International Journal of Science and Engineering Applications
Volume 5 Issue 1, 2016, ISSN-2319-7560 (Online)

Application of K-Means Clustering Algorithm for Classification of NBA Guards

Libao ZHANG
Department of Computer Science, Shandong University of Science and Technology, Qingdao 266590, China

Faming LU
Department of Computer Science, Shandong University of Science and Technology, Qingdao 266590, China

An LIU
Department of Computer Science, Shandong University of Science and Technology, Qingdao 266590, China

Pingping GUO
Department of Computer Science, Shandong University of Science and Technology, Qingdao 266590, China

Cong LIU*
Department of Computer Science, Shandong University of Science and Technology, Qingdao 266590, China

Abstract: In this study, we discuss the application of K-means clustering technique on classification of NBA guards, including determination category number, classification results analysis and evaluation about result. Based on the NBA data, using rebounds, assists and points as clustering factors to K-Means clustering analysis. We implement an improved K-Means clustering analysis for classification of NBA guards. Further experimental result shows that the best sample classification number is six according to the mean square error function evaluation. Depending on K-means clustering algorithm the final classification reflects an objective and comprehensive classification, objective evaluation for NBA guards.

Keywords: K-Means clustering algorithm, NBA guards, classification number

1. INTRODUCTION

In this study, K-means clustering technique is applied to the classification and evaluation for NBA guards. Recently, the classification of NBA guards is mainly based on the starting lineup, time, points and rebounding [10]. Meanwhile, starting point guard, reserve guard, point guard and offensive guard are also frequently used in traditional classification methods.

According to traditional classification methods, researchers needed to assign classification threshold to each indicator manually, which was so subjective that some particular players could not be classified in a logic situation. In this study, K-Means clustering technique origin's from machine learning field is applied to the classification of NBA guards. In order to realize the objective and scientific classification of NBA guards, this study depends on NBA 2014-15 season guards' data which is standardized and processed by mathematical models and Java language. In this way, the guards' type could be defined scientifically and properly based on classification result. Meanwhile, the guards' function in the team could be evaluated fairly and objectively. K-means clustering and improvements is widely used in present study, such as network intrusion detection [3], image segmentation [4], and customer classification [5] and so on. A cluster analysis of NBA players are very common, but their works mainly focus on the position of players.

2. K-MEANS APPLICATION

Cluster analysis is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups. It is an important human behavior. K-means algorithm [1, 2] is the most classic division-based clustering method, is one of the ten classical data mining algorithms. The basic idea of K-means algorithm is: k point in the space as the cluster centroids to cluster, classify their closest objects [8]. Through an iterative approach, in each successive update the value of cluster centroids until get the best clustering results so that the obtained clustering satisfy objects in the same cluster have high similarity and at the same time objects in the different cluster have low similarity. Therefore, based on K-means clustering algorithm one can identify the guard's function in the team, and helps people to obtain an objective evaluation about guard's ability.

2.1 K-means models establishment

2.1.1 Data filtering and processing

The data of tables obtained from DATA-NBA (www.stat-nba.com), as shown in Table 2-1, As the main task of guard is the score, rebounds and assists, so we can select these three data items as data factors for distance calculation of clustering analysis. In addition, assists and score are different, a player 10 assists in the difficulty, not less than 20 points, if not to take the measures standard, the cluster will not be fair, the score will become the main indicator, and rebounds and assists will become a secondary indicator. So we use the following equation to deal with the data processing.

$$SP_{ij} = \frac{C * n * P_{ij}}{\sum_{i=1}^{n} P_{ij}}$$

$SP_{ij} = the$ standard value of the player_{ij}

$P_{ij} = the$ original value of the player_{ij}

$C =$ auxiliary parameter for data amplification

$n =$ the total number of players in a dataset

2.1.2 K-means algorithm defect

K-means algorithm has some drawbacks [4]: First, the number of k cluster centers need to be given in advance, but in practice the selected k value is very difficult to estimate. It is extremely difficult to know how many types of data collection should be divided in advance. Second, K-means need to artificially determine the initial cluster centers, different initial cluster centers may lead to a completely different clustering results.

Table 2-1: The part of the original data

Player	Season	Team	Rebounds	Assists	Steal	Blocks	fault	foul	Scores
Russell - Westbrook	14-15	hunderbol	7.3	8.6	2.1	0.2	4.4	2.7	28.1
James - Harden	14-15	Rockets	5.7	7	1.9	0.7	4	2.6	27.4
Stephen Curry	14-15	Warriors	4.3	7.7	2	0.2	3.1	2	23.8
Kobe Bryant	14-15	Lakers	5.7	5.6	1.3	0.2	3.7	1.9	22.3
Carey - Owen	14-15	Knight	3.2	5.2	1.5	0.3	2.5	1.9	21.7
Klein - Thompson	14-15	Warrior	3.2	2.9	1.1	0.8	1.9	1.6	21.7
Dwyane - Wade	14-15	Heat	3.5	4.8	1.2	0.3	3.4	1.7	21.5
Damian - Lillard	14-15	Blazers	4.6	6.2	1.2	0.3	2.7	2	21
DeMar - DeRozan	14-15	Raptors	4.6	3.5	1.2	0.2	2.3	2	20.1
Kevin - Martin	14-15	mberwolv	3.6	2.2	0.8	0	1.9	1.9	19.5
Chris Paul	14-15	Clippers	4.6	10.2	1.9	0.2	2.3	2.5	19.1
Isaiah - Thomas	14-15	Celtics	2.1	5.4	0.6	0	2.6	2.1	19
Monta - Ellis	14-15	Mavericks	2.4	4.1	1.9	0.3	2.5	2.5	18.9
Victor - Oladipo	14-15	Magic	4.2	4.1	1.7	0.3	2.8	2.6	17.9
Kyle - Lori	14-15	Raptors	4.7	6.8	1.6	0.2	2.5	3	17.8

Considering the first defect, we need to evaluate different values of k in the k means clustering, and select the most reasonable k value.

Considering the second defect, we choose the initial center point by the remote-first algorithm [9]. The basic idea of the initial clustering center point lies in: the initial clustering centers should be as far as possible from the distance between each other.

Detailed steps of the k clustering center with remote-first algorithm is explained as follows:

Step1: Choose one center uniformly randomly from the data points.

Step2: For each data point x, compute D(x), the distance between x and the nearest center that has already been chosen.

Step3: Choose one new data point randomly as a new center, using a weighted probability distribution where a point x is chosen with probability proportional to D(x) 2.

Step4: Repeat Steps 2 ~ 3 until k centers have been chosen.

2.2 K-means algorithm

2.2.1 Data Preparation

In order to construct the K-means model, one needs to get the 14-15 season NBA guard data which includes 120 NBA guards' data. We standardize and filter the data, to prepare for the K-means analysis. The filtered data is stored in csv file & an excerpt of our processed data is shown in Table 2-2.

Table 2-2: 120 NBA Guard Regular Season Data

	Player	Team	Rebounds	Assists	Scores
1	Russell - Westbrook	Thunder	7.3	8.6	28.1
2	James - Harden	Rockets	5.7	7	27.4
3	Stephen Curry	Warriors	4.3	7.7	23.8
4	Kobe Bryant	Lakers	5.7	5.6	22.3
5	Carey - Owen	Cavaliers	3.2	5.2	21.7

6	Klein - Thompson	Warriors	3.2	2.9	21.7
7	Dwyane - Wade	Heat	3.5	4.8	21.5
8	Damian - Lillard	Trail Blazers	4.6	6.2	21
9	DeMar - DeRozan	Raptors	4.6	3.5	20.1
10	Kevin - Martin	Timberwolves	3.6	2.2	19.5
11	Chris Paul	Clippers	4.6	10.2	19.1
12	Isaiah - Thomas	Celtics	2.1	5.4	19
13	Monta - Ellis	Mavericks	2.4	4.1	18.9
……	……	……	……	……	……
……	……	……	……	……	……
114	Jose - Calderon	Knicks	3	4.7	17.3
115	Jason - Richardson	76ers	3.5	2	17.2
117	Quincy - Pondexter	Pelicans	3.1	1.5	17
118	Bojan - Bogdanovich	Nets	2.7	0.9	16.9
120	Marcus - Thornton	Celtics	1.9	0.9	16.6

2.2.2 Algorithm Design

Using K-means clustering algorithm for data analysis. The basic idea of K-means algorithm [11] is: allocating data set D into k clusters. To determine k clusters, we need to determine the k center C1, C2…Ck, calculate the distance to each point to the center for each point inside dataset, the point that the shortest distance from the center classified as represented by clusters.

K-means algorithm steps are explained follows:

Step1: Determine the number of K-means clustering center k;

Step2: The use of remote-first algorithm to initialize the center of k;

Step3: The points of dataset D assigned to the nearest center, forming a k clusters;

Step4: The calculation k Category cluster centroid obtained by [3], the nearest point of dataset D from the centroid as the new center;

Step5: Repeat [3] ~ [4], until the center remain stable.

Euclidean distance is calculated as follows:

$$D = \sqrt{\sum_{k=1}^{n}(P_{ik} - P_{jk})^2}$$

$$D = \text{the distance between } P_i \text{ and } P_j$$

$$P_{ik} = \text{the value of } P_i$$

$$P_{jk} = \text{the value of } P_j$$

2.2.3 K value determination

After calculation the results of the k are 2, 3, 5, 6, 7, and 8 by the k-Means algorithm, and then we use the Mean Squared Error to perform the comparison of results with different k values. The calculation formula is as follows:

$$MSE = \frac{\sum_{i=1}^{n}(P_i - PC_i)^2}{n}$$

$$n = \text{the total number of point in a dataset}$$

$$C = \text{the numbers of clustering center}$$

$$P_i = \text{the point } i$$

$$PC_i = \text{the center of the point } i$$

$$MSE = \text{the mean squared error}$$

According to Figure 2-1 and Table 2-2, we can see that as k-values gradually increase from 2 to 8, the mean square error getting smaller and smaller. Clustering result also gradually changed for the better, and the small changes of clusters to achieve a relatively stable state when the center points surpass six. This is the minimum mean squared error, it can be concluded that when the cluster number is 6, the mean squared error is becoming smaller, the similarity within the class is higher, and classification result is the best at the same time.

Table 2-2: Mean Square Error for Different Values of Time

K-Values	2	3	4	5	6	7	8
Mean square error	6.913933	5.701685	5.023356	6.27497	4.363918	4.335483	4.323653

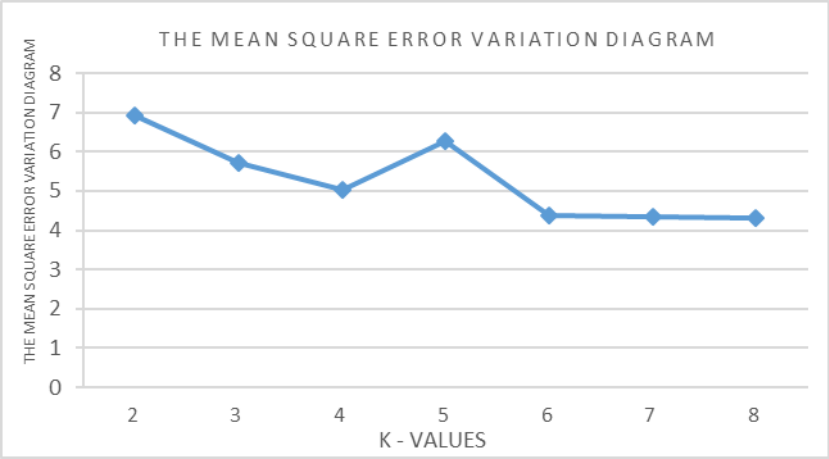


Fig2-1: The Mean Square Error Variation Diagram

3. RESULTS and EVALUATION

3.1 Classification Results

Based on the above analysis that the effect of clustering is the best when k = 6, the NBA guard can be divided into 6 categories, the classification results is shown in Table 3-1, classification and analysis of the results are as follows:

Category 1: The guards whose assist and score is well are the main shooting and point of the team. However, limited in playing time, the data is not particularly outstanding, such as Manu Ginobili, Tony Parker.

Category 2: The guards whose score ability and rebounds ability are high, with more playing time, is absolutely super guard and the core of the team, such as Harden, Curry and Westbrook.

Category 3: The guards whose score ability and rebounds ability are outstanding, assists ability is normal, are the guards of the Swingman type. They can make enough contribution to the team's defense and offense, such as Iman - Shumpert, Wesley - Matthews, etc.

Category 4: The guards who get 14.17 points and 8.27 assists, are typically assists madman, the initiator of the offense, the core and leader for a team such as Chris - Paul, John -Wall.

Category 5: The guards whose score ability are much higher than rebounds ability and assists ability, should be the team's point guard, the team's playmaker such as Wade, Owen.

Category 6: Compared rebounds ability and assists ability, score ability is the main contribution of this category guard, usually as the team's backup point guard, the outstanding ability of singles or long shot well, such as J.J-Redick, Nick - Young.

Table 3-1: K-Means Clustering Result

Classification Number	Category Centroid			Category members(Separated "|" between names)
	Rebounds	Assists	Scores	
1	2.47	4.49	11.77	Isaiah - Thomas | Brandon - Jennings | Tony - Tony Parker | Morris - Williams | Eric - Gordon | Terre - Burke | Brandon - Knight | Ismail - Smith | Jarrett - Jack | Morris - Williams | Jeremy | JR Smith | DJ- Augustin | Manu Ginobili | - Mario Chalmers | Zach - Lavin | CJ Watson | Dennis - Schroeder | Jameel-Nelson | DJ- Augustin | Gray Davis - Vasquez | Jose - Calderon |
2	5.43	6.87	20.36	Russell - Westbrook | James - Harden | - Stephen Curry | - Kobe Bryant | Damian-Lillard | Kyle - Lori | Eric - Bledsoe | Tyreke - Evans | Michael - Carter - Williams |
3	3.49	2.26	12.56	Kevin - Martin | Wesley-Matthews | Brad - Bill | Aaron - Afflalo | Avery-Bradley | Alec - Burks | Shabazz-Muhammad | A Long - Afflalo | JR Smith | Pop | Dion - Waiters | Rodney - Stuckey | Ben - Mark Lehmer | Gerald-Henderson | Jordan-Clarkson | Langston- Galloway | -Gary Neal | Will-Barton | Patrick- Beverly | Wayne-Ellington | Imran-Shumpert | Jason- Richardson | Quincy-Pondexter |
4	4.21	8.27	14.17	Chris Paul | Reggie Jackson | John Wall | Jeff - Teague | Thailand - Lawson | Zhu-Huo Ledi | Ricky - Rubio | Rajon-Rondo | Deron Williams - Williams |
5	3.68	4.78	16.61	Carey - Owen | Dwyane-Wade | Klein - Thompson | DeMar-DeRozan | Isaiah - Thomas | Monta-Ellis | Victor-Oladipo | Brandon - Knight | Derek - Ross | Kemba - Walker | Morris - Williams | Tony - Rothen | Golan - Dragic | Darren - Collison | George - Hill | - Mike Conley | Alec Frank - She Weide | Joe - Johnson | Evan - Turner |
6	2.18	1.96	11.2	J.J. Redick | Jamal - Crawford | Louis - Williams | Nick - Young | Isaiah - Buchanan | Avon - Fournier | Dion - Waiters | Aaron - Brooks | Tim - Hardaway II | OJ- Mayo | Jodi - Meeks | Anthony - Morrow | Aaron - Afflalo | AJ-Price | Alec Frank - She Weide | Courtney - Lee | - Gary Neal | Alec Frank - She Weide | Norris - Cole | Terrence - Rose | - Gary Neal | Marco – Marco Belinelli | Isaiah - Buchanan | Bojan - Bogdanovich | Marcus - Thornton |

3.2 Analysis and Evaluation

In news and media, guards are divided into point guard and shooting guard according to the arrangement in the team, and divided into key guard and reserve guard according to playing time order. Therefore, general guard has four categories: key point guard, key shooting guard, reserve point guard& reserve shooting guard. However, basketball is the athletic sports of constant adjustment and adaptation. Throughout the league process, every NBA guard assignment, as well as playing time, playing order required to make specific arrangements according to needs of the team and coach's strategy. Therefore, this intuitive classification is dependent on people's subjective judgment which is limited biased & changing. Because guards' function in the game would constantly adjustment, classification of guards should constantly adaptation, which caused a great disturbance to classification and evaluation of NBA guards macroscopically. Accordingly, the above classification and evaluation methods heavily depend on so many subjective factors, that the classification and evaluation of NBA guards are neither scientific nor objective.

In this study, the K-Means clustering analysis is applied to the classification of NBA guards. We take fully advantage of the statistical data of NBA guards to analyze data and standardize data rationally. Mining the authentic classified information, will get classification of NBA guards more scientifically and objectively. Find guards in the team's role, the ability to guards and defender in the team's performance has a comprehensive understanding and evaluation. Identify the guard's function in the team, can help people have a comprehensive understanding and objective evaluation about guard's ability and their performance has a comprehensive understanding and evaluation. Identifying the guard's function in the team could help NBA Sports News, NBA commentator and Basketball enthusiasts have a comprehensive understanding and objective evaluation about guard's ability and their performance. Furthermore，the classification results propose an effective solution for analysis the extremely big of NBA data, rather than just make statistical comparisons.

4. CONCLUSIONS

Traditionally, clustering is viewed as an unsupervised learning method for data analysis. In this study, we proposed a simple and qualitative methodology to classify NBA guards by k-means clustering algorithm and used the Euclidean distance as a measure of similarity distance. We demonstrated our research using k-Means clustering algorithm and120 NBA guards' data. This model improved some limitations, such as manual classification of traditional methods. According to the existing statistical data, we classify the NBA players to make the classification and evaluation objectively and scientifically. Experimented results show that this methodology is very effective and reasonable. Therefore, based on classification result the guards' type could be defined properly. Meanwhile, the guards' function in the team could be evaluated in a fair and objective manner.

