three consisted two objective functions ($f_{\mathbf{13}}(Q)$and $f_{\mathbf{14}}(Q)$) to maintain initial information, i.e., cosine similarity and term-occurrence of a context of the sentence. The last group consisted of one objective function $f_{15}(Q)$ to reduce the query length. After completing this step, ManQ moved to the second step, which was implementing the NSGA-III algorithm, which resulted in multiple solutions from which ManQ chose a final query through the union of selective queries from each function. The last step in ManQ was implementing bug localization, where after obtaining the optimal query, ManQ submitted the query to a popular search engine for document search called Lucene in order to find the buggy files in the source code [23].

In addition, a recently published empirical study [11] reported that the bug reports that can contain terms but not software information are useful and sufficient for the bug localization process and give high results. They compared the graph-based IRBL approaches that were used to extract the important terms with a number of other IRBL approaches. They concluded that the graph-based approaches were the best approach for extracting the important terms. In addition, the authors apply a comparison between the optimal high-quality queries, which achieve high results in retrieving information, and those of lower quality, and they claimed there are a set of attributes that make a query high-quality. These attributes are: keywords of high quality queries are less frequent within a bug report, are less ambiguous (i.e., have less entropy), are more likely to be found in the description section of a bug report, are more likely to be nouns, and the optimal candidates are likely to be a small number of important keywords. That study used the Genatic Algorithm (GA), which is widely used to solve complex optimization problems in various research domains, including SE.

This paper depends on the ManQ approach [15] as a base work, and it seeks to evaluate the performance of ManQ while reducing the 15 objective functions of ManQ that meet Rahman, et al. [11] study.

## 2. METHOD

In this paper, ManQ approach is employed as a base work. This paper seeks to evaluate the performance of ManQ while reducing its 15 objective functions supported by Rahman, et al. [11] study and excluding other objective functions that were not supported.

### 2.1 Many-objective optimization-based automatic query reduction (ManQ):

ManQ studied a set of IRBL specialized AQR approaches in order to find a high-quality subquery. It combines 15 optimization objective functions into one approach to achieve all objectives without neglecting any of them, which could negatively affect the outcome. It applies preprocessing to the bug report (stop word removal and splitting of dotted terms and camel case terms); then a set of terms is produced from the initial query and encoded as binary genes [15]. As shown in Figure 2, ManQ goes through the following steps:

#### 2.1.1 First step: calculating the Objective functions

It consists of 15 objectives that aim to improve query quality by combining the individual objectives of some studies. The objective functions are divided into four groups, as shown in Figure 2. The first group is QQP, which is a set of pre-retrieval query properties measures that are used in IR to evaluate query performance and are computed before executing the query [16, 20]. These QQP are specificity, coherency, similarity, and term-relatedness. They form eight objective functions ($f_1(Q) - f_8(Q)$). The second group maintains important keywords. It is a set of measures that include 1. determining whether the terms of a query are in the sentence describing Observed and Expected Behaviour (OEB), 2. determining whether the query contains the four keywords (words in the first, second, penultimate, or last position) of the bug report summary and the name of the source file; and 3. determining whether terms in the query are grammatically significant based on part-of-speech priority (POS). This group formed four objective functions (f_9 (Q)- f_12 (Q)). The third group is maintaining initial information, which is a set of measures that preserve the context of the sentences in the bug report based on the occurrence of the term using the PMI scale and compute the cosine similarity between subquery Q and the original query. This group formed two objective functions (f_13 (Q)and f_14 (Q)). The fourth group is minimize the query length by compare the results length of a sub query with the initial query

to evaluate how short the sub query is. It forms one objective functionf_15 (Q). All objective functions are illustrated in Figure 2.

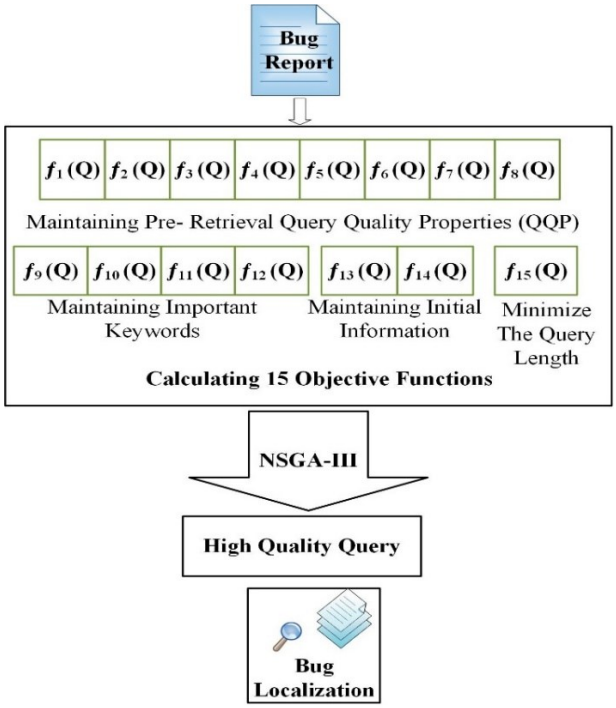


Figure. 2 An illustrated figure of ManQ Approach

**2.1.2 The second step: NSGA-III and final query selection**:

In this step, after NSGA III is implemented, multiple solutions are resulted where ManQ chooses a final query through the union of selective queries from each function according to next equation:

$$Q = \text{EuclideanDistance}(Q*, f) \cup \text{RankSum}(Q*, f), \quad (1)$$

$$\text{where} \quad f_i = \begin{cases} f_i = \frac{1}{|Q|} \cdot f_i & , \ if\ i<7 \\ f_i = f_i & , otherwise \end{cases} \quad (2)$$

$$\text{EuclideanDistance}(Q*, f) = \text{argmax}_Q \sqrt{\sum_{f_i \in f} f_i(Q)^2} \quad (3)$$

$$\text{RankSum}(Q*, f) = \text{argmin}_Q \sum_{f_i \in f} \text{rank}(f_i(Q)) \quad (4)$$

**2.1.3 The third step: Bug Localization:**

After obtaining the optimal query, the next step is to submit the query to a popular search engine for documents called Lucene in order to find the buggy files in the source code [5, 15, 16, 18].