5. REFERENCES

[1] Jiawei Han. Data Mining Concepts and Techniques [M].Beijing: Mechanical Industry Press .2006.

[2] http://en.wikipedia.org/wiki/K-means_clustering.

[3] Liu Chang Qian. K-means algorithm improvements and network intrusion detection application [J]. Computer simulation .2011.

[4] Yan Xinge .ISODATA and fuzzy K-means algorithm applied in image segmentation [C]. Chinese Optical Society 2004 Academic Conference.

[5]Qu Xiaoning .K-means clustering algorithm in commercial banking customers classification [J]. Computer simulation .2011.

[6] Raymond T. Ng and Jiawei Han, CLARANS: A Method for Clustering Objects for Spatial Data Mining, IEEE TRANSACTIONS ONKNOWLEDGE and DATA ENGINEERING. 2002.

[7] Zhu Xian based on simulated annealing Particle Swarm Optimization techniques of genetic data biclustering research [M]. Nanjing Normal University .2009.

[8]Yin Z.D .Based collaborative filtering Trusted Service Selection [M]. Nanjing University of Posts and Telecommunications.2013.

[9] Jiangwen Rui. Distributed machine learning framework based on cloud [M]. Xiamen University .2013.

[10] Data Source: http: //www.stat-nba.com/.

[11] Sun Jigui, Liu Jie, Zhaolian Yu clustering algorithm [J] Journal of Software 2008.

[12] Jin Ming. Optimization Selection and Evaluation of Technical Index Classification of NBA Elite Guard of. China Sport Science and Technology. 2005.

[13] Richard J. Roiger, Michael W. Geatz, Data Mining a tutorial-based primer, Addison-Wesley, 2003.

[14] Josef Cihlar, Rasim Latifovic, Jean Beaubien. "A Comparison Of Clustering Strategies For Unsupervised Classification," Canadian Journal of Remote Sensi.

International Journal of Science and Engineering Applications
Volume 5 Issue 1, 2016, ISSN-2319-7560 (Online)

Effect of AL_2O_3 Nanoparticles on the Rheological Properties of Water Based Mud

Amarfio, Eric Mensah
Department of Petroleum Engineering
African University of Science and Technology
Abuja, Nigeria

Abdulkadir, Mukhtah
Department of Petroleum Engineering
African University of Science and Technology
Abuja, Nigeria

Abstract: This research work investigates into the performance of Aluminium Oxide nanoparticles in water based bentonite drilling fluid at high temperature formations. We looked into the thermal stability effect of the Aluminium oxide nanoparticles on the drilling fluid at varying temperature conditions. We analyzed the interactive effects of temperature, the Aluminium Oxide nanoparticles and shear rates on the shear stress of the drilling fluid. Optimization of these parameters at the high and low point of the shear stress of the drilling fluid was analyzed. We also developed a predictive expression for Shear stress as a response variable for changes in temperature, Aluminium Oxide nanoparticle and shear rate.

Keywords: Thermal stability, Optimization, nanoparticles, bentonite drilling fluid, interactive effect, shear stress

1. INTRODUCTION

Drilling fluid plays very vital role in the drilling operations of oil and gas industries. It plays a multifunctional role such as removal of cuttings, lubrication of bits, maintenance of wellbore stability and prevention of inflow and outflow of fluids between borehole and the formation. The use of water-based muds (WBMs), oil based muds (OBMs) and other synthetic based muds (SBMs) in drilling of oil and gas wells has increased considerably over the past years. New mud systems are continuously being developed and existing systems are being refined to reduce exploration costs. Now the advancement of drilling operations into high temperature formations demand the usage of drilling fluid formulae that will withstand high temperatures by stabilizing the integrity of the rheology of the drilling fluid under such conditions [9, 11] . At high temperature conditions, better thermally stable drilling fluid is required to maintain the rheological properties of this multifunctional fluid [2, 4]. Equipping drilling fluid to perform its basic functions under such conditions requires engineering the fluid with additives. Developing a stable fluid to maintain the rheological properties is an important issue at this stage [5, 8].

This study therefore explored the use of Alpha Aluminium Oxide nanoparticles as stabilizing agent under the conditions of high temperatures.

Many other studies and operation on the field have employed different types of chemicals and polymers in designing the drilling mud to meet some functional requirements such as the appropriate mud rheology, density, mud activity, fluid loss control property. Studies of nanoparticles have shown their unique abilities in their functionalities such as thermal conductivity, electrical conductivity, optical features etc. [1, 3, 6,10]. Our aim and objectives for this work are; to evaluate the effect of temperature on the drilling fluid at varying mass fractions of the Aluminium Oxide nanoparticles and different shear rates; also to evaluate the interactive effect of temperature, Aluminium Oxide nanoparticle and the Shear rate on the Shear stress of the drilling fluid; and finally to develop a prediction model that predict the effect of the nanoparticles at higher temperature formations.

2. EXPERIMENTAL PROCEDURE

2.1 Chemical Synthesis of Iron Oxide Nanoparticles

Aluminium oxide nanoparticles used was obtained from stock with the following characteristic: Spherical morphology and particle size of 40nm and crystal size of 150nm

2.2 Procedure for Bentonite Drilling Fluid Formulation

350 ml of fresh water was measured using a measuring cylinder and was added to 22.5 g of bentonite and stirred in the bucket until no more lumps were observable by the help of an electric mixer. The drilling fluid was left to stay overnight (16 hours) to swell. The formulated bentonite mud was divided to four different samples. Aluminium Oxide Nanoparticles were then added to the formulated bentonite mud samples each in the following mass fractions 0.5 g, 1 g, and 1.5 g. The sample fluids are then stirred vigorously and homogenized with an electronic mixer for about 2 minutes to ensure stable and uniform dispersion of nanoparticles in the fluid for the study.
The bentonite fluid and the treated bentonite fluids were used for the various experiments.

3. EXPERIMENTAL RESULTS AND DISCUSSION

The viscometer analysis was completed for the aluminium oxide Np enriched drilling fluid. Analysis and discussion mainly centers on the effect on the fluid model and the thermal stability of the drilling fluids as they undergo temperature variation. Also, the statistical analysis of the parameters that affect the rheology of the drilling fluid.

3.1 Effect of Aluminum Oxide Nanoparticles on Water based Fluid Behaviour Model

Rheological properties such as yield point and plastic viscosity of the water-based drilling fluid with the presence of aluminum oxide were fairly stabilized. The behaviour can be linked to that of a Bingham plastic model as shown in figure 3-1 and 3-2 at 40 and 90 degrees Celsius temperature respectively. As we raised temperature to 90 degrees Celsius, the yield points and plastic viscosities only showed a slight variation for different mass fraction of the nanoparticles. But, the yield point for the zero nanoparticle concentration decreased as the temperature increased to 90 degrees Celsius.

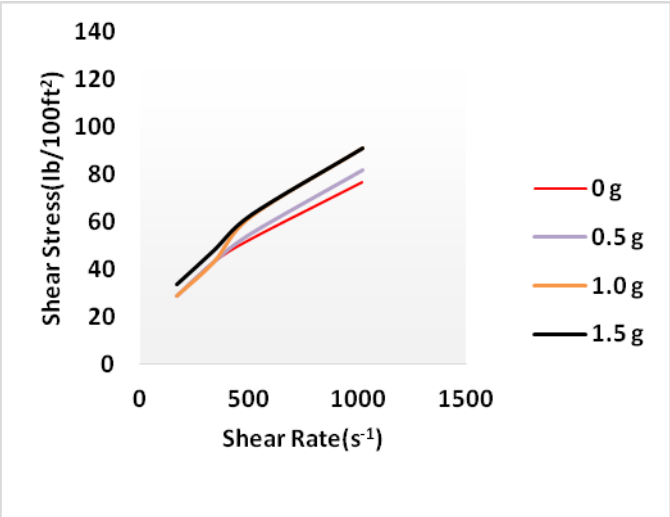


Figure 3-1: Rheogram Showing Non-Newtonian behaviour of Al_2O_3 Np Treated Water Based Mud @ 40C

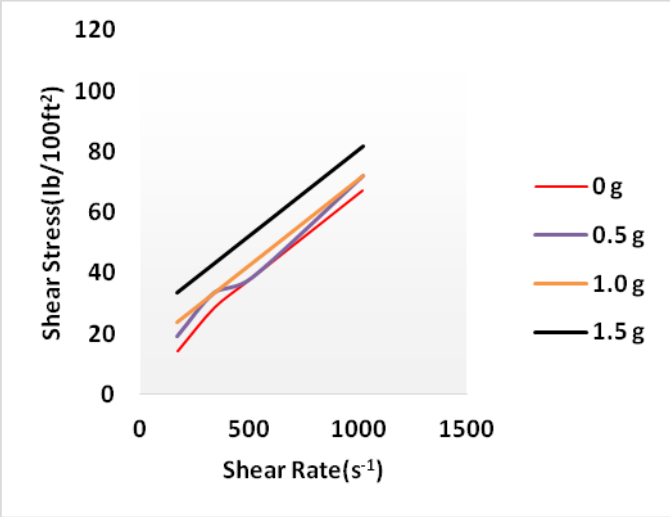


Figure 3-2: Rheogram Showing Non-Newtonian behaviour of Al_2O_3 Np Treated Water Based Mud @ 90C

3.1.1 Thermal Effect of Aluminium Oxide on Water Based Drilling Fluid

It can be seen clearly, especially in figure 3-3d, how the aluminium oxide Np actively maintained the rheology of the water based fluid across the temperature range. This indicates that for the same quantity of the drilling fluid, increasing the mass fraction of the nanoparticles increases the thermal stabilization of the drilling fluid

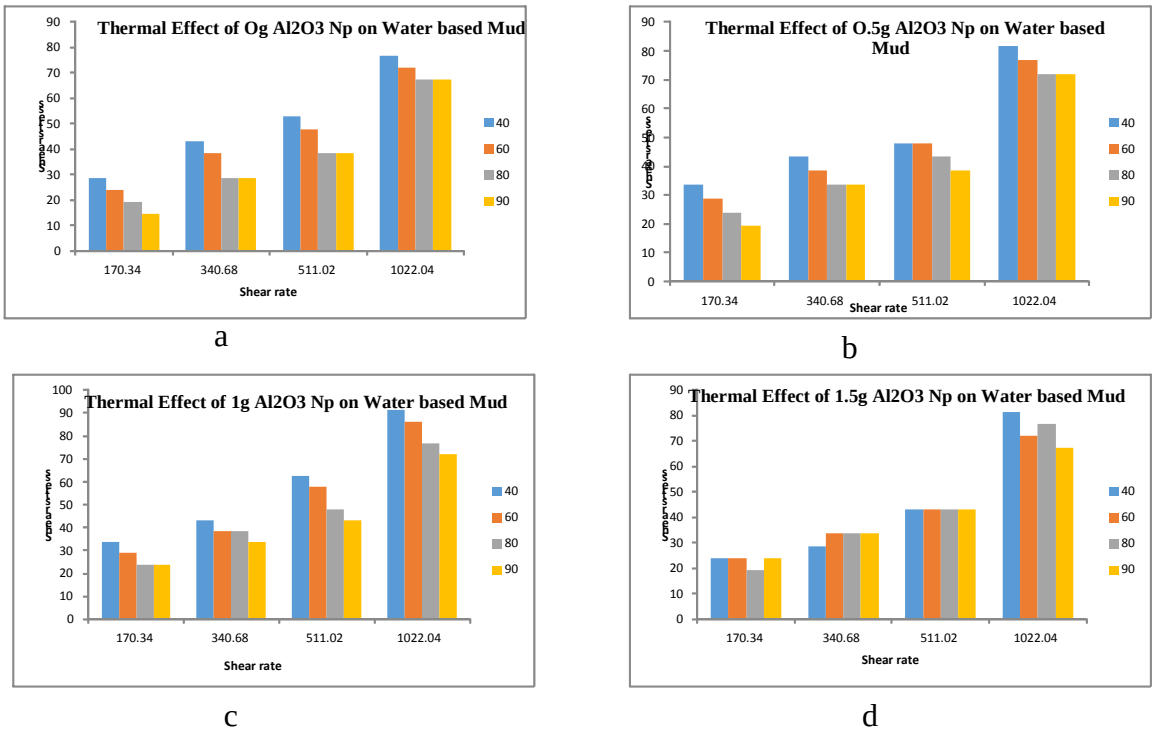


Figure 3-3: Summary of Thermal effect on Aluminium Oxide Nanoparticle treated Water Based Drilling Fluid at varying composition

3.2 Statistical Analysis

In this study we designed the experimental work using the factorial design approach[12]. This actually gave us the opportunity to analyze our results in ways that employs the multiple or combined effects of our research parameters. We used in this study the JMP software to analyze and observe the combined effects of the Aluminium Oxide nanoparticles, temperature and Shear rate on the shear stress. The following highlights the analysis of the results

Figure 3-4 below shows the actual plot of the rheology of the drilling fluid taking into consideration, the effect of Temperature and the Aluminium Oxide nanoparticles. The predictability of this experiment given by the R-square value of 0.91 indicates that the errors registered in our experiment are minimal. Therefore, the prediction model developed from this study can be used to project the effect of the nanoparticles at higher temperature regimes.

Equation 3-1 given below is our predictive factorial model or equation from the experiment.

$$S_S = 24 - 0.17T + 9.16K + 0.055B + (T - 67.5)[(K - 0.75) \times 0.036] + (T - 67.5)[(B - 511.02) \times 0.00018] + (K - 0.75)[(B - 511.02) \times 0.0057] \quad (1)$$

Where;

T= Temperature, ^{O}C

A= Shear rate, s^{-1}

k= Al_2O_3 Np, g

S_{ST}= Shear Stress, Ib/100 ft^2

Figure 3.5 is the cube plot which gives a model of how the Aluminium Oxide nanoparticles, the temperature and the sheat rate interplay at the various optimized points of the shear stress.

From this plot, it means that at the shear stress of 86.3 Ib/100ft^2, within a temperature zone of about 40 ^{o}C, a 1.5 mass fraction of the Aluminium Oxide nanoparticle must be uniformly dispersed in the drilling fluid and drilling operation set at the shear rate around1022 per second. On the other hand, minimum or no nanoparticles presence and the 90 ^{o}C temperature are required to obtain the least shear stress of 17.24 Ib/100ft^2 according to the optimization model.

The contour plot gives the performance of the Shear stress as the parameters change as in whether they increase or reduce in measure.

Figure 3-6 indicates that the Shear stress increases as Aluminium Oxide nanoparticles increases and temperature reduces.

Figure 3-7 indicates that the Shear stress increases as Shear rate increases and temperature decreases. And figure 3-8 shows that Shear Stress increases as Iron Oxide nanoparticles increases and Shear rate decreases.

Figure 3-9 is gives the interaction profile for all the parameters.

Table 3-1 shows the measured effect of the parameters on the Shear stress of the drilling fluid. The individual effects as well as the combined effect are shown. The Aluminium Oxide nanoparticles gave the highest impact which implies positive performance in stabilizing the temperature. All the combine parameters gave a negative impact on the shear stress.

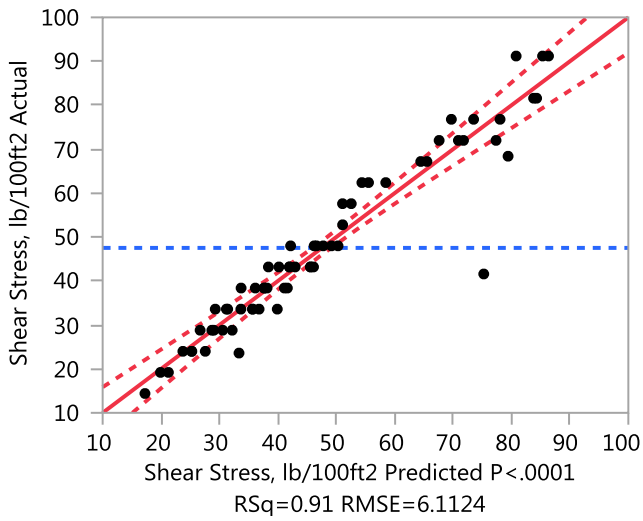


Figure 3-4: Actual by predicted plot for Shear Stress

Table 3-1: Estimated Parameters

Term	Estimate	Standard Error
Shear Rate, 1/s	**0.0554563**	**0.002398**
Al2O3 NP, g	**9.16**	**1.366763**
Temp, oC	**-0.174661**	**0.039788**
(Temp, oC-67.5)*(Shear Rate, 1/s-511.02)	**0.0001895**	**0.000125**
(Al2O3 NP, g-0.75)*(Shear Rate, 1/s-511.02)	**0.0057029**	**0.004289**
(Temp, oC-67.5)*(Al2O3 NP, g-0.75)	**0.0357966**	**0.071175**

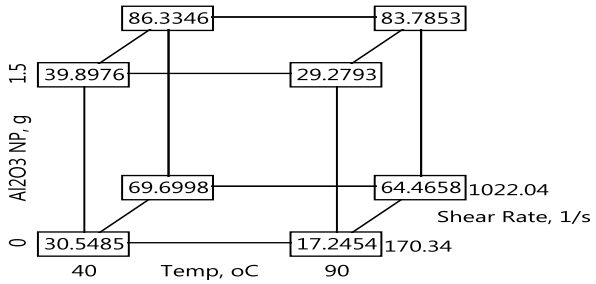


Figure 3-5: Cube plot showing optimum parameters

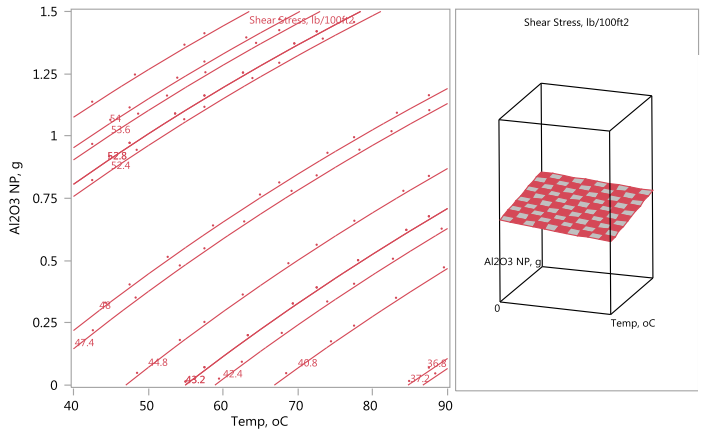


Figure 3-6: Contour Plot: Effects of Temperature/ Shear Rate on Shear stress

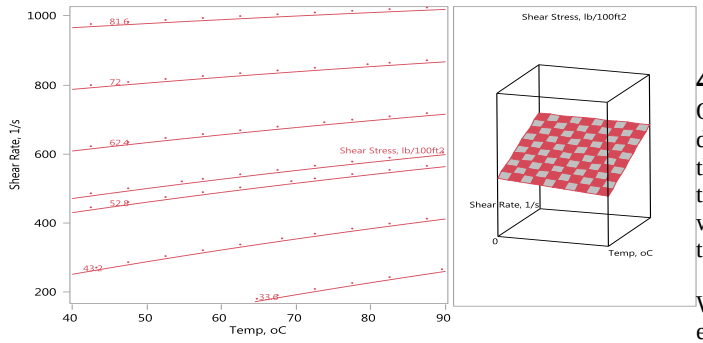


Figure 3-7: Contour Plot: Effects of Temperature/ Al_2O_3 Np on Shear stress

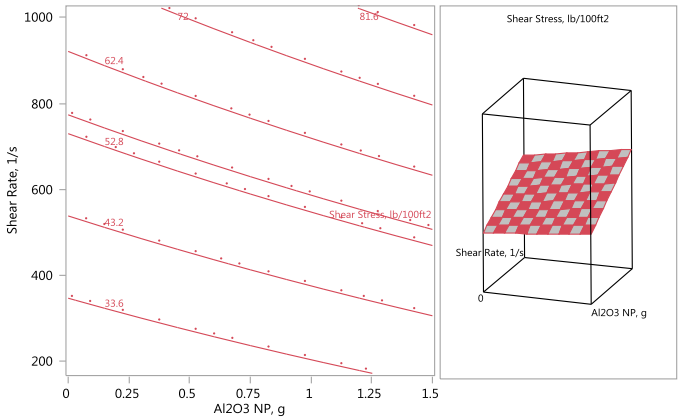


Figure 3-8: Contour Plot: Effects of Al_2O_3Np/ Shear Rate on Shear stress

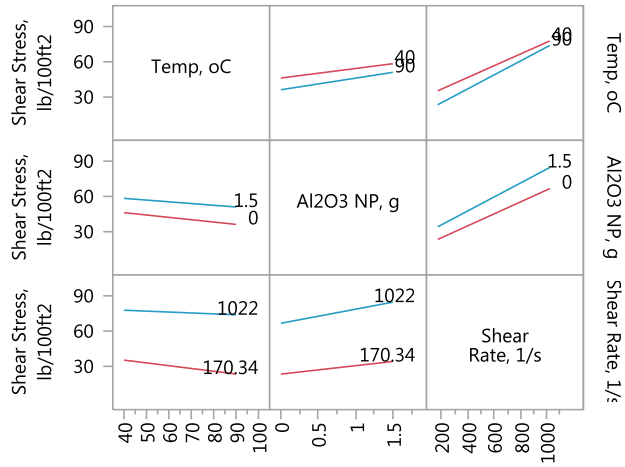


Figure 3-9: Parameters Interaction Profiles

4. CONCLUSION

Our study has shown that Aluminium Oxide nanoparticles dispersed in the water based bentonite drilling fluid provide thermal stabilization cover for the drilling fluid under the high temperature conditions. The Aluminium Oxide nanoparticles were able to maintain the shear stresses of the fluid as temperature increases at defined levels of shear rate.

We developed a predictive model to make engineering estimates of Aluminium Oxide nanoparticles mass fractions and Shear rates when drilling operations must be made at higher temperature zones above hundred degrees Celsius was generated for this study.

In this study we also showed the Interactive effects of the Aluminium Oxide nanoparticles, temperature and Shear rate and finally the cube plot that shows the optimization levels for all the parameters at the high and low levels of the Shear stress of the drilling fluid.

Nomenclature

Np= Nanoparticles

5. ACKNOWLEDGEMENT

The Authors wish to express appreciation to the University of Mines and Technology for Laboratory experiment support and African University of Science and Technology for Technical Assistance.

6. REFERENCES

1. Y. Yung et al, "Effect of Nanoparticle-Additives on the Rheological Properties of Clay-Based Fluids at High Temperature and High Pressure" AADE-11-NTCE-2, 2011

2. S. Agarwal et al, "Flow Behavior of Nanoparticle Stabilized Drilling Fluids and Effect of High Temperature Aging" AADE-11-NTCE-3, 2011.

3. V.C. Kelessidis, C. Tsamantaki, P. Dalamarinis, "Effect of PH and Electrolyte on the Rheology of Aqueous Wyoming Bentonite Dispersions," Applied Clay Science 36 (2007) 221-231.

4. N.J. Alderman, A. Gavignet, D. Guillot, and G.C Maitland, "High- Temperature, High-Pressure Rheology of Water-Based Muds," paper SPE 18035 presented at the 1988 SPE Annual Technical Conference and Exhibition, Houston, Texas.

5. C. Baird and J.Y. Walz, "The Effects of Added Nanoparticles on Aqueous Kaolinite Suspensions I. Structural effects," Journal of Colloid and Interface Science 297 (2006) 161–169.

6. T. Szabo, A. Bakandritsos, V. Tzitzios, S. Papp, L. Korosi, G. Galbacs, K. Musabekov, D. Bolatova, D. Petridis and I. Dekany, "Magnetic iron oxide/clay composites: effect of the layer silicate support on the microstructure and phase formation of magnetic nanoparticles," Nanotechnology 18 (2007) 285602-285610.

7. Hieu, P. and Quoc, P. N. (2014), "Effect of Silica nanoparticles on Clay Swelling and aqueous stability of nanoparticle dispersions", *Journal of Nanoparticle Research*, 16(1):2137

8. Kasiralvalad, E. (2014), "The great potential of nanomaterials in drilling and drilling fluid applications", *International Journal Nano Dimens*. 5(5):463-471

9. Shushant, A., Phuoc, T., Yee, S., Donald, M. and Rakesh, K. G. (2011), "Flow Behavior of Nanoparticle Stabilized Drilling Fluids and Effect of High Temperature Ageing", *AADE-11-NTCE-3*

10. Youngsoo, J., Mathew, B., Junk-kun, L., Phouc, T., Yee, S., Donald, M. and Minking, C. (2011), "Effect of Nanoparticle-Additives on the Rheological Properties of Clay-Based Fluids at High Temperature and High Pressure", *AADE-11-NTCE-2*

11. Makinde, F. A., Adejumo, A. D., Ako, C. T., and Efeovbokhan, V. E. (2011), "Modelling the effects of temperature and aging time on the Rheological Properties of Drilling Fluid", Petroleum and Coal, Vol. 3, pp. 169 - 182.

12. Douglas C. Montgomery, *Design and Analysis of Experiment*, Fifth Edition. John Wiley and Sons Inc. pp. 1-299

International Journal of Science and Engineering Applications
Volume 5 Issue 1, 2016, ISSN-2319-7560 (Online)

Isolation and Screening of Hydrogen Producing Bacterial Strain from Sugarcane Bagasse Yard Soil

Jonesh Swaminathan
Annamalai University
Annamalai Nagar- 608 002,
Tamil Nadu, India.

Meera Asokan
Annamalai University,
Annamalai Nagar- 608002,
Tamil Nadu, India.

Dhanasekar Ramasamy
Annamalai University,
Annamalai Nagar- 608002,
Tamil Nadu, India.

Abstract: The aim of this study is to isolate a highly competent bacterium with potent cellulose degrading capability and a better hydrogen producer. Soil sample from sugarcane bagasse yard was isolated, serially diluted and plated on cellulose specific nutrient agar plate. Four colonies have been isolated in which a single colony has potent cellulose degrading ability and the highest hydrogen productivity of 275.13 mL H_2 L^{-1}. The newly isolated bacterium was morphologically and biochemically characterized. The molecular characterization of the bacterium was carried out using 16S rDNA sequencing and the organism was identified as *Bacilllus subtilis* AuChE413. Proteomic analysis such as MALDI-TOF was carried out to differentiate the isolated *Bacillus subtilis* from *Bacillus thuringiensis* and *Bacillus amyloliquefaciens*. Phylogenetic tree was constructed to analyze the evolutionary relationship among different genus and species with the newly isolated strain.

Keywords: Hydrogen producer, cellulose degrader, biochemical and molecular characterization, phylogenetic analysis

1. INTRODUCTION

Global sustainable development is focused in our present condition due to environmental degradation and energy crisis. Nearly 80% of energy consumption mostly depends on the fossil fuels such as coal, oil and natural gas, which results in the rapid exhaustion of fossil fuels [1]. In addition to that, the process involved during the conversion of fossil fuels to energy results in global warming and drastic climate change problems in the environment. Energy gap should be filled with the new, alternative energy source around the world. In that way, the first generation biofuels like bioethanol and biodiesel have been progressed by the researchers using food crops such as corn, sugar crops and palm oil as organic substrates. This leads to indirect impact on the rise of food price and thus contributes to global food crisis. Further, the researchers are much focused in developing the second generation biofuels by utilizing the agricultural residues as organic substrates and resulting in the production of biofuels [2].

"Environmental biorefinery" is the optimization of process conditions for the production of wide range of products from biomass through designing the specific installations [2]. The suggestion of hydrogen is more advantageous due to its harmless nature to environment and the mammals [3,4,5,6]. At present, fossil fuels are mainly employed to generate 88% of H_2 production [7]. Hydrogen is the ideal and best alternative to fossil fuels due to its environmental friendly nature and zero green gas emissions. Its oxidative combustion can lead only to the formation of water and it has its higher energy content of 142 KJ g^{-1}[6]. At present, the different chemical methods have been employed for the production of H_2 such as water electrolysis, steam reforming by methane and thermocatalytic reformation of H_2 rich substrates [8,9,10,11,12]. The mostly used method for the synthesis of H_2 are water electrolysis and steam reformation by methane. About 4% of H_2 from the total hydrogen production have been produced by water electrolysis method. All the chemical methods of H_2 production are energy intensive and unsustainable but the operation of the process is simple. Now a-days, the energy source of H_2 is utilized for refining diesel by using H_2 as one of the chemical reactants and also in the synthesis of ammonia. But the commercialization of H_2 has not yet implemented due to problems in storage facilities and transportation methods and highest production costs [5,13,14,15] but the research on optimization is under progress [16]. The main emerging drawback of H_2 is the storage facility due to lesser volumetric density at normal condition is 3.00 k Wh m^{-3} compared to 9.97 kWh m^{-3} for methane [17].

Hence, the alternative best method for the synthesis of H_2 is through biological routes that require less energy, carry out the operations at ambient temperature and pressure and can be highly developed based on performance and design [7,10]. The different routes of biohydrogen production are direct and indirect biophotolysis by green algae and cyanobacteria respectively, photo fermentation using photosynthetic bacteria and dark fermentation using obligate or facultative anaerobic bacteria [8, 10, 11, 18, 19, 20, 21, 22, 23]. In contrast to different biohydrogen production processes, dark fermentative method is a key technology due to its no requirement of light source for operation, higher H_2 evolution rate (mmol L^{-1} h^{-1}) and the complex lignocellulosic wastes could be degraded biologically by cellulose and hemicellulose degrading bacteria resulting in an effective hydrogen production [12, 14, 24]. Major drawback of dark fermentation is the low yield (less mole of H_2 per mole of substrate consumed) which is mainly due to the poor knowledge on the fundamentals of metabolism of each microorganism [25].

Due to abundant availability of low cost effective feedstocks of lignocellulosic biomass such as grasses, water plants, soft woods, hard woods and agricultural residues like bagasse, straw and reeds, it can be effectively utilized for the conversion of biomass to biofuels and other chemicals with environmental, strategic and economic advantages [26]. Cellulose occupies a major component in the lignocellulosic biomass along with the smaller quantities of hemicellulose and lignin. Cellulose is a homopolymer composed of the linkage of 1→4 β D- glucose residues in a native form and these type of cellulose is classified as cellulose I. Each of the polymer chains of cellulose are assembled together as bundles which is called as microfibrils or fibrils [27]. The two phases of cellulose of cellulose I_a and I_b are present in the structure of cell wall of the plants along with amorphous cellulose [28,29,30]. The rare form of cellulose I_a is available only in few green algae [30] but the higher availability of cellulose I_b is present in the cell wall of higher plants and in some of the green algae.

Due to highest proportion of cellulose in the lignocellulosic wastes, it is used as a raw material for the production of pulp, paper and other value added products [31]. Hence for the utilization of complex lignocellulosic waste as organic streams, the microorganism must be competent in degrading cellulose and thus the bacterium must have an inbuilt cellulose degrading systems. The hydrolytic enzymes named as cellulases that perform the degradation of complex cellulose systems in nature. The complex systems of enzymes present in the cellulolytic systems such as endo- 1,4 β- D- glucanase (endoglucanase, EC3.2.1.4), 1,4- β-D-glucan- cellobiohydrolase (exoglucanase, EC 3.2.1.91) and β-D- glucoside glucanohydrolase (cellobiase, β-glucosidase, EC 3.2.1.21) which acts together for the degradation of cellulosic substrate [32, 33. Enzymatic cellulose degrading systems are available only in few of the bacterial species.

In this article, the challenging effective biohydrogen producing bacteria with inbuilt cellulose degrading ability was isolated from the soil and was further characterized. The basic microbiological investigations of the bacterium was carried out to improve the overall strain performance and further it will be helpful in the bioprocess engineering research during the optimization of different operating parameters in biohydrogen productivity.

2. MATERIALS AND METHODS

2.1 Medium components and analytical instruments used

Nutrient broth and other chemicals used for mineral salt medium (MSM) preparation, genomic DNA isolation, MALDI-TOF analysis were procured from Himedia Limited, Mumbai, India. The gas products mainly hydrogen was analyzed using Chemito 7610 gas chromatography, Chemito Instruments private limited, India. UV spectral analysis were carried out using UV/Vis Biospectrophotometer (EliCo private limited, India) to read the absorbance. MALDI-TOF mass spectral analysis was done using Microflex, Bruker Daltonics Inc., USA. The forward and reverse primers for the amplification of template DNA were procured from Eurofins, Bangalore, India. Purification and sequencing of PCR product was carried out using a wizard PCR DNA purification system (Promega, Madison, Wis.,) and ABI PRISM 310 automated sequencer (PERKIN- ELMER, Conn.) as described in the manual for the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit respectively [34].

2.2 Media preparation

The composition of MSM is as follows (g L^{-1}): Na_2HPO_4, 3.6; KH_2PO_4, 1.6; $(NH_4)_2$ SO_4, 1.0; Mg SO_4, 1.0; $CaCl_2$. $2H_2O$, 0.10; Fe (NH_4) citrate, 0.01; agar, 15.0 and trace elements solution, 10 ml. The trace element solution contains the following constituents (mg L^{-1}): $ZnSO_4$. $7H_2O$, 10.0; $CoCl_2$. $6H_2O$, 1.0; $MnCl_2$. $4H_2O$, 3.0; NO_2MoO_4. $2H_2O$, 3.0; $NiCl_2$. $6H_2O$, 2.0; $CuCl_2$. $2H_2O$, 1.0 and H_3BO_3, 3.0 and pH of the medium was adjusted to 7.0 ± 0.05. Nutrient broth contains the following composition in g L^{-1}: beef extract, 1.0; yeast extract, 2.0; peptone, 5.0 and sodium chloride, 5.0 and pH of the medium was adjusted to 7.0 ± 0.2.

2.3 Strain isolation

The strain used in this study was isolated from soil samples collected from sugarcane bagasse storage yard. One gram of soil sample was mixed with 50 ml sterile distilled water. The supernatant was serially diluted upto 10^5 dilution and spread plated on CMC agar (MSM supplemented with 0.6% (w/v) CMC). The plates were incubated under aerobic condition at 37 ℃ for 48 h. The grown colonies were isolated and inoculated separately in fermentation medium supplemented with 0.6% CMC as sole carbon source. The percent hydrogen content present in the biogas was estimated by Gas Chromatography. It was equipped with thermal conductivity detector (TCD) connected with packed column. 20µL of the gas was taken from the headspace with the help of air tight syringe and injected into the column. Nitrogen was used as a carrier gas at a flow rate of 2 ml/min. The operational temperature of the injection port, detector and oven were maintained at 150℃, 200℃ and 80℃ respectively. The highest hydrogen producing strain was selected for further studies.

2.4 Identification of the bacterial strain using 16S rDNA sequencing

Morphological and biochemical characterization of the isolated strain was carried out by following the standard procedures [36]. An isolated strain was inoculated in 1.5 ml Luria broth (LB) medium and grown overnight in a shaking incubator at 37 ℃. The grown culture was centrifuged at 10,000 rpm for 10 minutes. From the pellet, chromosomal DNA was isolated using the following procedure. The procedure includes disruption of cells by cell-lysing solution, RNase treatment, phenol/chloroform extraction and ethanol precipitation [37]. The extracted DNA was run in agarose gel electrophoresis and visualized under UV light. The quantity and the purity of the extracted DNA were estimated by measuring the UV absorbance at 260 and 280 nm spectra.