**2.2 Proposed Work: R-ManQ**

As mentioned in the Rahman, et al. [11] study, it concluded there are several attributes that make high-quality queries, and the performance of IRBL approaches is highly efficient. They are: keywords of that query are less frequent within a bug report, are less ambiguous (i.e., have less entropy), are more likely to be found in the description section of a bug report, are more likely to be nouns, and the optimal candidates are likely to be a small number of important keywords. Therefore, R-ManQ aimed to reduce the objective functions that make up the ManQ by adopting the objective functions supported by Rahman's study [11] and dispensing with the rest of the other objectives. Therefore, these attributes were represented in seven objective functions and were kept within the ManQ approach while dispensing with the rest of the other objective functions, and they are:

$f_4(Q)$and $f_5(Q)$to fulfill the keywords of the optimal query are less frequent within a bug report.

$f_6(Q)$to fulfill the keywords of the optimal query are less ambiguous (i.e., have less entropy).

$f_9(Q)$ and $f_{10}(Q)$to fulfill the keywords of the optimal query are more likely to be found in the description section of a bug report.

$f_{12}(Q)$to fulfill the keywords of the optimal query are more likely to be nouns.

$f_{15}(Q)$to fulfill the optimal candidates are likely to be a small number of important keywords.

As shown in Figure 3, the blue rectangles are represented the remaining objective functions in ManQ, only seven objective functions, whereas the rest objective functions are have been removed from the original ManQ and therefore create R-ManQ.

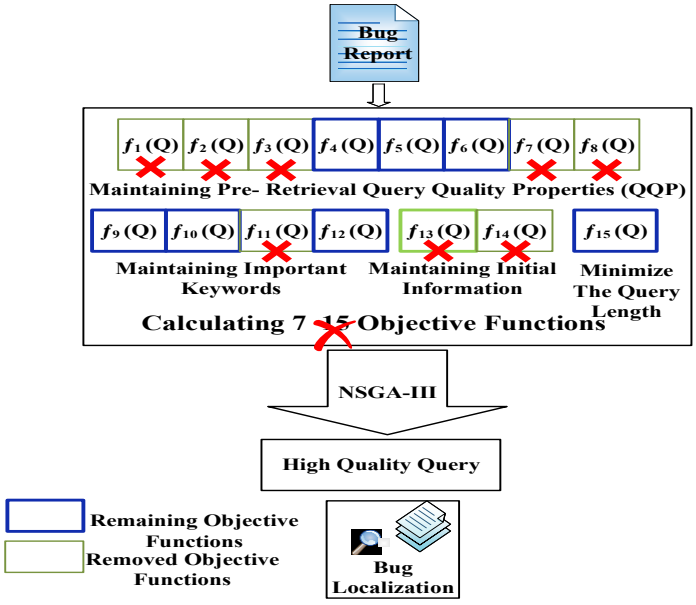


Figure. 3 proposed work R-ManQ

## 3. RESULTS AND DISCUSSION

This section discusses the evaluation metrics, used data set, and the obtained results from the proposed work and compares them with the base work ManQ, as well as the previous studies identified in [15] , which are, BLIZZARD, STRICT, and an initial query (INIT).

### 3.1. Evaluation Metrics:

The proposed work, R-ManQ, has been evaluated by a set of evaluation metrics in order to validate its effectiveness and efficiency. They are:

#### 3.1.1 Performance Evaluation Metrics:

To evaluate the effectiveness of the proposed work, performance measures are to be used, which are used in most of the related work to evaluate the performance of IRBL. These metrics are the Top-N (N = 1, 5, 10), mean average precision at 10 ranks (MAP@10), and mean reciprocal rank at 10 ranks (MRR@10).

**TOP N Rank:** It returns the percentage of queries that contains at least one relevant file (buggy files) in the first N files of the resulted list. The values of N (1,5,10) will be used in this work [3, 5, 6, 12, 14, 15, 18, 24].

**Mean Average Precision@10 (MAP@10)** calculates the mean of the average precision of all queries, whereas Precision @ 10 calculates the precision when each related result occurs in the sorted list (in 10 ranks) [3, 5, 6, 12, 13, 15, 18, 24].

$$\text{MAP@10} = \frac{\sum_{i=1}^{10} AP_i}{N}, \text{ where } \quad \text{AP@K} = \sum_{k=1}^{D} \frac{P_k \times \text{buggy(k)}}{|S|} \tag{5}$$

**Mean Reciprocal Rank@10 (MRR@10)** is the mean of the reciprocal of the position of the first buggy file within the top-10 results. The reciprocal rank for a query is the inverse rank of the first relevant document found. This metric evaluates how quickly the developer finds the first buggy files [3, 5, 6, 12, 13, 15, 18, 24].

$$\text{MRR@10} = \frac{1}{|Q|} \sum_{q \in Q} \frac{1}{\text{rank(q)}} \tag{6}$$

#### 3.1.2 Execution Time:

It means how long the approach took to finish its work and produce its results. it calculates the total number of the number of system clock seconds that it took.

#### 3.1.3 Query performance changed (|Qc|):

It computed the number of queries whose performance changed (|Qc|), improved ($|Q+|$) and worsened ($|Q-|$) according to the highest rank of the buggy files [11, 15].

#### 3.1.4 The average length of the queries ($|Q|$):

It computed the average length of the queries ($|Q|$) according to the number of terms in the final query of the approach [15].

### 3.2. Data set:

The used data set is the ManQ data set itself, which consists of 1546 poor queries from six open-source Java-based subject systems from two popular bug tracking systems—BugZilla and JIRA. Poor queries do not include stack traces or program entity queries [15].

### 3.3. Experiment

#### 3.3.1 Experimental settings

This study used the Eclipse IDE tool to create R-ManQ, which was written in pure Java, and also used it to obtain the results. It was implemented in an AMD Ryzen 5 processor with 2.10 GHz and a RAM of 8.00 GB.

#### 3.3.2 Experimental results and discussion

R-ManQ has reduced the objective functions that make up ManQ by adopting the objective functions supported by the empirical study [11] and dispensing with the rest of the other objectives. This section discusses the results obtained from the proposed work compared with the original ManQ, INIT, STRICT, and BLIZZARD. Table 1 shows the results of this comparison.