A large fragment of 16S rRNA gene was amplified by PCR using universal primers as 27F (5"- AGA GTT TGA TCM TGG CTC AG- 3") and 1522R (5"- AAG GAG GTG WTC CAR CC-3"). PCR amplification was carried out using the master mix containing dNTP at a concentration of 200 µM, each of the primers at a concentration of 4 µM, 100 ng of template DNA, and 2.5U DNA polymerase in a total volume of 50 µL. PCR reaction conditions were as follows: (i) 2 min at 95 ˚C, (ii) 30 cycles of 30 s at 95 ˚C, 30 s at 50-55 ˚C, and 5 min at 72 ˚C, (iii) 5 min at 72 ˚C. Further, the PCR product was purified and sequenced. After sequencing, the closest known relatives of the new isolate were determined using the Basic Local Alignment Search Tool (BLAST) algorithm (BLAST tool) available at National Center for Biotechnology Information (NCBI) by performing a nucleotide database search.

2.5 Confirmation of the species using MALDI-TOF analysis

Strain identification was carried out by MALDI-TOF MS analysis. The procedure is as follows: A portion of strain at the exponential growth phase was smeared onto a 96-well target plate. After drying, 1 µl of α- cyano- 4- hydroxy cinnamic acid (CHCA) matrix solution was used to cover the surface of the microorganism. The dried target plate was loaded into the machine, which was equipped with a 337-nm nitrogen laser. In the mass range of 2 to 20 kDa, the spectra were recorded in the linear mode and subsequently analyzed using MALDI Biotyper Automation Control and Biotyper 2.0 software.

2.6 Phylogenetic tree construction

For making evolutionary studies, the sequence homology was obtained by nucleotide- nucleotide BLAST (BLASTn) search tool. Phylogenetic analysis was performed with the Clustal X program [38]. Phylogenetic tree construction and bootstrap analysis were performed using the Mega 7 program. Phylogenetic tree was constructed using neighbour-joining method available in the Mega 7 program [39,40].

3. RESULTS AND DISCUSSION

The soil sample was serially diluted and spread plated on mineral salt medium containing CMC as sole carbon source and further the plates were incubated at 37 ℃. In 10^5 dilution, the four colonies were grown on the specific growth media and each of the colony showed different morphologies. The four colonies can effectively degraded the cellulosic substrate, CMC. After isolation, the hydrogen productivity of the four colonies was analyzed by using glucose as carbon source. Among the different colonies, cellulose degrading single colony with highest hydrogen yield was selected for further studies.

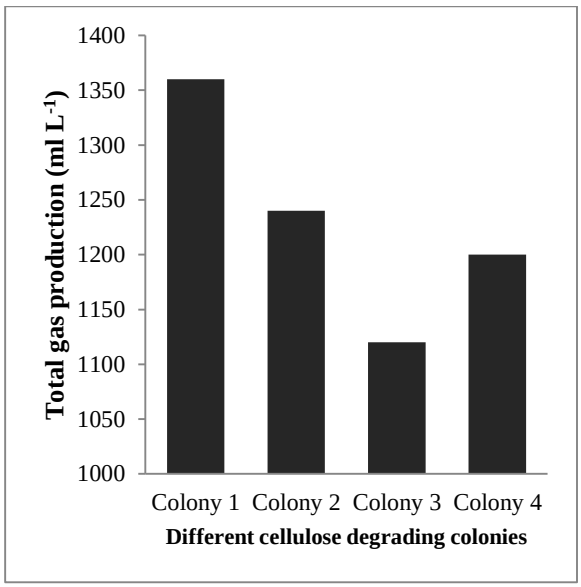


Figure 1 Total gas production of different cellulose degraders and hydrogen producers

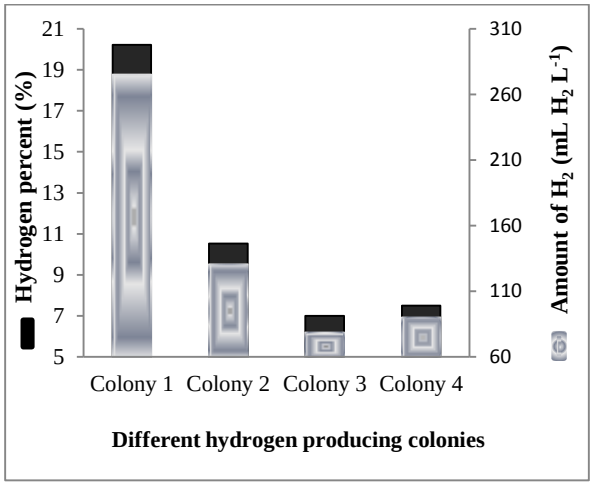


Figure 2 Hydrogen percent in total gas production and hydrogen productivity of the different isolated colonies

The isolated four colonies have the potent to degrade cellulose and capable of producing hydrogen. Among the four colonies, colony 1 showed the highest hydrogen productivity of 275.13 mL H_2 L^{-1}. 20.23% of H_2 was present in total gas produced of 1360 mL L^{-1} (Figure 1 and 2). The highest hydrogen produced strain was further biochemically and morphologically characterized.

3.1 Characterization of the highest hydrogen producing strain

During primary identification, the colony morphology was thoroughly analyzed. Colony forming units (CFU) of the isolated bacterium were appeared as cream coloured, the

edges are circular, flat and it formed an undulate margins. The shape of the colony was appeared as like bacilli when cultured on the nutrient agar plate. In the nutrient broth culture, the isolate was in the form of pellicle with little or no turbidity. The isolated highest hydrogen yielding strain is a Gram positive and rod-shaped bacterium (Fig. 1.3) which was visualized under 1000X total magnification (100X lens × 10X eyepiece magnification) oil immersion light microscope. Differential staining technique such as endospore staining is mainly to differentiate the vegetative cells from endospore formation. After staining, both vegetative cells and endospore cells were observed under 1000X total magnification (100X lens × 10X eyepiece magnification) oil immersion light microscope (Fig. 1.4). The spore formation was occurre at the end of the exponential growth phase or during the time of substrate depletion. From the motility test, it was observed that the isolate is motile.

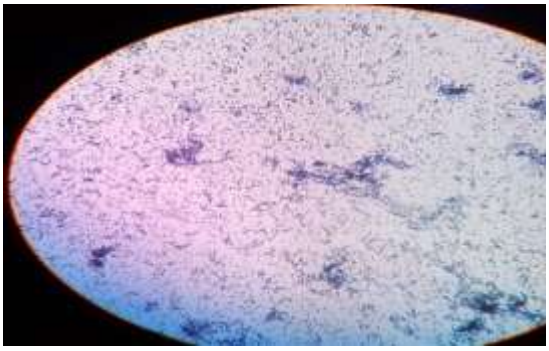


Figure 3 Gram staining of the highest hydrogen producing colony

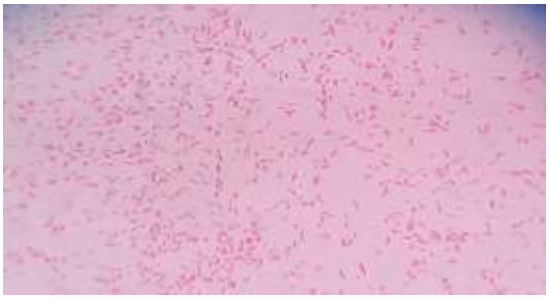


Figure 4 Endospore staining of the highest hydrogen yielding strain

The different secondary biochemical characterization methods such as IMViC (Indole, methyl red (MR), Voges-Proskauer (VP), citrate utilization) tests, sugar utilization, nitrate reduction, starch, urea and gelatin hydrolysis, oxidase and catalase tests were carried out to identify the exact genus and species (Table 1).

The isolate has the ability to produce organic acids by utilizing different sugars such as glucose, sucrose, lactose and mannitol. Nitrate could be reduced by the isolate to nitrite. The isolated microorganism can hydrolyze starch which was identified by the formation of black colour in the agar plate containing minimal medium with starch as sole carbon source. Urease and gelatinase were produced by the microorganism that was confirmed through the hydrolysis of urea and gelatin. The organism showed negative result in the oxidase test. Catalase is the enzyme responsible for the detoxification of hydrogen peroxide into water and oxygen gas (Eq. 1). The isolate showed positive result in the catalase test.

$$2H_2O_2 \xrightarrow{\text{Catalase}} 2H_2O+O_2 \text{ ---------- } (1)$$

Table 1 Biochemical characterization of the highest hydrogen producing isolate

S. No.	Biochemical tests	Results observed
01.	Gram staining	Gram positive
02.	Endospore staining	Spore formation occurs
03.	Motility test	Motile
04.	Indole test	Negative
05.	MR test	Negative
06.	VP test	Positive
07.	Citrate Utilization test	Positive
08.	Starch hydrolysis	Positive
09.	Nitrate reduction test	Positive
10.	Gelatin hydrolysis	Positive
11.	Oxidase test	Negative
12.	Catalase test	Positive
11.	Glucose fermentation	Acid formation occurs
12.	Sucrose fermentation	Acid formation occurs
13.	Lactose fermentation	Acid formation occurs after 48 h of growth
14.	Mannitol fermentation	Acid formation occurs after 48 h

3.2 Growth curve of an isolate, *B. subtilis* AuChE413

The viable cells present at different time intervals were analyzed using cell count and optical density measurements. The growth curve of the bacteria was plotted using incubation time and its corresponding optical density at 600 nm (Figure 5). The cells grew favorably at 37 °C and there is a minimum lag phase in the simple substrate, glucose.

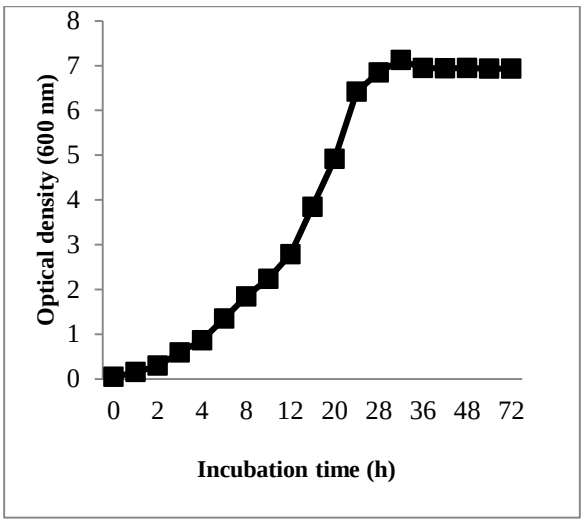


Figure 5 Growth curve of an isolate

The doubling time (or) the generation time of the bacteria is calculated by using Eq. (2) and (3).

$$k = \frac{\log N_t - \log N_0}{0.301\, t} \text{ ---------- (2)}$$

$$g = \frac{1}{k} \text{ ---------- (3)}$$

Based on the growth rate constant (k), the mean generation time or doubling time was calculated [40]. The doubling time of an isolate is 120 min i.e., 120 min/ gen or 0.5 generations h^{-1}. Researchers have also reported that the doubling time of *B. subtilis* is 120 min [41,42]. The microorganism attains the sporulation stage 12 hrs after the end of exponential phase which was observed through phase contrast microscope. During this stage, the spores are mainly formed to compensate the unfavorable conditions created in the medium due to the production of primary and secondary metabolites and the depletion of medium components [43]. The isolate, *B. subtilis* AuChE413 grows well in cellulose and starch medium rather than in glucose medium. It grows well in aerobic as well as in anaerobic environment. Thus, the strain is a facultative anaerobic microorganism.

3.3 Specific growth rate of an isolate at different temperatures and pH

The specific growth rate of the isolate at different temperatures was obtained from the viable cells at different time intervals corresponding to different temperatures. Specific growth rate (h^{-1}) was calculated for each temperature using Eq. (4).

$$log_{10} \frac{N_2}{N_1} = K.\frac{(t_2 - t_1)}{2.303} \text{ ------ (4)}$$

Table 2 Growth parameters of the isolate incubated at different temperatures at pH 7

Temperature (℃)	Specific growth rate (h^{-1})	Regression (R^2)
20	0.160	0.999
30	0.188	0.987
32	0.215	0.989
35	0.297	0.992
37	0.347	0.994
40	0.184	0.991
45	0.182	0.972
50	0.174	0.986

Based on the specific growth rate at different temperatures, the isolate exhibited significant growth upto 40-50 ℃ and the maximum growth falls between 30-35 ℃ [44]. The maximum growth of the bacterium was achieved at an incubation temperature of 35 ℃ (Table 2).

Table 3 Growth parameters of the isolate at different pH conditions with a constant temperature of 37 ℃

pH	Specific growth rate (h^{-1})	Regression analysis (R^2)
4	0.180	0.987
5	0.196	0.996
6	0.204	0.995
6.5	0.297	0.990
7	0.347	0.992
8	0.189	0.986
9	0.147	0.991

Maximum specific growth rate was achieved at pH 7 and hence it is considered as an optimum pH for the growth of the strain (Table 3). Compared to pH 6.5, there was 14% more growth rate was obtained at pH 7. Hence, pH plays an important role in obtaining the maximal growth.

3.4 Molecular characterization of the highest hydrogen yielding strain

To identify the exact species, the isolated highest hydrogen yielding strain was molecularly characterized using 16S rRNA sequencing.

Figure 6 shows the FASTA format of the gene sequence. The obtained gene sequence of the isolate was compared with the known gene sequence of the microorganism found in the National Center for Biotechnology Information (NCBI) database using Basic Local Alignment Search Tool (BLAST). To find the similar sequences, BLAST uses heuristic method in which the short matches between query and known database sequences has been find out. After seeding, local alignment between the sequences was carried out. Finally based on the Expectation or expect value "E" (also identified as E- score), the comparison results were analyzed which signifies the high- scoring segment pair (HSP). Through BLAST analysis, the gene sequence of the isolate showed 99% identity with *Bacillus subtilis* and the novel strain was named as *Bacillus subtilis* AuChE413. The obtained gene sequence was uploaded in NCBI with the reference identification number JX471147.1 .

>gi|406822415|gb|JX471147.1| *Bacillus subtilis* strain AuChE413 16S ribosomal RNA gene, partial sequence
CGGGCGATTGGGGCGGCGGTGGCCCTTATACAT
GCAAAGTCGAGCGGACAGATGGGAGCTTGCTC
CCTGATGTTAGCGGCGGACGGGTGAGTAACACG
TGGGTAACCTGCCTGTAAGACTGGGATAACTCC
GGGAAACCGGGGCTAATACCGGATGGTTGTTTC
AACCGCATGGTTCAAATATAAAAGGTGGCTTCG
GCTACCACTTACAGATGGACCCGCGGCGCATTA
ACTAGTTGGTGAGGTAACGGCTCACCAAGGGA
ACGATGCGTAGCCGACCTGAAAGGGTGATCGG
CCACACTGGGACTGAAACACGGCCCAAACTCCT
ACGGGAGGCAGCAGTAGGGAATCTTCCGCAAT
GGACGAAAGTCTGACGGAACAACGCCGCGTGA
GTGATGAAGGTTTTCGGATCGTAAAACTCTGTT
GTTAGGGAAGAACAAGTACCGTTCGAATAGGG
CGGTACCTTGACGGTACCTAACCAGAAAGCCAC
GGCTAACTACGTGCCAGCAGCCGCGGTAATACG
TAGGTGGCAAGCGTTGTCCGGAATTATTGGGCG
TAAAGGGCTCGCAGGCGGTTTCTTAAGTCTGAT
GTGAAAGCCCCCGGCTCAACCGGGGAAGGTCA
TTGGAAACTGGGGAACTTGAGTGCAGAAGAAG
AGAGTGGAATTCCACGTGTAGCGGTGAAATGCG
TAGAGATGTGGAAGAACACCAGTGGCGAAGCG
ACTCTCTGGTCTGTAACTGACGCTGAAGAGCGA
AAGCGTGGGGGAGCGAACAAGATTAGATACCC
CTGGTAGTCCACGCC

Figure 6 16S rDNA sequencing of the newly isolated bacterium

3.5 Confirming the isolated strain by MALDI-TOF analysis

Ample of literature studies have reported on difficulty in differentiating *B. subtilis* from *B. amyloliquefaciens* due to similarity in 16S rRNA sequences (Wang et al., 2007; Fernandez- No et al., 2013). The complementary method through phylo proteomic analysis to the genetic analysis such as 16S rRNA sequencing is Matrix- assisted laser desorption ionisation- time of flight mass finger printing (MALDI- TOF) analysis. This method can accurately identify and differentiate the *Bacillus* at its species level particularly *B. cereus* and *B. subtilis* from *B. thuringiensis* and *B. amyloliquefaciens* (Fernandez-No et al., 2013). The isolated bacterium was confirmed at its species level using MALDI- TOF analysis. The spectral result of the isolate was obtained and it was shown in Figure 7. By comparing the mass profiling ratio (m/z) and intensity of the peak with the reference database, score value was provided by the Real-Time Classification Software (RTC) installed in Bruker Daltonics MALDI- TOF which confirms the microorganism as *Bacillus subtilis.*

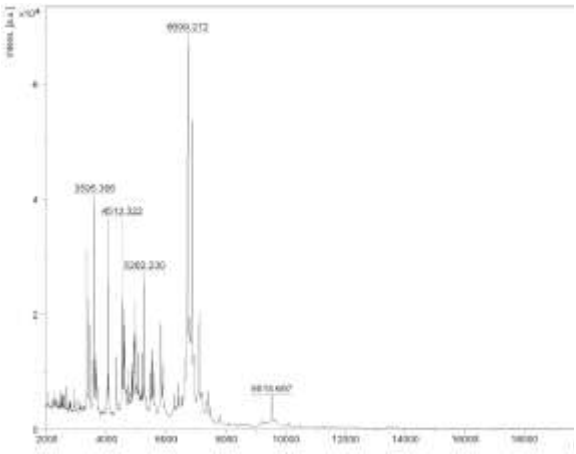


Figure 7 MALDI- TOF analysis to confirm the species level

3.6 Comparison of new isolate with different *Bacillus* strain

The isolate was compared with different other *Bacillus* sp., using 16S rRNA sequencing by multiple sequence alignment.

Table 4 Comparison of 16S rRNA sequencing of different *Bacillus sp.*, strains with *B. subtilis* AuChE413 (new isolate) by Multiple Sequence alignment (MSA)

Microorganism	Strain	Accession No.	% similarity
New isolate, *B. subtilis*	**AuChE413**	**JX471147**	**100**
Brevibacterium halotolerans	DSM_8802	NR042638	97
Bacillus mojavensis	IFO15718	NR024693	97
Bacillus subtilis subsp. spizizenii	NRRL_B-23049	NR024931	97
Bacillus vallismortis	DSM11031	NR024696	97
Bacillus atrophaeus	JCM9070	NR024689	97
Bacillus amyloliquefaciens	NBRC15535	NR041455	96
Bacillus sonorensis	SMPGMb7	JX280500	96
Bacillus sonorensis	NRRL_B-23154	NR025130	95
Bacillus aerius	UAC-16	JX475117	94

The isolated soil bacterium was highly correlated with *B. subtilis* strains and with other *Bacillus sp.*,. It showed 97% similarity with *B. mojavensis, B. vallismortis, B. subtilis subsp. spizzenii, B. atrophaeus* and *Brevibacterium halotolerans.* The organism showed 96% similarity with *B.*

amyloliquefaciens and *B. sonorensis* and 94% similarity with *B. aerius* (Table 4).

3.7 Phylogenetic analysis

The phylogenetic analysis of the isolated strain was compared with different genus such as different *Bacillus sp., Clostridium sp., E. coli, Pantoea sp.,* and *Rhodoplanes sp.,* and is represented in Figure 8. Initially, the 16S rRNA sequence of all the strains were collected from the database NCBI. The multiple sequence alignment was carried out using CLUSTAL W and the phlyogenetic tree was constructed using maximum likelihood method. The isolated strain was closely related with some of the *Bacillus sp.,* such as *Bacillus sonorensis, aerius and licheniformis.* The isolated strain was much different from few of the microorganisms such as *Clostridium, Rhodoplanes sereus, Pantoea agglomerans and E. coli.*

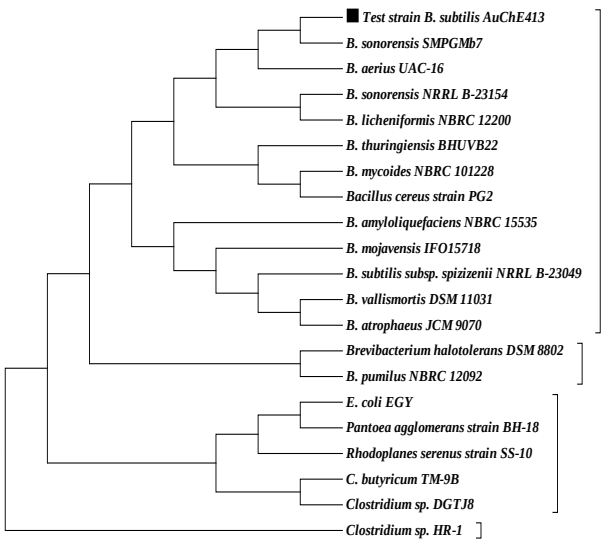


Figure 8 Phylogenetic relationship of an isolate with other *Bacillus sp.,*

CONCLUSION

A competent strain with the ability of hydrogen producing and cellulose degrading was the major target during isolation of bacterium from the soil. Among the different colonies, a single colony with an ability of degrading the complex substrate, cellulose obtained the highest hydrogen productivity of 275.13 ml H_2 L^{-1}. The isolated bacterium is a Gram positive, *Bacillus sp.,* which was inferred from the morphological and biochemical characterization. Further, the bacterium was molecularly characterized using 16 S rDNA analysis. The obtained sequence was compared with the known gene sequence available in the NCBI database using BLAST tool and named as *Bacillus subtilis* AuChE413. Further, the bacterium was confirmed using proteomic analysis by MALDI-TOF. Phylogenetic tree was constructed to compare the evolutionary analysis between the species and different genus.

Conflict of Interest

The authors did not declare any conflict of interest.

ACKNOWLEDGEMENTS

The authors would like to thank Annamalai University, India for providing the lab facilities.

REFERENCES

[1]. Ni, M., Leung, D. Y. C., Leung, M. K. H., and Sumathy, K. 2006. An overview of hydrogen production from biomass. Fuel Process Technol. 87(5):461-72.

[2]. Angelidaki, I., and Kongjan, P. 2007. Biorefinery for sustainable biofuel production from energy crops; conversion of lignocellulose to bioethanol, biohydrogen and biomethane. In: 11th IWA world congress on anaerobic digestion, Brisbane, Australia.

[3]. Zajic, J. E., Margaritis, A., and Brosseau, J. D. 1979. Microbial hydrogen production from replenishable resources. Int J Hydrogen Energy 4:385–402.

[4]. Bicelli, P. L. 1986. Hydrogen: a clean energy source. Int. J. Hydrogen Energy 11:555–562.

[5]. Bockris, J. O. M. 2002. The origin of ideas on a hydrogen economy and its solution to the decay of the environment. Int. J. Hydrogen Energy 27:731–740.

[6]. Piera, M. 2006. Safety issues of nuclear production of hydrogen. Energy Convers. Manag. 47(17):2732-9.

[7]. Nath. K., and Das, D. 2003. Hydrogen from biomass. Curr. Sci. 85(3): 265-71.

[8]. Nandi, R., and Sengupta, S. 1998. Microbial production of hydrogen: an overview. Crit. Rev. Microbiol. 24:61–84.

[9]. Momirlan, M., and Veziroglu, T. 1999. Recent directions of world hydrogen production. Renew. Sust. Energy Rev. 3:219–231.

[10]. Das, D., and Veziroglu, T. N. 2000. Hydrogen production by biological processes: a survey of literature. Int. J. Hydrogen Energy 26:13–28.

[11]. Lopes Pinto, F. A., Troshina, O., and Lindblad, P. 2002. A brief look at three decades of research on cyanobacterial hydrogen evolution. Int. J. Hydrogen Energy 27:1209–1215.

[12]. Hallenbeck, P. C., and Benemann, J. R. 2002. Biological hydrogen production; fundamentals and limiting processes. Int. J. Hydrogen Energy 27:1185–1193.

[13]. Cherry, R. S. 2003. A hydrogen utopia? Int. J. Hydrogen Energy 29:125–129.

[14]. Levin, D. B., Pitt, L., Love, M. 2003. Biohydrogen production: prospects and limitations to practical application. Int. J. Hydrogen Energy 29:173–185.

[15]. Das, D., Khanna, N., and Veziroglu, T. N. 2008. Recent developments in biological hydrogen production processes. Chem. Ind. Chem. Engg. 14:57–67.

[16]. Boddien, A., Mellmann, D., Gaertner, F., Jackstell, R., Junge, H., Dyson, P. J., Laurenczy, G., Ludwig, R., and Beller, M. 2011. Efficient Dehydrogenation of Formic Acid Using an Iron Catalyst. Science 333:1733–1736.

[17]. Wasserstoff Daten - Hydrogen Data. http://www.h2data.de/.

[18]. Benemann, J. R. 2000. Hydrogen production by microalgae. J. Appl. Phycol. 12:291–300.

[19]. Madamwar, D., Garg, N., and Shah, V. 2001. Cyanobacterial hydrogen production. World J. Microbiol. Biotechnol. 16:757–767.

[20]. Melis, A., and Melnicki, M. R. 2006. Integrated biological hydrogen production. Int. J. Hydrogen Energy 31:1563–1573.

[21]. Melis, A. 2002. Green alga hydrogen production: progress, challenges and prospects. Int. J. Hydrogen Energy 27:1217–1228.

[22]. Schuetz, K., Happe, T., Troshina, O., Lindblad, P., Leitao, E., Oliveira, P., and Tamagnini, P. 2004. Cyanobacterial H_2 production - a comparative analysis. Planta 218:350–359.

[23]. Lee, D. J., Show, K. Y., and Su, A. 2011. Dark fermentation on biohydrogen production: Pure culture. Bioresource Technol. 102:8393–8402.

[24]. Hallenbeck, P. C. 2005. Fundamentals of the fermentative production of hydrogen. Water Sci. Technol. 52:21–29.

[25]. Thauer, R. K., Jungermann, K., and Decker, K. 1977. Energy conservation in chemotrophic anaerobic bacteria. Bacteriol. Rev. 41:100–180.

[26]. Adsul, M. G., Ghule, J. E., Singh, R., Shaikh, H., Bastawde, K. B., Gokhale, D. V., and Varma, A. J. 2004. Polysaccharides from bagasse: Applications in cellulose and xylanase production. Carbohydrate Polymers 57:67–72.

[27]. Zykwinska, A. W., Ralet, M. C. J., Garnier, C. D., and Thibault, J. F. J. 2005. Evidence for in vitro binding of pectin side chains to cellulose. Plant Physiology 139:397–407.

[28]. Zykwinska, A. W., Ralet, M. C. J., Garnier, C. D., and Thibault, J. F. J. 2005. Evidence for in vitro binding of pectin side chains to cellulose. Plant Physiology 139:397–407.

[29]. Atalla, R. H., and Van der Hart, D. L. 1984. Native cellulose-a composite of 2 distinct crystalline forms. Science 223:283–285.

[30]. Imai, T., and Sugiyama, J. 1998. Nanodomains of I-alpha and I-beta cellulose in algal microfibrils. Macromolecules 31:6275–6279.

[31]. Mansfield, S. D., and Meder, R. 2003. Cellulose hydrolysis: The role of monocomponent cellulases in crystalline cellulose degradation. Cellulose 10:159–169.

[32]. Singhania, R. R., Sukumaran, R. K., Patel, A. K., Larroche, C., and Pandey, A. 2010. Advancement and comparative profiles in the production technologies using solid-state and submerged fermentation for microbial cellulases. Enzyme Microbiol. Technol. 46:541-549.

[33]. Chandra, M., Kalra, A., Sangwan, N. S., Gaurav, S. S., and Darokar, M. P. 2009. Development of a mutant of *Trichoderma citrinoviride* for enhanced production of cellulases. Bioresour. Technol. 100:1659-1662.

[34]. Weisburg, W. G., Barns, S. M., Peltier, D. A., and Lane, D. J. 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. 173(2):697-703.

[35]. Bergey, D. H. 1985. Bergey's Manual of Systematic Bacteriology. Int. J. Syst. Bacteriol. First edition, 408.

[36]. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY.

[37]. Tompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.

[38]. Kumar, S., Tamura, K., and Nei, M. 2004. MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings Bioinf. 5:150–163.

[39]. MEGA program tool: http://www.megasoftware.net/

[40]. Prescott, L. M., Harley, J. P., and Klein, D. A. 2002. Microbiology, 5th edition, Chapter 7: Microbial growth, The McGraw- Hill publishers, ISBN: 0-07-282905-2.

[41]. Burdett, D. J., Kirkwood, T. B. L., and Whalley, J. B. 1986. Growth Kinetics of Individual *Bacillus subtilis* Cells and Correlation with Nucleoid Extension. Journal of Bacteriology 167(1):219-230.

[42]. Leitch, J., and Collier, P. J. 1996. A new chemically-defined medium for *Bacillus subtilis* (168) NCIMB 12900. Letters in applied microbiology 22(1):18-20.

[43]. Warriner, K., Waites, W. M. 1999. Enhanced sporulation in *Bacillus subtilis* grown on medium containing glucose: ribose. Letters in applied microbiology 29: 97-102.

[44]. Korsten, L., and Cook, N. 1996. Optimizing Culturing Conditions for *Bacillus subtilis*, South African Avocado Growers' Association Yearbook 19: 54-58.

International Journal of Science and Engineering Applications
Volume 5 Issue 1, 2016, ISSN-2319-7560 (Online)

An IT Approach to Improve the Compilation of Clinical Access Indicators and Disease Prevalence Determination for District Health Administrators

The Case of East Mamprusi District Directorate of Ghana Health Service

Zakaria Abukari
School of Applied Science
Dept. of Computer Science
Tamale Polytechnic
Tamale, Ghana

Abeo Timothy Apasiba
School of Applied Science
Dept. of Computer Science
Tamale Polytechnic
Tamale, Ghana

George Abraham
District Director
Ghana Health Services
Tatale-Sanguli District
Tatale, Ghana

ABSTRACT: The proportion of persons visiting a health facility reflects the level of access of that centre to its catchment area in terms of Out-Patient Department (OPD) per capita. These attendances come with diagnoses which give an indication of the diseases pattern and prevalence within the catchment area as well as patients who are insured and not insured. Though data of this nature are undisputedly crucial to public health processes, morbidity returns from most health facilities, particularly public health facilities are generated manually making it cumbersome and stressful. In addition, this method is error-prone and as such poses a strong threat to disease prevention, control and information management. This research therefore uses an Information Technology approach to improve the process achieving over 90% time gain. The Gambaga Health Centre in the East Mamprusi District was selected for the simulation.

Keywords: OPD; Morbidity; Database; Clinic; Health Information system; Health Metrics Network

1. INTRODUCTION

Health care is an information intensive industry in which reliable and timely information is a critical resource for the planning and monitoring of service provision at all levels of analysis be it organizational, regional, national and international (Paolo, Nicola, Luca, & Mariano). As a consequence, *Information and Communication Technologies* (ICTs) have become enabler in the health care sector. Healthcare institutions are not excluded in the use of advance technologies derived from the proliferation of the silicon chip for record keeping. The records kept include patients' attendances, diagnoses, treatments, health insurance etc. These records are used by the Public Health Units for planning, controlling and preventing diseases.

The Ghana Health Service together with its partners and other health agencies are responsible for the provision of health care delivery to the people of Ghana. This involves the planning, implementation, monitoring and performance assessment of health programmes and services (Ghana Health Service, 2010). In order to fulfil this role, there is the need for accurate, relevant and timely information on the health status and health services in Ghana (Ghana Health Service, 2010). Well-kept patient data also assists healthcare managers in determining how much resources are needed to manage such occurrences.

Over the past decades, information systems have played crucial roles in the development process of most of the advance World (Erickson, 2002). According to (Gregor & Jones, 2006) information systems are said to be the wheels on which steady and sustainable development strives. The United States of America, Germany, China and most of the countries with modern health care systems are advancing daily because of current technological development (Gregor & Jones, 2006).

The health sector in Ghana is facing the challenges of adequately sourcing and keeping information on what pertains in the sector just because the systems that are supposed to be put in place for such information retention are in small scale if not lacking. This situation is even worse at the rural-urban health centres. Government and policy makers have emphasized the importance of computer-base information systems (CBIS) noting that achieving universal access to quality health care is highly dependent on modern health information systems (Andries & Bob, 2008).

The National Health Insurance since its introduction has led to increases in utilization of OPD services across all the regions. With these increases, there has not been equally significant improvement in infrastructure of most of the facilities to accommodate this phenomenon (Ministry of Health, 2013).

According to a Health Metrics Network (HMN) assessment report of Ghana's health information system (DHIMS), there is an urgent need for major improvements in the way the nation's health system collects and uses data (Ministry of Health, 2006).

The use of modern systems of communication, data storage and retrieval at all levels or facet of society especially the Ghana Health Services through the use of computers is being emphasized (Hall, 2003)

It is against these backgrounds that this research realization became preponderant to provide a rudimentary solution to accurate data compilations at the OPDs of the source health facilities. The wise men say that "a thousand-mile journey begins with the first step".

The study is limited to the East Mamprusi District in the Northern Region of Ghana. The scope further focused on the Out-Patient Department, Public Health Unit and as well as a random sampling and interviewing of some clinical and records staff of the Gambaga Health Centre and others from the East Mamprusi District Health Administration

2. STATEMENT OF THE PROBLEM

At the end of every month, the patient attendance books are taken from the consulting rooms for manual generation of morbidity and OPD returns reports by the Public Health Unit of the Health facilities. These reports are then forwarded to the District and Regional Health Directorates for actions. Normally, diseases and persons are grouped by sex, age, NHIS (National Health Insurance Scheme) insured and not insured. Whilst the morbidity report gives the breakdown of the disease, taking into consideration that some clients may have co-morbidity, the OPD statement gives a breakdown of persons who attend the facility within the reporting period. The compilation of the OPD and the morbidity reports are currently done manually which never yields the desired results for the Ghana Health Service (GHS). Major problems of the manual compilations included:

i. Time consuming exercise
ii. During the period of compilation, the consulting rooms' registers are not available for further entries i.e. if the compilation takes a week, then consultation entries for that week will not be recorded and thus will not reflect in that months reports.
iii. In some cases the OPD is more than the morbidity which is logically impossible since the OPD report counts individuals whilst the morbidity counts diagnoses presented by each patient.

In the light of these challenges and considering the crucial importance of the morbidity and OPD returns to the Public Health Processes in the GHS, this research comes in to solve the problem using a state of the art Information Technology tools.

3. OBJECTIVES OF THE STUDY

This work seeks to examine the feasibility and implementation of Health Information Systems technology in the compilation of the clinical access indicators contained in the OPD and morbidity reports which are monthly generated in all public health centres in the country. The specific objectives include:

i. Conduct an in-depth research exploration on Health Information Systems in Ghana with emphasis on OPD and Morbidity reports compilation in the public health sector
ii. Develop a working software solution of the OPD and morbidity reports compilation for implementation in the East Mamprusi District.
iii. Showcase the robustness of ICTs for effective information processing, storage and retrieval in the Public Health Units of the GHS.