Tabel 1 R-ManQ, ManQ, STRICT, INIT, BLIZZARD results

| Evaluation Metrics | R-ManQ | ManQ | STRICT | INIT | BLIZZARD |
|---|---|---|---|---|---|
| **TOP-1** | 28.57% | 25.71% | 22.86% | 25.71% | 27.62% |
| **TOP-5** | 42.86% | 45.71% | 42.86% | 42.86% | 45.71% |
| **TOP-10** | 53.33% | 56.19% | 55.24% | 52.38% | 58.10% |
| **MAP@10** | 34 % | 33.25% | 30.69% | 32.05% | 34.09% |
| **MRR@10** | 36% | 35.02% | 31.42 % | 33.49% | 35% |
| **\|Q+\|** | 33% | 33% | 33% | ___ | 31.42% |
| **\|Q-\|** | 20% | 18% | 30% | ___ | 26% |
| **\|Qc\|** | 53% | 51% | 63% | ___ | 57.42% |
| **\|Q\|** | 20 | 24 | 14 | 29 | 42 |
| **Time Execution** | 15 | 38 | ___ | ___ | ___ |

According to the obtained results, the proposed work, R-ManQ, outperformed the original ManQ in all performance evaluation metrics except for TOP-5 and TOP-10. In addition, R-ManQ outperformed INIT in all performance evaluation metrics and STRICT, except in the TOP-10. Further, R-ManQ outperformed BLIZZARD in all performance evaluation metrics, except in TOP-1 and 10.

R-ManQ achieved 28.57% in TOP-1, higher than 25.71% in ManQ, 22.86% in STRICT, 25.71% in INIT, and 27.62% in BLIZZARD. Furthermore, R-ManQ achieved 42.86% in TOP-5, the same as in STRICT and INIT. R-ManQ achieved 53.33% in the TOP-10, higher than 52.38% in the INIT. However, in the TOP-5, R-ManQ obtained less value than 45.71% in ManQ and 45.71% in BLIZZARD. In addition, R-ManQ obtained (53.33% <56.19%) less than ManQ in the TOP-10, as well as less than 55.24% in STRICT and 58.10% in BLIZZARD. All those comparisons are illustrated in Figure 4.

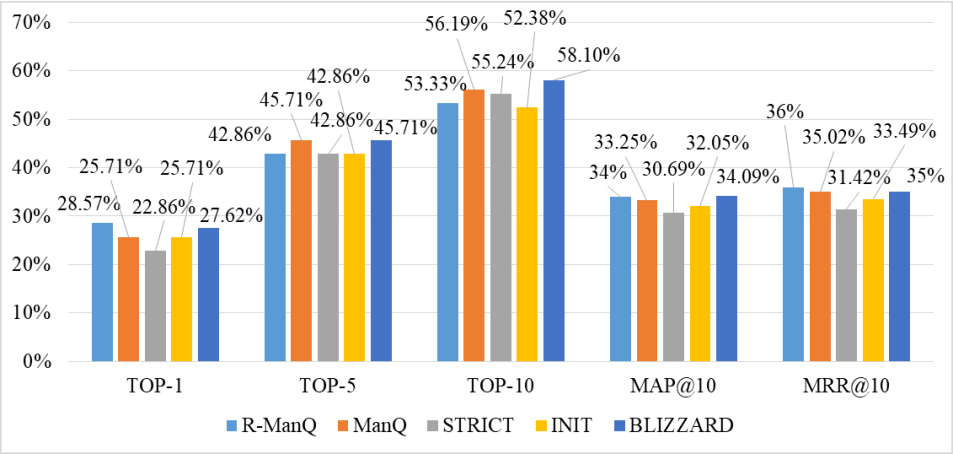


Figure. 4 The performance metrics of R-ManQ, ManQ, STRICT, INIT, BLIZZARD

However, the |*Q*+| overall queries of R-ManQ was the same as it was in ManQ and STRICT (33%), and was higher than BLIZZARD (31.42%). As shown in Table 1, the proposed work obtained better results (20%) in |Q-| than (30%) in STRICT and (26%) in BLIZZARD, while it was lower compared to ManQ (20% > 18%). Figure 5 illustrated |Qc| metrics values.

According to the number of terms, |Q|, R-ManQ decreased it to 20 compared to ManQ, INIT, and BLIZZARD (24, 29, and 42, respectively), while it was 14 in STRICT. Regarding time execution, ManQ took too long time more than the proposed work (38 > 15). Figure 6 illustrates the |Q| and execution time metrics of the proposed work compared to ManQ.

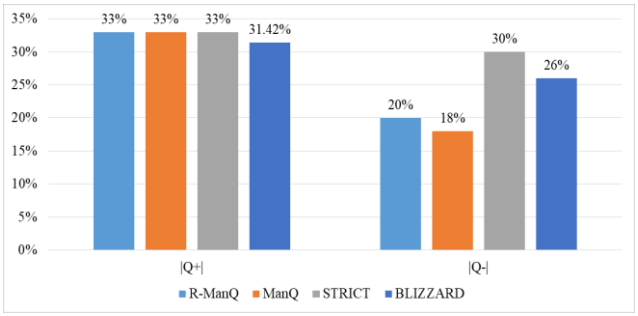


Figure. 5 |Qc| metrics R-ManQ, ManQ, STRICT, BLIZZARD

Despite reducing the number of objective functions that make up ManQ approach, the obtained results of R-ManQ were not low compared to the original ManQ, On the contrary, the proposed work has produced better results than ManQ in the three performance evaluation metrics (TOP-1, MAR@10, and MRR@10) with much less execution time than ManQ. The reason for this reduction in time is that the required processing time is reduced due to reducing the number of objective functions. However, the effect of reducing these objective functions appeared in the values of the TOP-5 and TOP-10, where their values were lower than in ManQ, which indicates that the objective functions that were excluded had an impact on the results obtained by the ManQ approach.