4. SIGNIFICANCE OF THE STUDY

The significance and relevance of this research to the Public Health Units of the GHS are numerous. These include but not limited to:

i. Time spent in generating the OPD and morbidity reports will be reduced to a few button clicks giving the health workers more time to attend to other issues.
ii. The generated reports will be accurate, reliable and timely for district and regional health directorates.
iii. The Health Administrators in the East Mamprusi District for the first time can lay hands on timely and accurate morbidity reports leading to better health care planning, improved diagnosis, disease control, prevention and eradication in its catchment area.
iv. Pave way for the creation and adoption of IT enabled services in other functional units in the public health centres.
v. The research report shall constitute an academic reference to other researchers in the subject matter.

5. METHODOLOGY

Two principal research techniques were adopted. Initially, a variety of formal and informal participatory methods were used to obtain the background information needed to constructively carryout the study. The informal methods involved observation, the use of key informants and in-depth interviews of key stakeholders whilst the formal methods included the use of Public Health Nurses to manually compile OPD and morbidity reports for thirty(30) clients on sampled attendances for quantitative and qualitative analyses of the compilation processes. The next technique was an extensive literature reviews on HIS and its initiatives in the GHS to have a comprehensive view of the problem and its solution to be piloted at the Gambaga Health Centre.

6. LITERATURE REVIEW

Various literatures on the thematic areas of the research have been reviewed. Clinical Access Indicators and their compilation in the GHS were first looked at. The researched targeted a problem base solution, and for that reason, technologies and methodologies for developing functional health information systems such as database management systems, systems analyses and design with emphasis on the traditional software development life cycle and the choice of either desktop application or web based software were rigorously studied as the major theoretical underpinning. In order to pilot the solution, the review culminated in the analyses of the ICT status in the Gambaga Health Centre to help focus on implementable information system based on available resources.

6.1 Clinical Access in Ghana

There are many factors that influence access to clinical and health services in Ghana. Access to clinical services is defined as living within thirty (30) minutes to any kind of modern health facility. It is also stated that access to health care services indicates the degree to which individuals are inhibited in their ability to gain entry to and receive care and services from the health care system (Ministry of Health, 1996). Geographical, financial, level of education, availability of health facilities, health professionals, transportation etc. are all factors which influence access to health care. Geographical access indicates the spatial interaction of people and health facilities. Financial access refers to the ability of the people to demand and pay for health services when the need arises. The level of an individual's education affects his/her access to health services. Most illiterates have beliefs and myths that make them belief that modern health care poses problems to them hence they do not access the facilities (Ministry of Health, 2006). A Core Welfare Indicator survey conducted by the Ghana Statistical Services (GSS) in 2003 indicated that between 1997 and 2003, the percentage of people stating that they had access to a health facility increased from 37% to 58% (GSS, 2003).

The 2004 Ghana Demographic and Health Survey (GDHS) gave the following health status indicators: infant mortality rate: 68 per 1000 live births; under 5 mortality rate: 112 per 1000 live births; maternal mortality: 540 per 100,000 live births and life expectancy: 57.2 years. Malaria remains the most frequently reported cause of morbidity and a major cause of childhood mortality. Other frequently reported diseases are diarrhoea, acute respiratory infections, skin diseases, pregnancy related complications, anaemia and malnutrition (GSS, 2005).
The most common chronic diseases are hypertension and diabetes (GSS, 2005)

6.2 Out-Patient Department (OPD) Per Capita

The proportion of persons visiting a health facility reflects the level of access of that centre to its catchment area in terms of OPD per capita. Normally, persons visiting health facilities are grouped by diseases, gender, and age and as well as patients who are insured and not insured by NHIS. These indicators are all determined from the collected data in the OPDs of the health facilities (clinics) to estimate the level of access of that facility to its clients.
The National Health Insurance since its introduction has led to increase in utilization of OPD services across all the regions (figure 1). The number of outpatients per capita continued increasing and in 2012, the relative increase was 11%. OPD per capita reached 1.17, more than doubling 2006 figure (Figure 2). Over 60% of those attending outpatient were females and overall women between the ages of 20 and 34 years were seen more often at the OPD than any other age group for males or females (Ministry of Health, 2013) .

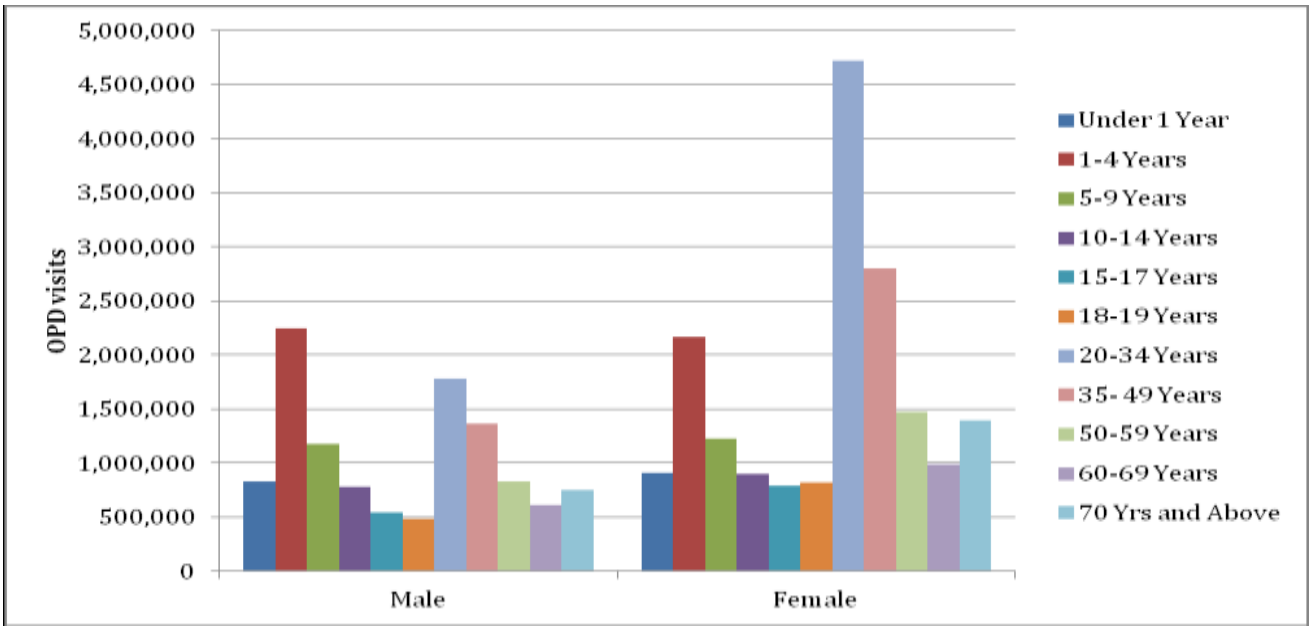


Figure 1 OPD visits by gender and age-group 2012, source GHS

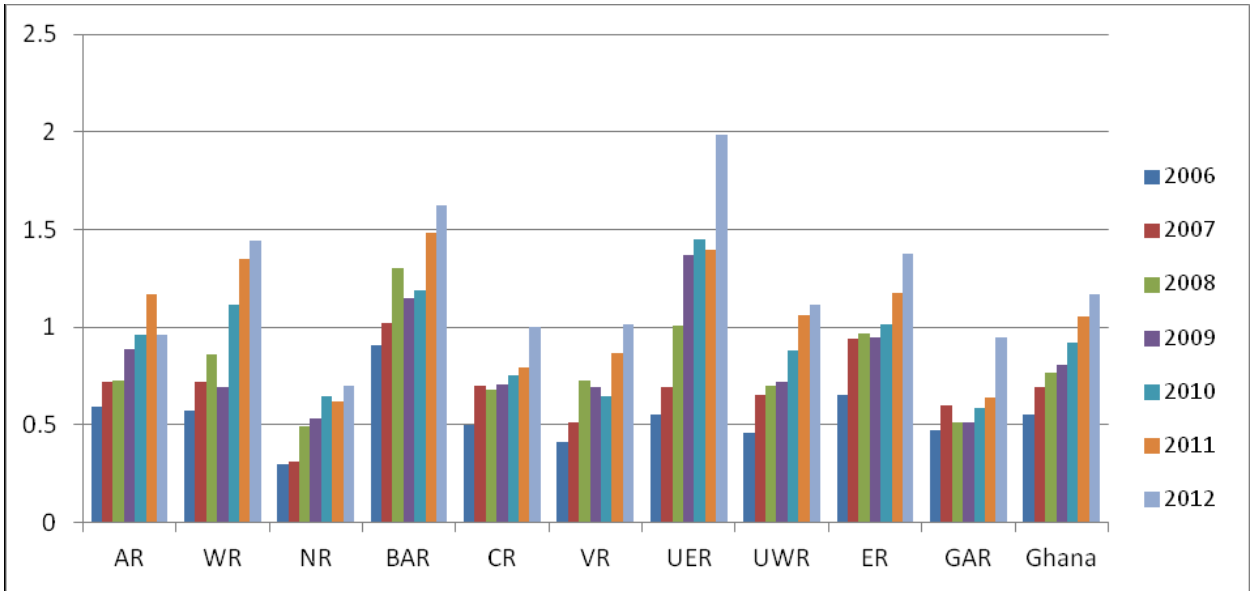


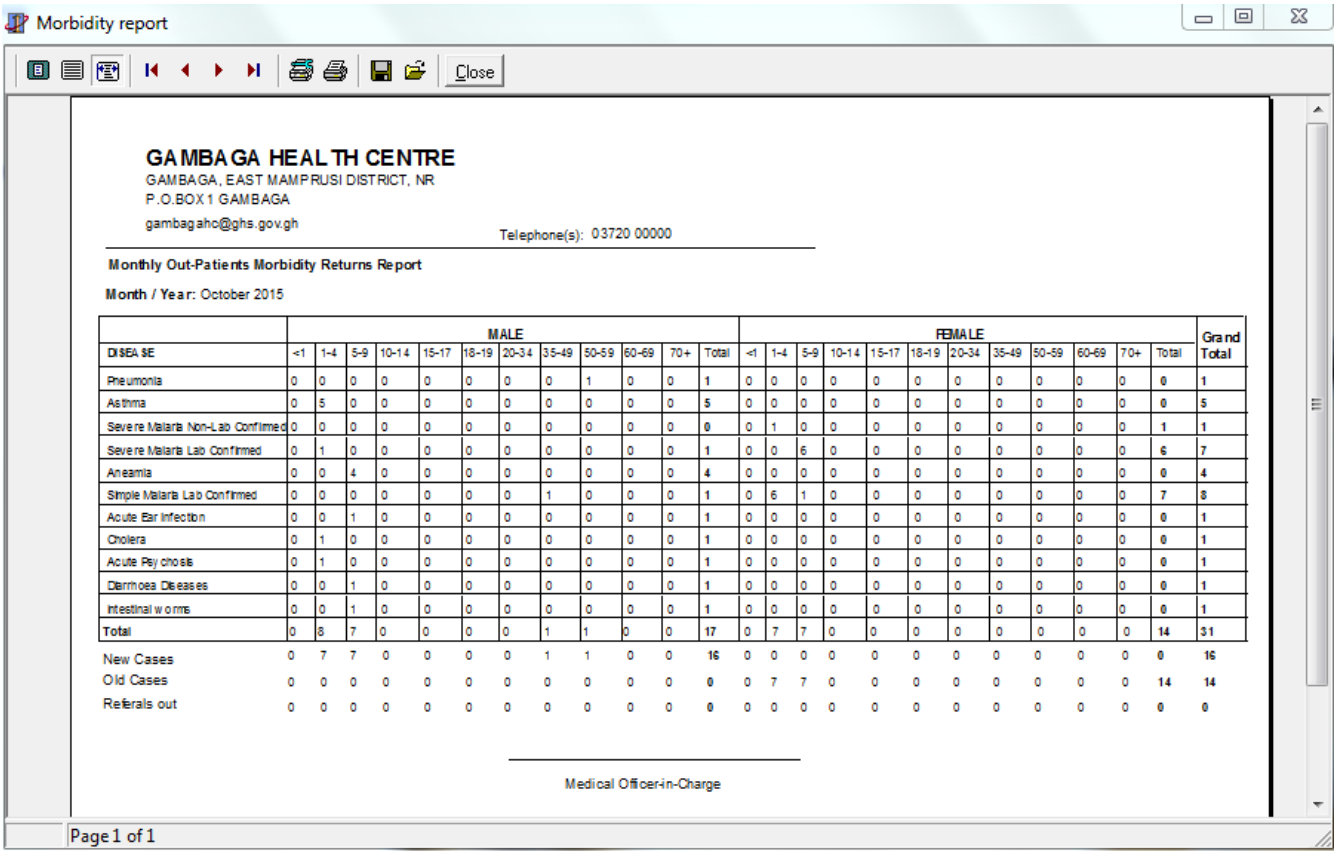
Figure 2: OPD per capita by region, 2006-2012, Source GHS

6.3 Compilation of Clinical Access Indicators in Ghana Health Services

At the end of each month, the consulting room registers are sent to the Public Health Units to compile the OPD and Morbidity reports. Each patient (entry) in the register is classified and counted according to the date, diseases, gender, age group, NHIS Insurance, case type (new or old), referral status. Malaria cases of children below five and above five are specially counted. Figure 3 shows a few entries in a consulting room register whilst figure 4 shows sample morbidity report.

Date	Sex	Age	Case type	Reattendance	Referal	NHIS Status	Diagnosis1	Diagnosis2
15-10-15	F	2	O	Y	N	Y	Simple Malaria Lab Confirmed	
15-10-15	M	2	N	N	N	Y	Asthma	
15-10-15	F	2	O	Y	N	Y	Severe Malaria Non-Lab Confirmed	
15-10-15	F	8	O	Y	N	N	Severe Malaria Lab Confirmed	
15-10-15	M	6	N	Y	N	Y	Aneamia	
15-10-15	M	39	N	N	n	y	Simple Malaria Lab Confirmed	
15-10-15	F	8	O	Y	N	N	Severe Malaria Lab Confirmed	
15-10-15	F	2	O	Y	N	Y	Simple Malaria Lab Confirmed	
15-10-15	M	2	N	Y	N	Y	Severe Malaria Lab Confirmed	
15-10-15	M	6	N	Y	N	Y	Acute Ear infection	
15-10-15	F	8	O	Y	N	N	Severe Malaria Lab Confirmed	
15-10-15	F	2	O	Y	N	Y	Simple Malaria Lab Confirmed	
15-10-15	M	2	N	N	N	Y	Cholera	Acute Psychosis
15-10-15	M	50	N	N	N	Y	Pneumonia	
15-10-15	F	8	O	Y	N	N	Severe Malaria Lab Confirmed	
15-10-15	M	6	N	Y	N	Y	Intestinal worms	
15-10-15	M	2	N	N	N	Y	Asthma	
15-10-15	M	6	N	Y	N	Y	Aneamia	
15-10-15	F	8	O	Y	N	N	Severe Malaria Lab Confirmed	
15-10-15	F	2	O	Y	N	Y	Simple Malaria Lab Confirmed	
15-10-15	M	2	N	N	N	Y	Asthma	
15-10-15	M	6	N	Y	N	Y	Aneamia	
15-10-15	F	8	O	Y	N	N	Severe Malaria Lab Confirmed	
15-10-15	F	2	O	Y	N	Y	Simple Malaria Lab Confirmed	
15-10-15	M	2	N	N	N	Y	Asthma	
15-10-15	M	6	N	Y	N	Y	Aneamia	
15-10-15	F	8	O	Y	N	N	Simple Malaria Lab Confirmed	
15-10-15	F	2	O	Y	N	Y	Simple Malaria Lab Confirmed	
15-10-15	M	2	N	N	N	Y	Asthma	
15-10-15	M	6	N	Y	N	Y	Diarrhoea Diseases	

Figure 3 consulting room register sample



Morbidity report

Close

GAMBAGA HEALTH CENTRE
GAMBAGA, EAST MAMPRUSI DISTRICT, NR
P.O.BOX 1 GAMBAGA
gambagahc@ghs.gov.gh
Telephone(s): 03720 00000

Monthly Out-Patients Morbidity Returns Report

Month / Year: October 2015

	MALE												FEMALE												Grand Total
DISEASE	<1	1-4	5-9	10-14	15-17	18-19	20-34	35-49	50-59	60-69	70+	Total	<1	1-4	5-9	10-14	15-17	18-19	20-34	35-49	50-59	60-69	70+	Total	
Pneumonia	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Asthma	0	5	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	5
Severe Malaria Non-Lab Confirmed	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1
Severe Malaria Lab Confirmed	0	1	0	0	0	0	0	0	0	0	0	1	0	0	6	0	0	0	0	0	0	0	0	6	7
Anaemia	0	0	4	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	4
Simple Malaria Lab Confirmed	0	0	0	0	0	0	0	1	0	0	0	1	0	6	1	0	0	0	0	0	0	0	0	7	8
Acute Ear Infection	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Cholera	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Acute Psychosis	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Diarrhoea Diseases	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Intestinal worms	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Total	0	8	7	0	0	0	0	1	1	0	0	17	0	7	7	0	0	0	0	0	0	0	0	14	31
New Cases	0	7	7	0	0	0	0	1	1	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	16
Old Cases	0	0	0	0	0	0	0	0	0	0	0	0	0	7	7	0	0	0	0	0	0	0	0	14	14
Referals out	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Medical Officer-in-Charge

Page 1 of 1

Figure 4 A sample morbidity reports

6.4 Health Information Systems (HIS)

Health Information Systems constitute the switch from conventional recordkeeping to computer-controlled systems in medical offices. In the Western countries, patient record keeping and management which were traditionally handled by medical assistants and health administrators resulting concentrated efforts on office procedures that generated paper records of patient data, payments, insurance claims, statistics etc. are now automated. Although paper-based systems aren't likely to disappear, many practices now have affordable Health Information Systems for scheduling, recording patient information, billing, filing insurance claims etc. (Sanderson, 2009).