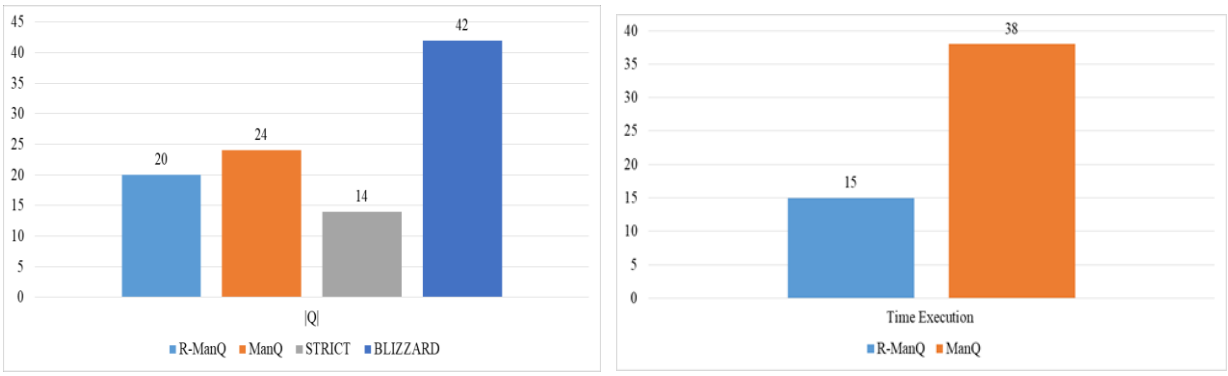


Figure. 6 Execution Time and |Q| metrics R-ManQ, ManQ, STRICT, INIT, BLIZZARD

## 4. CONCLUSION

Although many IRBL approaches were provided in order to make the bug localization process faster and with less effort, their performance is still not efficient. ManQ is one IRBL approach that improves the performance of IRBL by enhancing the quality of queries through a multi-objective optimization approach. Although ManQ consists of fifteen objective functions, [11] study claimed there are several attributes that make high-quality queries and the performance of IRBL approaches highly efficient. Therefore, R-ManQ has been proposed. It was evident that a similar performance could be achieved by R-ManQ as ManQ with a much lower execution time, a much lower number of objective functions (only seven objective functions), and a smaller number of terms in the query, where it obtained high results on TOP-1, MAP@10, and MRR@10 while failing on TOP-5 and TOP-10. Therefore, in the future, new objective functions related to graph-based term weighting algorithms will be added within the components of the proposed work R-ManQ as a step toward improving its performance. It was concluded that the attributes identified in [11] are the most important attributes that make the query high quality, which leads to an improvement in the process of locating the buggy files that contain the IRBL.

## REFERENCES

[1] A. Goyal and N. Sardana, "Performance Assessment of Bug Fixing Process in Open Source Repositories," *Procedia Computer Science,* vol. 167, pp. 2070-2079, 2020.

[2] H. Mohsin and C. Shi, "SPBC: A self-paced learning model for bug classification from historical repositories of open-source software," *Expert Systems with Applications,* vol. 167, p. 113808, 2021.

[3] R. Almhana, M. Kessentini, and W. Mkaouer, "Method-level bug localization using hybrid multi-objective search," *Information and Software Technology,* vol. 131, p. 106474, 2021.

[4] K. Kevic and T. Fritz, "Automatic search term identification for change tasks," in *Companion Proceedings of the 36th International Conference on Software Engineering*, 2014, pp. 468-471.

[5] M. M. Rahman and C. K. Roy, "STRICT: Information retrieval based search term identification for concept location," in *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, 2017, pp. 79-90.

[6] K. C. Youm, J. Ahn, and E. Lee, "Improved bug localization based on code change histories and bug reports," *Information and Software Technology,* vol. 82, pp. 177-192, 2017.
[7] M. Erşahin, S. Utku, and D. Kilinç, "Bug Localization by Using Information Retrieval and Machine Learning Algorithms," *International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES'18),,* 2018.
[8] A. Sood, S. Sharma, A. Khanna, A. Tiwari, D. Gupta, V. Madan, and S. Doss, "Bug Localization Using Multi-objective Approach and Information Retrieval," in *International Conference on Innovative Computing and Communications*, 2021, pp. 709-723.
[9] B. Dagenais and M. P. Robillard, "Creating and evolving developer documentation: understanding the decisions of open source contributors," in *Proceedings of the eighteenth ACM SIGSOFT international symposium on Foundations of software engineering*, 2010, pp. 127-136.
[10] M. Ficco, R. Pietrantuono, and S. Russo, "Bug localization in test-driven development," *Advances in Software Engineering,* vol. 2011, 2011.
[11] M. M. Rahman, F. Khomh, S. Yeasmin, and C. K. Roy, "The forgotten role of search queries in IR-based bug localization: an empirical study," *Empirical Software Engineering,* vol. 26, pp. 1-56, 2021.
[12] R. K. Saha, M. Lease, S. Khurshid, and D. E. Perry, "Improving bug localization using structured information retrieval," in *2013 28th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2013, pp. 345-355.
[13] O. Chaparro, J. M. Florez, and A. Marcus, "Using bug descriptions to reformulate queries during text-retrieval-based bug localization," *Empirical Software Engineering,* vol. 24, pp. 2947-3007, 2019.
[14] T. Dao, L. Zhang, and N. Meng, "How Does Execution Information Help with Information-Retrieval Based Bug Localization?," in *International Conference on Program Comprehension (ICPC)*, 2017.
[15] M. Kim and E. Lee, "ManQ: Many-objective optimization-based automatic query reduction for IR-based bug localization," *Information and Software Technology,* vol. 125, p. 106334, 2020.
[16] C. Mills, G. Bavota, S. Haiduc, R. Oliveto, A. Marcus, and A. D. Lucia, "Predicting query quality for applications of text retrieval to software engineering tasks," *ACM Transactions on Software Engineering and Methodology (TOSEM),* vol. 26, pp. 1-45, 2017.
[17] S. Henninger, "Using iterative refinement to find reusable software," *IEEE software,* vol. 11, pp. 48-59, 1994.
[18] M. M. Rahman and C. K. Roy, "Improving ir-based bug localization with context-aware query reformulation," in *Proceedings of the 2018 26th ACM joint meeting on European software engineering conference and symposium on the foundations of software engineering*, 2018, pp. 621-632.
[19] Q. Wang, C. Parnin, and A. Orso, "Evaluating the usefulness of ir-based fault localization techniques," in *Proceedings of the 2015 International Symposium on Software Testing and Analysis*, 2015, pp. 1-11.
[20] S. Haiduc, G. Bavota, R. Oliveto, A. De Lucia, and A. Marcus, "Automatic Query Performance Assessment during the Retrieval of Software Artifacts," in *In Proceedings of the 27th IEEE/ACM international conference on Automated Software Engineering*, 2012, September, pp. 90-99.
[21] G. Kumaran and J. Allan, "Effective and efficient user interaction for long queries," in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, 2008, pp. 11-18.
[22] G. Kumaran and V. R. Carvalho, "Reducing long queries using query quality predictors," in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, 2009, pp. 564-571.
[23] M. Ruchika, S. Aggarwal, R. Girdhar, and R. Chugh, "Bug Localization in Software using NSGA-II," *IEEE,* 2018.
[24] G. Liu, Y. Lu, K. Shi, J. Chang, and X. Wei, "Mapping bug reports to relevant source code files based on the vector space model and word embedding," *IEEE Access,* vol. 7, pp. 78870-78881, 2019.

Al-Ahgaff University Journal of Computer Science and Mathematics
مجلة جامعة الأحقاف لعلوم الحاسوب والرياضيات

Vol. 1, Issue. 1, Jan 2023, pp. 25~33

# Collecting Complex Human Activity Dataset

**Mohammed Mobark Wahdeen**
Information Technology Department, Al-Ahgaff University, Mukalla, Yemen
mabomobark@yahoo.com

**Article Info**

***Article history:***

Received jan, 9, 2023
Revised jan,23, 2023
Accepted jan 25 2023

***Keywords:***

Complex Human Activity Dataset
Third keyword

**ABSTRACT**

Existing human activity datasets involve simple activities or were collected using standalone sensors. So they do not properly match the requirement to evaluate the classifiers of the complex activities that were collected using smartphone sensors. The author collected a dataset (i.e., the Complex Activity Dataset (CAD)) to solve this problem. A group of 20 subjects was selected for this task. Data was collected in the scenario where the subject prepares breakfast. The subject performs three complex activities: preparing breakfast, preparing tea, and preparing a sandwich. Those activities are categorized into two levels in the form of a hierarchy, so that the complex activities would be placed at the high level and the simple activities at the low level. CAD was collected using the accelerometer and gyroscope sensors of smartphones. This paper presents the protocol for collecting, labeling, and filtering CAD. Also, this paper evaluated the variation property of a CAD dataset and the ability to recognize its complex human activities. The result supports the variation property of the CAD dataset and presents the ability to recognize its activities with greater accuracy than other datasets.



**الخلاصة**

بيانات النشاط البشري المجمعة الحالية تتضمن أنشطة بسيطة أو تم جمعها باستخدام أجهزة استشعار قائمة بذاتها. لذلك فهي لا تتطابق مع متطلبات تقييم المصنفات للأنشطة المعقدة التي تم جمعها باستخدام مستشعرات الهواتف الذكية. جمع المؤلف مجموعة بيانات (CAD) لحل هذه المشكلة. تم إختيار مجموعة من 20 شخصا لهذه المهمة. حيث يقوم الشخص بإعداد وجبة الإفطار. و الذي يشمل القيام بثلاثة أنشطة معقدة: تحضير الإفطار ، تحضير الشاي ، تحضير الساندويتش. يتم تصنيف هذه الأنشطة إلى مستويين في شكل تسلسل هرمي ، بحيث يتم وضع الأنشطة المعقدة على مستوى عالٍ والأنشطة البسيطة على المستوى المنخفض. تم جمع CAD باستخدام مستشعرات التسارع والجيروسكوب للهواتف الذكية. تقدم هذه الورقة بروتوكول لتجميع وتصنيف وفلترة CAD. أيضًا ، قيمت هذه الورقة خاصية التباين لمجموعة بيانات CAD والقدرة على التعرف على أنشطتها البشرية المعقدة. تدعم نتيجة هذا البحث خاصية التباين لمجموعة بيانات CAD وتوفر القدرة على إمكانية التعرف على أنشطتها بدقة أكبر من مجموعات البيانات الأخرى.

## 1. INTRODUCTION

Using mobile phones for Human Activities Recognition (HAR) is very helpful in observing the daily habits of the user and detecting health diseases or accidents early. In real-world situations, human activities are often performed in complex ways. Complex human activities are composite activities that occur concurrently or interleave. In those activities, the existence and variations of each activity, as well as the order and length, may vary. In this research, the composition and variations of human activity were considered as factors that impact the complexity of human activities.

Existing human activity datasets involve simple activities or were collected using standalone sensors. So they do not properly match the requirement to evaluate the classifiers of the complex activities that were collected using smartphone sensors. For example, the Opportunity [1] and UCI-HAR [2] datasets, probably the two most popular, are cases in point. The Opportunity dataset, which contains complex activities, was performed by 12 subjects. But the activities were collected using the inertial measurement unit, which contains a standalone accelerometer and gyroscope.

On the other hand, the UCI-HAR dataset contains inertial data collected from 30 subjects who performed a set of common daily activities while carrying a smartphone. It provides data collected from the smartphone's accelerometer and gyroscope. But the subjects performed simple activities such as walking (straight, upstairs, downstairs), sitting, standing, and lying down. The author collected CAD to solve this problem. The following sections present the protocol for collecting, labeling, and filtering CAD. Also they

show the experiments to evaluate the variation property of CAD dataset and the ability to recognition its complex human activities.

## 2. METHODOLOGY

CAD was collected using an Android (Samsung SM-G935F) that is commonly utilized [3]. It was attached to the upper arm and forearm of the subject. This smartphone contains a tri-axial accelerometer and a gyroscope derived from the findings of [4]. We want to see the effect of using the smartphone gyroscope sensor in addition to the accelerometer to smooth the data and increase recognition accuracy. Smartphone sensors collected data about 30 subjects who perfomed complex activities.

In addition to the requirement that the human activities should be collected using mobile phone sensors, the collected activitiy should be complex. There are several factors that make recognising complex human activities a challenge. In this research, the composition and variations of human activity were examined as factors that impact the complexity of the human activity recognition. For example, the complex activity comprises more than one activity that might be performed in changing order such as in an interleave or parallel manner (composition property). However, the activities should be in particular structures and sequences to be recognized by current recognition methods [5]. Also, several factors can affect the performance of the activity such as physical body differences or the environmental state in which the activity is performed. Hence, the same activity may be performed differently by different subjects (variation property).

In our own dataset, the complex activities are organized into levels to reflect the composition property of human activities whereby the high level consists of complex activities such as making sandwiches or preparing tea. Meanwhile, the low level contains meaningful, elementary (basic) movements of a person's body parts to perform the complex activity, for instance, stretching an arm or raising a leg. Table 1 shows the levels of activities in our own dataset.