Similar to other industries, the nature of healthcare industry has changed over time from a relatively stable industry to a dynamic one leading to health information systems evolving through several different technologies. The aim of Health Information Systems is to contribute to a high-quality and efficient patient care, disease control and prevention.

The current status of HIS varies among countries. There are 193 member countries of the World Health Organization (WHO). In 2009; 114 of them participated in the global survey on e-health (Muhammadi & Mohd, 2011). Most developed countries have fully utilized HIS in their systems as they have the resources, expertise, and capital to implement them. However, in the developing countries, HIS have not been fully utilized yet according to WHO (2012). HIS are complex applications which have demonstrated benefits. Their complexity makes it imperative to have good application design, training, and implementation. Studies have evaluated Health Information Systems and reported on various benefits and limitations of these systems including increases in immunization rates, improved data collection, increased staff productivity, increased visitor satisfaction with services, improved communication, quality of care, access to data, reduced medical errors, and more efficient use of staff time (WHO, 2012).

6.4.1 *Health Information Systems in GHS*

Since 2005 Ghana has effectively implemented the National Health Insurance Scheme (NHIS) in all Districts/Municipalities/Metropolitans of the country. The NHIS strives to cover the entire population with equal universal access to the basic health services package. A key part of these health sector reforms is the need to develop better health information systems (Vital Wave Consulting, 2009).

In parallel with the NHIS implementation, the Ministry of Health (MOH) initiated various programs to strengthen its inadequate health information system (HIS). As a responds, the MOH's Ghana Health Service (GHS) launched its automated system, called **DHIMS**. It attained full country coverage in 2007 after

original piloting in 20 districts across the country (Vital Wave Consulting, 2009). Data flows are vertical, with accumulated data from health facilities in districts flowing to the regional and national levels. The GHS has expressed plans to extend it to the sub-districts where untapped capacity for data entry is present.
DHIMS has faced numerous challenges, such as Inconsistent information requirements, lack of qualified information officers in many district hospitals, inadequate funding for deployment and poor equipment availability at the district level (Vital Wave Consulting, 2009)

7. DESIGN OF THE SYSTEM

It is often said that "if you fail to plan, you plan to fail" (GoodReads, 2011) . Systems analyses and design are critical planning stages for effective information systems development and implementation. Advance software technologies are available to facilitate the development of health information system which resolves the data management challenges in data-centred organizations such as the hospitals. The traditional Software Development Life Cycle (SDLC) has been identified in the literature as the best practice strategy to develop both large and small scale data centred applications. At each step details of the activities carried out are captured.

Following the interviews conducted, it was well established that the problem was understood to be a data management challenge due to large volumes of data and the time-consuming processing requirements. The primary source of data for diseases surveillance, control and prevention in the GHS are the OPDs of hospitals, clinics and CHPS compounds. The details of patients' data captured and the processed reports are sampled in Figure 3 and 4.

The potential users were variously interviewed to ascertain the access levels, inputs, processing requirement and output reporting. We debriefed the Management of the Gambaga Health Centre on the project document which addressed the minimum requirements to create a solution to such a long standing problem. Management commitment, which is vital to systems development success and implementation, was guaranteed and as such the next steps in the SDLC were followed.

Funding and available ICT resources at the pilot site suggested we drew a conclusion on developing a client/server desktop application. A centralized database to store the OPD data was created in a relational database management system (figure 8)

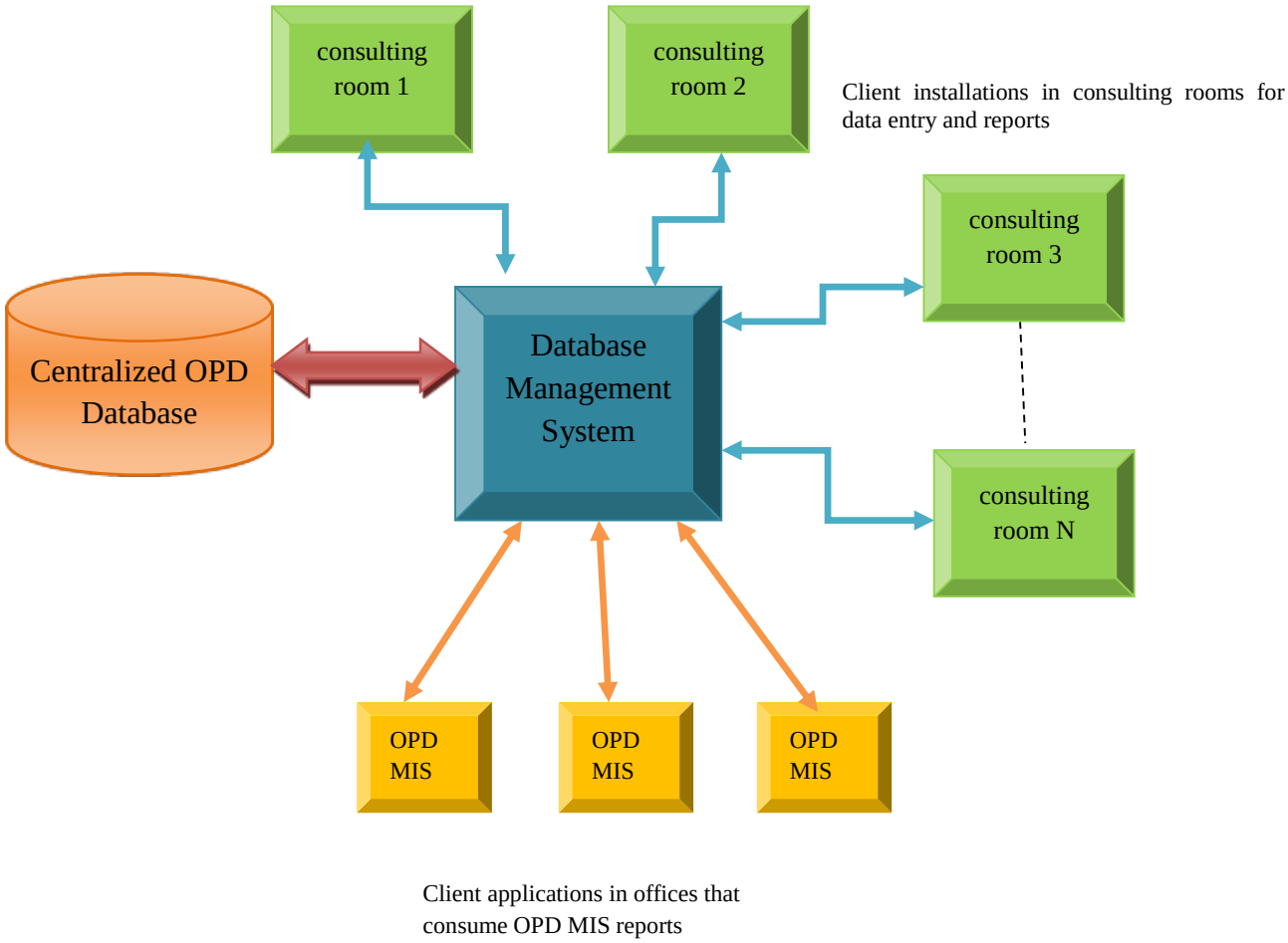


Figure 5 Architecture of the solution

The client side of the application is connected to the data server in a secured manner for data entry, processing and generation of the appropriate MIS reports (Figure 5). All users go through a login process to access the Information System. This architecture allows centralized but distributed database through the provision of a network environment. The complete system can equally be deployed on a standalone PC which will act as the server and the client at the same time.

The front and the back ends of the solution are distinctions which separate the user interface layer (presentation layer) and a data access layer respectively. The front-end code was written in Borland Delphi, an object oriented language which has a rich IDE for Rapid Application Developments (RAD). Code for creating graphical components are encapsulated and embedded in visual component libraries that make graphical and event driven applications easy to write.

Using the Data Manipulation Language (DML), the Transact-SQL of the Microsoft SQL Server 2008, back-end interactions are done with the underlying data of the information system.

8. SUMMARY OF FINDINGS

Research methodologies vary from problem to problem. We have particularly used experimental research design in this work to manipulate some of the independent variables in order to observe the final results. A cross-sectional slice of public health nurses were experimented on manual compilation of the OPD and Morbidity reports of sampled attendance from the Gamabaga Health Centre for quantitative and qualitative analyses. Variables measured were limited to: Effect of years of work experience of a public health nurse on clinical access indicators compilation; Accuracy of the manual compilation; and effect of manual compilation on GHS deliverables.

8.1 Effect of Years of Work Experience of a Public Health Nurse on Clinical Access Indicators Compilation

Fifteen public health nurses with varying years experience from one to twenty worked on thirty patients records to measure these variables (Table 1 and 2). Some exclusive findings made included:

i. During the experimentation, it was realized that practitioners with five or more years in the job performed far better than the others (figure 9).
ii. Processing just thirty (30) patients, the average time used by the novice (one to four years experience) was about two hours (120 minutes) whilst the experience nurses used approximately one hour twenty three minutes (83 minutes). It implies that a facility attending to a minimum of 150 patients a day requires seven and half hours of an experienced public health nurse to come out with the morbidity reports for the day. Meanwhile the 2012 annual report of the GHS reveals that the public health nurses are not many in the country (GHS Annual report, 2012)
iii. It is also worth noting that from eight years experience, the Respondents used almost the same time. This implies that for the manual method to mature, a facility needs to have the public health nurses who are above seven years of experience.

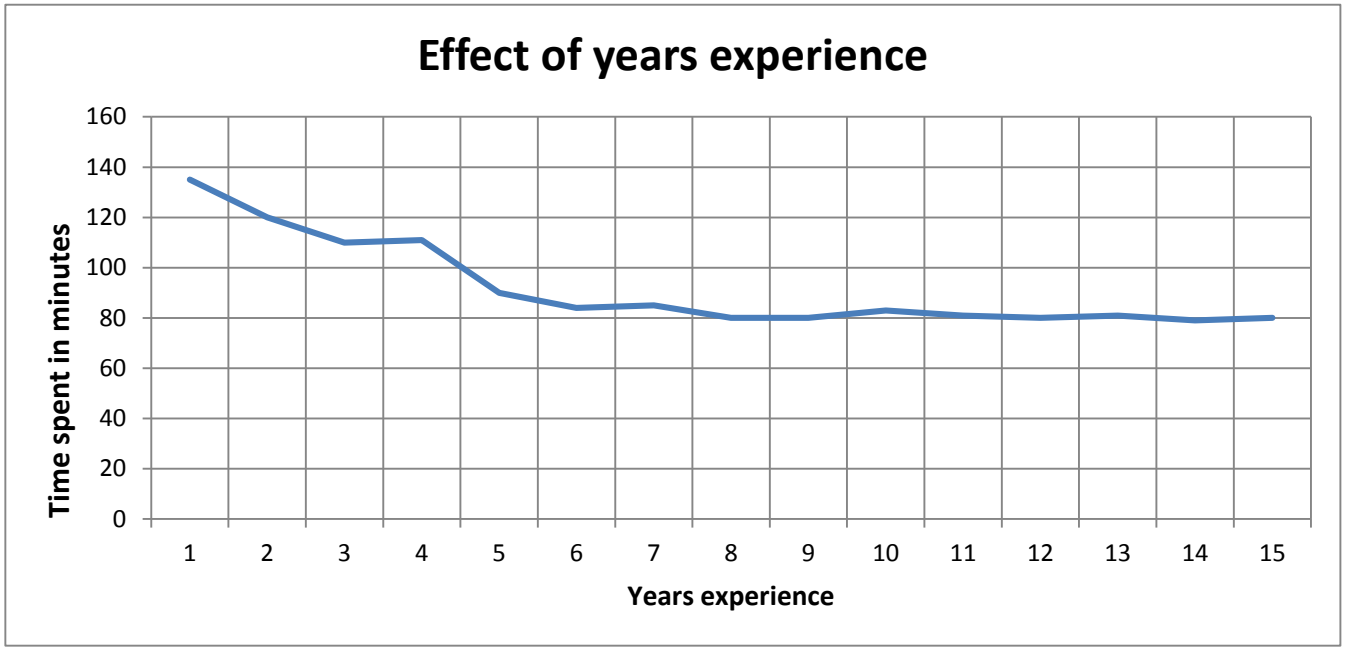
Table 1: Years of work experience distribution of participants in the experiment

Years experience	Frequency	Percent
0-1	1	6.7
2 - 4	3	20.0
5 - 8	3	20.0
9 - 12	4	26.7
13 - 15	1	6.7
15+	3	20.0
Total	15	100.0

Table 2: Respondents' performance

Respondent	Years of work experience	Time used
Respondent #01	12	81
Respondent #02	10	83
Respondent #03	13	80
Respondent #04	6	84
Respondent #05	1	135
Respondent #06	4	111
Respondent #07	2	80
Respondent #08	3	110
Respondent #09	8	85
Respondent #10	16	81
Respondent #11	5	90
Respondent #12	2	120
Respondent #13	9	80
Respondent #14	19	79
Respondent #15	10	80

Figure 10: Morbidity Compilation Times



8.2 Accuracy of the Manual Compilation

Accuracy could not be established until simulations were done using the developed software. In general the experienced nurses had 85% accuracy whilst the others obtained 72%.

8.3 Effect of Manual Compilation on GHS Deliverables

The effects of the manual compilations were obvious: The public health nurses will spend most of their duty hours in producing OPD morbidity reports if they really want accurate results. The ripple effect this situation is that the nurses are overburdened or overwhelmed in the discharge of their duties. The current remedy is that the GHS is striving to keep increasing the numerical strengths of public health nurses but salaries and wages budgetary constraints thwart this approach.

9. SIMULATION RESULT ANALYSIS

The sampled patients data used in the experiments were fed into the developed information system using a data entry clerk who had an average typing speed of 30 WPM. This is a very slow typist yet the data for thirty clients were entered in approximately seven minutes. Computation time on a standalone PC that met the minimum requirement for the thirty patients was in the range of two to five seconds! One can conclude that using the new Information System, the compilation time is reduced by 93%.

10. RECOMMENDATIONS AND FUTURE ENHANCEMENT

The research findings suggest the following recommendations:

The Ghana Health Services and the Ministry of Health should linkup with the academia to research and produce vibrant, robust and inexpensive computing solutions to the various data management challenges currently inundating the sector.

The research results should be piloted in other health centres to fine-tune the information system for national implementation.

Special incentives should be given to individuals who accept postings to the deprived districts.

All health facilities should be empowered to use ICTs to enhance their data management processes since their industry is data-centred.

11. REFERENCES

[1] Andries, J. d., & Bob, B. (2008). The Changing Role of Human Resource Managers for International Assignments. International Review of Business Research Papers Vol.4.No.5. October-November2008 , 166-181.

[2] Arif, M. (2015, March 1). The best software for small businesses (SMEs) - Essential Guide. Retrieved 3 15, 2015, from Tech Target Weekly Computer Magazine: http://www.computerweekly.com/feature/The-best-software-for-small-businesses-SMEs-Essential-Guide

[3] Erickson, T. (2002). Some problems with the notion of context-aware computing, ask not for whom the cell phone tolls. Communications of the ACM, vol. 45, no. 2 , 102-104.

[4] Ghana Health Service. (2006). Annual Report. Accra, Ghana: Ghana Health Service.

[5] Ghana Health Service. (2010). THE HEALTH SECTOR IN GHANA FACTS AND FIGURES. Accra: Ghana Health Services.

[6] Gregor, S., & Jones, D. (2006). Improving the specification of Information Systems design theories, Working Paper,. The Australian National University.

[7] GSS. (2005). Ghana Demographic and Health Survey, Republic of Ghana. ACCRA: GSS.

[8] GSS. (2003). GSS, (2003), Core Welfare Indicator Questionnaire Survey, Republic of Ghana. ACCRA: GSS.

[9] Hall, D. (2003). Building a theoretical foundation for a learning-oriented management system. Journal of Information Technology Theory and Application, vol. 5, no. 2 , 63-85.

[10] Haux, R. (2006). Health Information Systems - Past, Present, Future. International Journal of Medical Informatics , 268—281.

[11] Haux, R. (2006). Health information systems – past, present, future. International Journal of Medical Informatics , 75, 268—81.

[12] Laudon, K. C., & Laudon, J. P. (2012). Management Information Systems, managing the digital firm, 12th Edition. New Jersey, USA: Pearson Education, Inc.

[13] Ministry of Health. (2006, November). Ghana Health Information System: National Level Review and Assessment Report. Retrieved February 15, 2015, from Ministry of Health Ghana Web site: http://www.moh-ghana.org/moh/docs/hmn_act/National%20HIS%20Assessment%20Report.doc

[14] Ministry of Health. (1996). Ghana Health Sector: Facts and figures. Accra, Ghana: Ministry of Health.

[15] Ministry of Health. (2013). Holistic Assessment of the Health Sector Programme of Work 2012. Accra, Ghana: Ministry of Health.

[16] Muhammadi, A., & Mohd, N. A. (2011). "Evaluating CRM Implementation in Healthcare Organization". International Conference on Economics and Business Information (pp. 1-5). Bangkok, Thailand: IACSIT Press.

[17] Nonna, R. (1979). Managing The Crisis In Data Processing. Harvard Business Review 57 (2) , 115–126.

[18] Paolo, L., Nicola, R., Luca, G., & Mariano, C. (n.d.). Health Care Information Systems: Architectural Models and Governance. Milano, Italy: Fondazione Politecnico di Milano and IHCO.

[19] Sanderson, S. M. (2009). Computers in the Medical Office (6th Edition). New York: McGraw-Hill.

[20] Silberschatz, A., Henry, K., & S, S. (2011). Database System Concepts Sixth Edition. NY, USA: McGraw-Hill.

[21] Stair, R. M., & Reynolds, G. W. (2010). Principles of Information Systems (9th Edition). Boston, USA: CENGAGE Learning.