Table 1 Hierarchal labelling of complex activity

| High level activities | Medium level activities | Low level activities |
|---|---|---|
| Preparing Tea | Get boiling water | Shoulder Extension |
| | Add tea | Shoulder Internal Rotation |
| | Mix the tea | Wrist rotation |

### 2.1. Data collection setup

The selected smartphone to conduct our experiment was an Android (Samsung SM-G935F) that is commonly utilized [3]. Table 2 shows the specifications of the smartphone used. It was attached to the upper arm of the subject. The smartphone contains a tri-axial accelerometer and a gyroscope derived from the findings of [4]. The sensors record timestamp motion data at the "fastest" sampling rate which can reach a maximum of 80 Hz [6]. The selected sampling rate for acquiring the body movement is contained within frequency components below 20 Hz as recommended by [3]. It is equipped with a SensorDataCollector program for collecting subject data and for storing it in a log file at SD card.

Table 2 Specifications of the smartphone used

| Device | Smartphone |
|---|---|
| Brand | Samsung SM-G935F |
| CPU | Exynos 8890<br>8 Cores (Octa-Core) |
| ROM Memory | 64GB |
| RAM Memory | 4GB |
| Operating System | Android v6.0.1 (Marshmallow) |

| Device | Smartphone |
|---|---|
| Accelerometer And Gyroscope | Sample rate is set to fastest which can reach a maximum of 80 Hz. |
| Make | STM. |
| Model | K6DS3TR. |
| Power | 0.2500 mA. |
| Rang | 8.0 g. |
| Resolution | 0.002394 m/s^2. |
| Mobile Network Type | HSUPA (High-Speed Uplink Packet Access). |
| Battery Capacity | 3600mAh. |

## 2.2. Data collection protocol

The experiment was carried out to obtain the HAR datasets. A group of 20 subjects were selected for this task based on the findings of [7]. Data of the subjects is presented in Table 3. Each subject was instructed to follow a protocol of activities while carrying the selected smartphone in his upper arm to infer overall body motion. Data was collected in the scenario where the subject prepares breakfast. This scenario has been used extensively in other works in literature [8]. The subject performs three complex activities: preparing breakfast, preparing tea, and preparing a sandwich. Those activities are categorized into two levels in the form of a hierarchy so that the complex activities would be placed in the high level and the simple activities in the low level [9]. The details of the low level activities in each complex activity are as follows:

**A. Preparing:**
Lying down on the deckchair.
Getting up.
Retrieving bread, cheese, cup, tea, sugar, plate, spoon, and knife from the cupboard and putting them on the cooking table.
**B. Preparing Tea (Pre. Tea):**
Getting a cup of water from the water boiler machine.
Adding the tea and sugar.
Mixing the tea.
Putting the cup on the dining table.
**C. Preparing Sandwich (Pre. Sandwich):**
Making bread and cheese sandwich at the cooking table.
Heating it in the microwave.
Putting it on the dining table.

Each activity lasted a minute and was repeated twice for each subject. The duration of the entire experiment was around 15 minutes per person excluding the setting up of the sensors and the repetition of the protocol. The collected human activities were designed to closely represent the natural world in both the style and time of action classes executed. The subject is free to perform the sequence of activities, so we will get activities with wide range of variations. Also, there was no time limitation on the execution of each task, so some tasks took naturally longer than others. Table 4 shows the proportion of classes in our dataset.

Table 3 Data of the subjects

| Sex | Male (11)+female (9) |
|---|---|
| Age | 24-49 years |
| Average length of upper arm | 29.28 cm |

| | |
|---|---|
| Average length of Forearm | 23.11 cm |

Table 4 Proportions of our dataset classes

| Class | Instances | Proportion |
|---|---|---|
| Preparing | 1383 | 19% |
| Preparing Tea | 2300 | 31% |
| Preparing Sandwich | 3648 | 50% |
| Total | 7331 | 100% |

### 2.3. Data Labeling

Once data was collected from the experiment, the log files were filled in order to generate the HAR datasets. Firstly, smartphone and video signals were synchronized manually by specifying the start and end of the basic movements and complex activities. All the labels of the experiments were collected in a file (labels file) which was used as one of the inputs for the dataset generation process.

### 2.4. Signal filtering

Most of the time, raw sensor signals from the accelerometer and the gyroscope are noise, so they should be preprocessed by a series of filters. We used the following filters to utilize the best performance of each sensor after carefully examining the sensor's dynamic models:

- Low pass filter that only allows signals with lower frequencies than certain cutoff frequencies. It was used to extract the low frequency of the accelerometer. The break frequency of low-pass filter was chosen at 2.5 rad/s.
- High pass filter that only allows signals with higher frequencies than certain cutoff frequencies. It was used to extract the high frequency of gyroscope. The break frequency of high pass filter was selected at 3.3 rad/s.

### 2.5. Time window size

The labeled and preprocessed signals are segmented into time window samples. Every window supposedly has an associated activity. We used fixed-width sliding windows and assigned them into activity label with 50% overlap between windows. The overlap avoids any missing activity data that begins during the time window and continues into the next one when splitting the data into segments. We evaluated the set of sizes {1, 2, 4, and 8} seconds following the recommendations of [10]. We chose a one second time window to segment our data. This decision is based on preliminary experiments which showed that using one second time window yields the best recognition accuracy.

## 3. PERFORMANCE EVALUATION

In addition, this research evaluated the variation property of CAD dataset and the ability to recognition its complex human activities. For this purpose, the following two experiments were conducted:

### 3.1. Experiment I –Evaluating the variation property of CAD

The experiments were conducted to check the variation property of the collected human activities in our own dataset. Different subjects were chosen to perform three selected tasks namely boiling water, adding tea, and mixing the tea. These tasks were chosen because each task is represented by basic movements of different arm joints as shown in Table 5. The F-test measure of analysis of variance (ANOVA) was used to investigate the effect of the variations of those task and subject factors in the wrist velocity. The two factors were tested and verified statistically by 20 x 3 (subject x task) ANOVA analysis at probability levels ($p<0,05$).

Table 5 Hierarchical labelling of complex activity

| High level activities | Medium level activities | Low level activities |
|---|---|---|
| Preparing Tea | Get boiling water | Shoulder Extension |

| | |
|---|---|
| Add tea | Shoulder Internal Rotation |
| Mix the tea | Wrist rotation |

The result of the F-test for one task performed by four subjects is presented in Table 6. Meanwhile, the result of the F-test for the three tasks performed by the same subject is shown in Table 7. Table 6 presents the result of the F-test for the (get boiling water) task performed by four subjects. It shows that the variations between the subjects (3.82E+09) who performed the same task are greater than the variations inside the task (2.93E+08). Also, it displays that the P -value (1.74E-8) is more than 0,05 showing that there is no significant difference in the four subjects when performing the same task.

Table 6 F-test for one task performed by four subjects

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Between Groups | 1.15E+10 | 3 | 3.82E+09 | 13.05485 | 1.74E-8 | 2.606929 |
| Within Groups | 1.29E+12 | 4396 | 2.93E+08 | | | |
| | | | | | | |
| Total | 1.3E+12 | 4399 | | | | |

Table 7 presents the result of the F-test for the three tasks performed by the same subject. It shows that the variations between the tasks (2.07E+12) are greater than the variations inside the same task (1.21E+08). Also, it displays that the P-value (0) is less than 0,05 showing that there is a significant difference in the three tasks when performed by the same subject.

Table 7 F-test for three tasks performed by the same subject

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Between Groups | 4.15E+12 | 2 | 2.07E+12 | 17127.18 | 0 | 2.998456 |
| Within Groups | 3.99E+11 | 3297 | 1.21E+08 | | | |
| | | | | | | |
| Total | 4.55E+12 | 3299 | | | | |

The result supports the variation property of the collected human activities in our own dataset. There are variations in the performed task more than the variations in the subject.

### 3.2. Experiment II –Evaluating the ability to recognize the complex activities of CAD

In addition, experiments were conducted to check how far we could recognize the complex human activities collected in our datasets. Our recognition system, the Complex Activity Recognizer through Wrist Velocity (CARWV) [14], was used to recognize the complex human activities that were collected using the accelerometer and gyroscope of smartphones in CAD. The result was compared, to check the ability to recognize them, with the Oppurnity dataset [1] that was collected using the Inertial Measurement Unit that contains a standalone accelerometer and gyroscope.

#### 3.2.1 Opportunity dataset

We used the Opportunity dataset to test our system's capability (CARWV) in generalizing the recognized complex human activities that were collected using standalone sensors. The Opportunity dataset was collected from four subjects who performed 17 different Activities of Daily Life (ADLs). In our experiments, we chose the scenario that consists of four high level activities i.e. Early morning moving, Coffee time, Sandwich time, and Cleanup. These high level activities and their low level activities are further described below.

**A. Early morning moving:**

Getting up

Opening the door

Closing the door

Walking

**B. Coffee time:**

Sipping

Opening drawer

Closing drawer

Reaching for an item

**C. Sandwich time:**

Slicing

Opening the fridge

Closing the fridge

Reaching for an item

**D. Cleanup:**

Opening dish washer

Closing dish washer

Reaching for an item

Moving item

Releasing item

Wiping

The Opportunity dataset includes a high number of instances of different gestures recorded by a high number of on-body, environmental and object-attached sensors at a sampling frequency of 30Hz. In our experiments, we chose the Inertial Measurement Unit that contains standalone accelerometer and gyroscope to evaluate our system. It was placed at the Right Upper Arm (RUA) of the subjects.

### 3.2.2 Experiment setup

To evaluate our system (CARWV), a 5-fold cross validation method was used. In 5-fold cross validations, the dataset is randomly divided into 5 groups (folds) of equal sizes. Each time, one fold is taken as a testing set whilst the remaining is used for training our system. This process is repeated 5 times before arriving at the final performance by taking the average of test errors that resulted from each step. The K folds cross validation system incurs less computational cost compared to other validation systems.

The recognition performance of our system was measured by two performance metrics: accuracy and F1 measure. We used the recognition accuracy metric to measure the performance of our system because it is a popular measure in the literature of human activity recognition ([11], [12], and [9]). But recognition accuracy is affected by imbalanced classes in the dataset, so we also used the F1 measure that is independent of the class distribution and measures the effect of false negatives and false positives. These two metrics (i.e. accuracy and F1 measure) had been used in previous works ([5], [13], and [9]) which makes the comparison easier. The F1 measure is the mean of precision and recall metrics. The evaluation process was simulated using MATLAB R2018b on a notebook computer with Intel i7-7700K CPU and 8GM RAM.

### 3.2.3 Experiment result

Firstly, this experiment tests the capabilities of our system to recognize complex human activities in our own dataset. For this purpose, the 5-fold cross validation method was used. The results are shown in Table 8 which shows the confusion matrix of complex activities in our own dataset when applying our system using the 5-fold cross validation. The total recognition accuracy of applying the proposed system to recognize complex activities in our own dataset is 86.2 percent and all classes obtained more than 55 percent.

The complex activity of Preparing Sandwich consists of three simple activities (i.e. Making bread and cheese sandwich at the cooking table, Heating it in the microwave, Putting it on the dining table) in which each one involves a number of basic arm movements. It obtained the highest recognition accuracy (93 percent) with little confusion with other classes. The misclassification of Preparing Sandwich activity with other activities (42 percent with Preparing Tea and 20 percent with Preparing activity) may be because it is a dominant class that shares similar abduction and adduction basic arm motions to reach items such as the bread, cheese, tea, or cupboard.

The other complex activity in our dataset is (Preparing) which consists of three simple activities (i.e. Lying down on the deckchair, Getting up and Retrieving bread, cheese, cup, tea, sugar, plate, spoon, and knife from the cupboard, and Putting them on the cooking table). Each simple activity involves a set of basic arm movement. The CARWV recognized it with a 73 percent accuracy with 20 percent confusion for the Preparing Sandwich activity. The lowest percentage of recognition accuracy was for the Preparing Tea activity with 55 percent and 42 percent confusion for the Preparing Sandwich activity. The complex activity (Preparing Tea) consists of four simple activities with basic arm movements (i.e. Getting a cup of water from the water boiler machine, Adding the tea and sugar, Mixing the tea, and Putting the cup on the dining table). The results show the ability of our system to recognize complex human activities with relatively high recognition accuracy (92-55 percent) in our own dataset.

Table 8 Confusion matrix of our dataset classes

| Classes of Complex Activity dataset | Cross validation (Folds=5, Total accurate=86.2%) | | |
|---|---|---|---|
| | Preparing | Preparing Tea | Preparing Sandwich |
| Preparing | **73** | 7 | 20 |
| Preparing Tea | 3 | **55** | 42 |
| Preparing Sandwich | 2 | 5 | **93** |

Table 9 Confusion matrix of Opportunity classes

| Classes of Opportunity | Cross validation (Folds=5, Total accurate =87%) | | | |
|---|---|---|---|---|
| | Coffee time | Cleanup | Sandwich time | Early morning moving |
| Coffee time | 78 | 0 | 0 | 22 |
| Cleanup | 0 | 58 | 2 | 40 |
| Sandwich time | 0 | 2 | 48 | 50 |

| | | | | |
|---|---|---|---|---|
| Early morning moving | 2 | 2 | 4 | 92 |

Secondly, Table 9 shows the confusion matrix of classes in the Opportunity dataset when the 5-fold cross validation system was applied. The total recognition accuracy of applying the proposed system to recognize complex activities in the Opportunity dataset is 87 percent compared to the total recognition accuracy in our own dataset. It was noticed that the total recognition accuracy of our system in the Opportunity dataset (87 percent) is better than that in our own dataset (86.2 percent). This might be in part due to the use of standalone sensors to collect the Opportunity dataset instead of using smartphone sensors as how it was carried out in our own dataset. Dernbach et al. (2012) noted that the capabilities of standalone sensors are better than the ones used in smartphones for acquiring data.

All classes obtained more than 48 percent. The Early Morning Moving complex activity obtained the highest recognition accuracy with 92 percent. It consists of four simple activities (i.e. Getting up, Opening the door, Closing the door, and Walking) in which each one involves a number of basic arm movements. The misclassification of the Early Morning Moving activity with other activities (50 percent with Sandwich time, 40 percent with Cleanup, and 22 percent with Coffee time activity) may be because it is a dominant class that shares similar flexion and extension basic arm movements when the arm swings during walking, opening doors, cutting bread, wiping, and moving hand near mouth to sip coffee.

The next activity was Coffee time which obtained 78 percent with a 22 percent confusion with Early Morning Moving activity. It consists of four simple activities (i.e. Sipping, Opening drawer, Closing drawer, Reaching for an item). Each simple activity involves a set of basic arm movements. Following that is the complex activity of Cleanup which consists of six simple activities with their basic arm movements (i.e. Opening dish washer, Closing dish washer, Reaching for an item, Moving item, Releasing item, and Wiping). It received a 58 percent accuracy with a 40 percent confusion with Early Morning Moving activity. The lowest percentage was 48 percent for Sandwich time activity with 50 percent confusion with Early Morning Moving activity with 44 percent difference between its accuracy and the one of the best performing class. This activity consists of four simple activities (i.e. Slicing, Opening the fridge, Closing the fridge, and Reaching for an item) in which each one involves a number of basic arm movements. The results of this experiment show the ability of our system in recognizing complex human activities with relatively high recognition accuracy on the two datasets.

The hierarchical structure of activities and result of F-measure of Anova in our own dataset shows the sufficiency of our dataset in representing the required two factors for evaluating the complex activities which are the ability to recognize complex, with variation human activities which were collected using mobile phone sensors. The experiments show also the ability to recognize its activities with accuracy more than other datasets.

## 4. CONCLUSION

The protocol for collecting, labeling, and filtering CAD is presented in this work. It also assessed the CAD dataset's variation property and its ability to recognize complex human activities. The outcome validates the CAD dataset's variation property and demonstrates the capacity to recognize its activities with more accuracy than other datasets. As a result, it might be used to evaluate the classifiers of complex activities recorded using smartphone sensors.

## ACKNOWLEDGEMENTS

The author thanks Alahgaff University for giving him this opportunity to present and publish his research.

## REFERENCES

[1] Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S. T., Tröster, G., Millán, J. D. R., & Roggen, D. "The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition", Pattern Recognition Letters, 34(15), 2033-2042. 2013.
[2] Ortiz, J. L. R. *Smartphone-based human activity recognition*. Springer. 2015.
[3] Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. "A survey of online activity recognition using mobile phones", *Sensors*, 15(1), 2059-2085. 2015.
[4] Wu, W., Dasgupta, S., Ramirez, E. E., Peterson, C., & Norman, G. J. "Classification accuracies of physical activities using smartphone motion sensors", *Journal of medical Internet research*, 14(5), e130. 2012.
[5] Liu, L., Peng, Y., Wang, S., Liu, M., & Huang, Z. "Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors", *Information Sciences*, 340, 41-57. 2016.
[6] Dernbach, S., Das, B., Krishnan, N. C., Thomas, B. L., & Cook, D. J. Simple and complex activity recognition through smart phones. *Intelligent Environments (IE), 2012 8th International Conference*. June. IEEE. 214-221. 2012.
[7] Su, X., Tong, H., & Ji, P. "Activity recognition with smartphone sensors", *Tsinghua Science and Technology*, 19(3), 235-249. 2014.

[8] Lukowicz, P., Pirkl, G., Bannach, D., Wagner, F., Calatroni, A., Förster, K., ... & Doppler, J. Recording a complex, multi modal activity data set for context recognition. *Architecture of Computing Systems (ARCS), 2010 23rd International Conference*. February. VDE. 1-6. 2010.
[9] Saguna, S., Zaslavsky, A., & Chakraborty, D. "Complex activity recognition using context-driven activity theory and activity signatures", *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(6), 32. 2013.
[10] Bao, L., & Intille, S. S. Activity recognition from user-annotated acceleration data. *International Conference on Pervasive Computing*. April. Berlin, Heidelberg: Springer. 1-17. 2004.
[11] Zhan, Y., & Kuroda, T. "Wearable sensor-based human activity recognition from environmental background sounds", *Journal of Ambient Intelligence and Humanized Computing*, 5(1), 77-89. 2014.
[12] Liu, L., Peng, Y., Liu, M., & Huang, Z. "Sensor-based human activity recognition system with a multilayered model using time series shapelets", *Knowledge-Based Systems*, 90, 138-152. 2015.
[13] Vaka, P., Shen, F., Chandrashekar, M., & Lee, Y. PEMAR: A pervasive middleware for activity recognition with smart phones. *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference*. March .IEEE. 409-414. 2015.
[14] M. M. Wahdeen, "Task oriented feature extraction for complex human activity recognition," Ph.D. dissertation, UTM, Malaysia, 2022.