[22] Subhash, V. (2008, October). GNotice: Limitations of Web Applications. Retrieved February 10, 2015, from GNotice Web site: http://www.gnostice.com/nl_article.asp?id=145&t=Limitations_of_Web_Applications

[23] Terry, Q. (2001). Introduction to the Unified Modeling Language. Rational Developer Network , 5-7.

[24] Vital Wave Consulting. (2009). Health Information Systems in Developing Countries,A Landscape Analysis. Vital Wave Consulting.

[25] WHO. (2010). Health financing and access to effective interventions, World Health Report (2010), Background Paper, 8. Geneva: World Health Organization.

[26] WHO. (2012). Management of patient information: trends and challenges in Member States: based on the findings of the second global survey on eHealth. Switzerland: World Health Organization.

[27] WHO. (2000). The World Health Report 2000: Health Systems: Improving Performance. Geneva: World Health Organization.

[28] WHO. (2011, March). World Health Organization. Retrieved February 16, 2015, from World Health Organization Publications: http://www.who.int/goe/publications/en

[29] WIKIPEDIA FREE ENCYCLOPEDIA. (1996). Front and back ends. Retrieved February 20, 2015, from Wikipedia web site: http://en.wikipedia.org/wiki/Front_and_back_ends

International Journal of Science and Engineering Applications
Volume 5 Issue 1, 2016, ISSN-2319-7560 (Online)

Mobile Device Application to locate an Interest Point using Google Maps

Leo Willyanto Santoso
Informatics Department
Petra Christian University
Surabaya, Indonesia

Alvin Wilistio
Informatics Department
Petra Christian University
Surabaya, Indonesia

Lily Puspa Dewi
Informatics Department
Petra Christian University
Surabaya, Indonesia

Abstract: Location-based service is a service mechanism that generates geographical location information. This service will certainly give much benefit as there are many people who would like to find public facilities quickly and easily. The needs to know the existence of public facilities quickly and easily is of great demands in the society. Based on the above problems, it is necessary to develop an application that can get any access to the geographical location in mobile phones. Information from this application is expected to help searching for such public facilities. This Searching for interest point information was read using the Location API in Xcode IOS 4.2. The resulting information was displayed on the application of the iPhone. From the test results, the information of the position was influenced by the availability of the information service provider, carrier signal strength, and mobile phone capabilities.

Keywords: location based service, Xcode iOS 4.2, Google Maps, public service, interest point

1. INTRODUCTION

Information on the location of public facilities at an area is very much needed, especially for a person who is in that area for the first time. Traditionally, a person searches for the information on direction or a specific location is by asking people in the vicinity. The increasing development of technology in the field of telecommunication has driven people to keep abreast with the current development. Many advanced companies in the field of mobile multimedia are competing in introducing new products to meet the needs of the consumers. One of that mobile multimedia is called Navigation Global Positioning System (GPS) or digital map which acts as a sophisticated device on signal receiving and positioning determinant. That product has started to be developed by a multimedia production company due to some frequent phenomena in the society. Some riders or pedestrians often do not know their position when they are in an unrecognized area. They often get lost and do not know the direction to go home. This situation makes them confused, worried and panic. This digital map or GPS navigation device will help a rider or pedestrian not to be afraid of getting lost on the way. In general, the location searching technique usually makes use of two kinds of techniques.

The first technique is by connecting the telecommunication device with the satellite. The second technique is by connecting the telecommunication device with the database server containing the geographical situation of a location which is usually called Geographical Information System (GIS). The first searching method is using the satellite connection or what is usually called Global Positioning System (GPS) [1]. This technique is using a device which can do asynchronization with the satellite to determine the positioning of a location. This application is usually used to find a location based on the coordinate or the latitude and longitude position using satellite as the mediator. The second searching method is using an application used in a browser using computers or mobile phones connected to SIG database server, for instance the foursquare and Google maps. Currently, the positioning technology has already been of much used in the mobile phones. It is even a compulsory capability in a smart phone. Therefore, the use of location searching technology is growing in line with the rapid growth of today software and hardware technology supported by various smart phone vendors. Based on the needs of today society who needs fast and accountable information, a mobile device application to search interest point using Google maps is very much needed [2, 3]. A visit to a gas station to refuel the vehicle is a routine work in the society. The location of ATM is also needed to withdraw some money fast and easy. Hotels, supermarkets and hospitals are also of much needed on certain conditions. This application is expected to help the users not to bother asking people if they want to go somewhere. The goal of this research is to develop an application for iPhone that can help users determining the closest location to the gas stations, hotels, hospitals, supermarkets and ATMs.

2. LOCATION BASED SERVICE (LBS)

Location-based service (LBS) is a service to determine the location of the user and provide certain information needed [4]. LBS can be classified into 3 types. They are Local Information, Traffic and Tracking Information and General Services. Local Information enables users to search services in the vicinity. Traffic and tracking information focuses at assets or people locating. Meanwhile, General Services do not provide information to users. They use the data of the user location. The example of General Services, among others, is emergency services. Various kinds of services have been provided by the use of LBS, such as [5]:

- Providing public facility information services, such as hospitals, gas stations, ATM, Hotels etc.
- Determining and finding the location of someone or something.
- Showing the closest route to go somewhere.
- Direction to go to a certain address.
- Traffic information system services.

The most important thing in LBS is the method in locating users. In general there are two methods used, using Global Positioning System (GPS) or using Cellular Based stations. The technology using Cellular-Based Station is usually called Cell id.

This technology is usually used at certain locations within limited range as it depends on the BTS (Base Transceiver

Station) to receive and transmit signals. Therefore GPS technology is currently used due to its high accuracy and its ability to reach the entire surface of the earth. As this navigation device depends entirely on the satellite, the satellite signal is of utmost important. This satellite-based navigation device can not work with a maximum capacity when there are disturbances at the satellite signal.

3. DESIGN OF THE SYSTEM

The software developed, in general consists of two parts, the application system to be installed at the mobile phone, and the querying system installed at the Google server in taking and generating data into the application. The main function of the application installed at the mobile phone is to identify the geographical coordinate point of the mobile phone to be sent to the server. Whereas the main function of the querying system installed at the server is to display the geographical coordinate data in the form of information to the application. Fig 1. shows the system design in general. The user can access the location information displayed at the application.

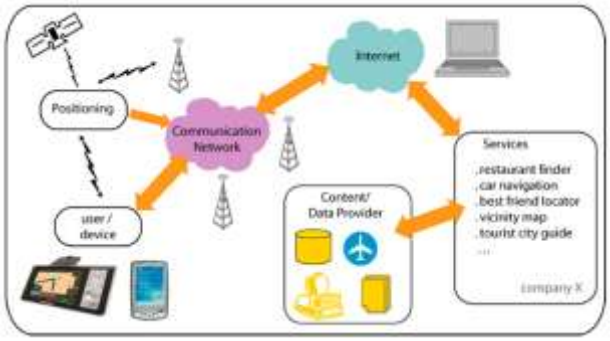


Figure 1. General System Design

3.1 The design of the software system installed at the Mobile phone

Fig. 2 shows the flowchart diagram at the software system of the mobile phone. As the mobile phone application is being run, the main menu using the application icon will be displayed first. Centering the application page, the application will make a checking procedure at the device whether the location service has been activated or not. If it has been activated, the application will be ready to receive the input of the location that is going to be located, such as ATM, Hotels, and hospitals.

If the location service has not been activated yet, the application will display the information that the location service has to be activated first. The use of the location service is to locate the initial position of the user to be converted at the form of a coordinate at the system.

At the main form, the user can set the data device_id which is functioning as the mobile phone identifier towards the data at the database server. The user can also set time interval to update the geographical location and arrange the application to be activated or to be deactivated. If the location reader is activated, then the application can run the location reader at the mobile phone. The process that occurs is when the mobile phone application has managed to read the coordinate of the mobile phone location, the data of the coordinate will be displayed at the mobile phone and sent to the server.

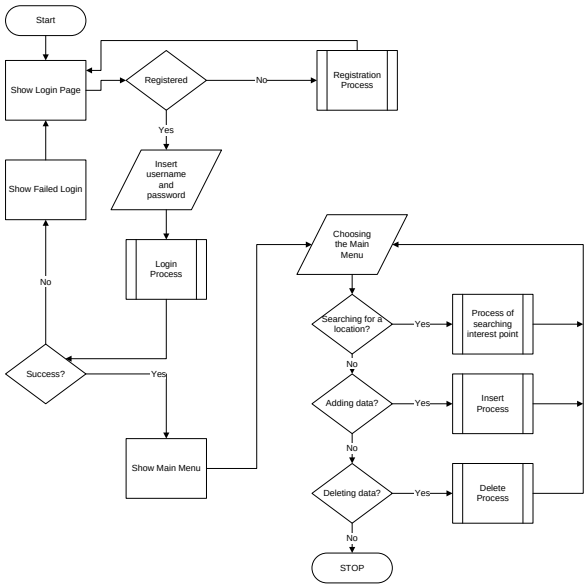


Figure 2. The Flowchart diagram of the software installed at the mobile phone

3.2 The design of the software installed at the Server

The software system installed at the server is a form of API from Google Maps. Its main function is to display the geographical location information at the mobile phone using maps, direction from Google Maps. It can add and save the data at the database. It can also delete the location data from the database. The second database is using PHP. This server is used to save the user database which has been logon and registered at the application.

4. IMPLEMENTATION

In general the implementation was divided into two: the implementation at the iPhone and at the server. The implementation at the iPhone was using Xcode 4.02 and iOS 4.3. The implementation at the server was using:

- Server: Google.com and ta9.petra.ac.id
- Apache/2.2.9 (Debian) DAV/2 PHP/5.2.6-1+lenny16 with Suhosin-Patch
- MySQL client version: 5.0.51a
- PHP extension: My sql.

5. TESTING

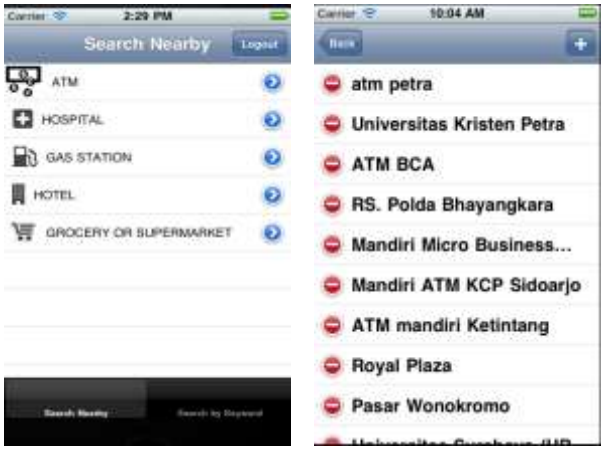
The first step in running this application was by doing the login as seen at Fig. 3 (a) and registering as a member. Without the registration, the application could not be used. To register, the user could select the register button. If the link succeeded to be touched, then a form of registration would be displayed as seen at Fig. 3 (b)

If the login succeeded to be done correctly, the application would proceed to the main menu that was the search for public service location. There would be a selection to do the search nearby and search by keyword.

(a) (b)

Figure 3. Application: (a) Login Form, (b) Register Form

5.1 Steps to search the nearest Public Services Part 1



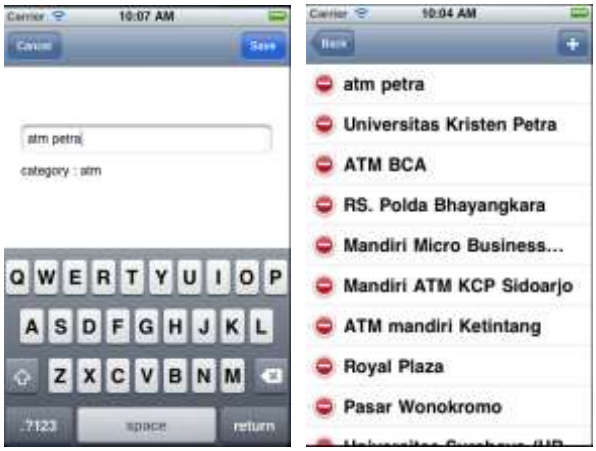
(a) (b)

(c)

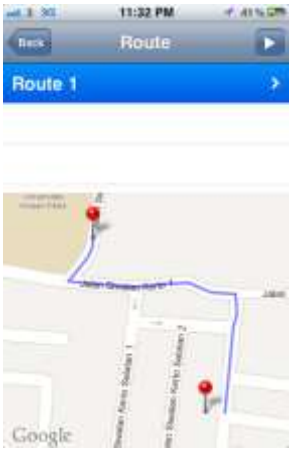
Figure 4. Application: (a) Main Menu Form, (b) List of nearest public services, (c) Empty

If the user would like to search some location nearby, then the menu search nearby should be touched. Then the application would display the nearest location with the category selected before. If there was no data of the target location at Google, then the application would display an empty table view. It can be seen in Fig. 4.

If there was no data at Google, new data could be added by touching the button with the symbol "+". After entering the names of the location in accordance with the category, the button save should be touched.



(a) (b)



(c)

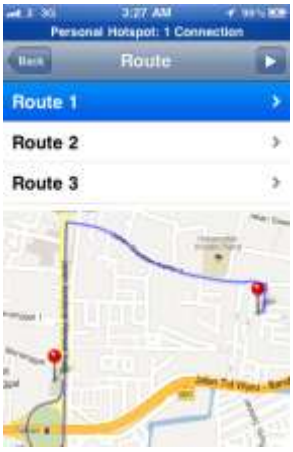
Figure 5. Application: (a) Adding New Place Form, (b) List of nearest public services after adding process, (c) Routing Form

After the process of adding data was successful, the application could proceed to the list menu as followed. After being added to the Google database, a routing procedure could be made by touching the target location. A map with certain route would be displayed. The application would display the route to the tested location towards the nearest ATM. There was also a narration if the button with the symbol ">" was touched, showing how to reach the target location at the map. It can be seen in Fig. 5.

5.2 Steps to search the nearest Public Services Part 2

This testing was using the location at Petra Christian University in searching for a hospital. It can be seen in Fig. 6.

(a)

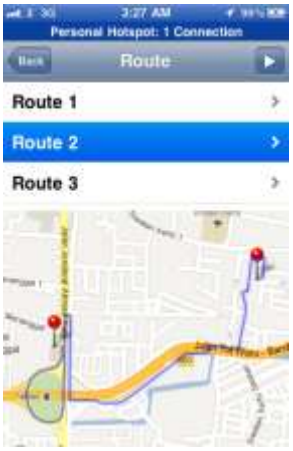


(b)

(c)

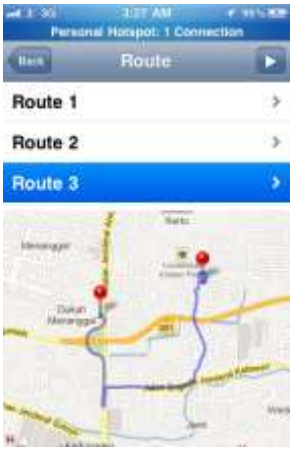
Figure 6. Application: (a) Result after Searching Process, (b) 1st Alternative Routing Process, (c) Direction

The search for the hospital was similar to that search of ATM. The symbol of Hospital at the category menu was touched. A map showing the route to the location of the hospital would be displayed.



(a)

(b)



(c)

(d)

Figure 7. Application: (a) 2nd Alternative Routing Process, (b) Direction, (c) 3rd Alternative Routing Process, (d) Direction

5.3 Steps to Add New Locations to the Application

The following process was the way of adding a new location at the application. The Fig. 8 (a) is the example of inserting a new location at the Supermarket category.

(a)

(b)

(c)

Figure 8. Application: (a) Adding New Place Form, (b) Confirmation after adding process, (c) List Form

A form would appear, ready to be inserted the name of a location into the database. After the save button was touched, the name of the location would be saved and added to the database. A pop up result would appear as seen at Fig. 8 (b). The data of the location added would be displayed at the search list as seen at Fig. 8 (c).

6. CONCLUSION AND SUGGESTION

6.1 Conclusion

Based in the development and testing of the application by detecting the geographical location installed at the mobile phone using API Google Location, it was concluded that:

- The mobile phone application was able to identify the coordinate of the geographical location of the mobile phone at a certain time.
- The success of the application to read the coordinate of the geographical location was influenced by the availability of the information service provider, carrier signal strength, GPS signal strength and the availability of the mobile phone capabilities.
- The geographical coordinate could be read accurately based on the testing of the geographical coordinate accuracy in Surabaya.
- The application requires a stable signal and data connection to guarantee the success of the process of the transmission of the coordinate data to the server.
- The application could be installed at some mobile phones simultaneously based on the testing.
- The mobile phone application required a lot of battery power source.
- The application can not be free open source software, as Google limits the use of API to a maximum of 25.000 map load. Payment to Google is required if more map load is needed.
- A long approval process allowed the possibility of delays in inputting and managing the location that has been input.
- A long querying process of inputting a new location to the Google server, was allowing the possibility of the application developer to have to create a database first.
- It took some time to route Google to do the search for the first time. Therefore the sophisticated database availability is a necessity in routing

6.2 Suggestion

The following suggestions are expected to encourage further developments, such as:

- The mobile phone application can be developed at other platforms supporting Google API Location such as Android, Symbian and Windows phone.
- The mobile phone application can be developed further to save the required battery resources.
- It might be necessary to try other API locations such as API Bing, yellow pages etc.
- Before developing this application, a complete database on public services has to be created first. Otherwise it will create problems in running the application.

7. REFERENCES

[1] X.Y. Zhaoa, “On processing GPS tracking data of spatio-temporal car movements: a case study”, Journal of Location Based Services, Vol. 9, No. 4, 2015, pp. 235-253.

[2] Udell, S. 2009. Beginning Google Maps Mashups with Mapplets, KML, and GeoRSS. New York, NY: Apress.

[3] Peterson, M. P. 2012. Online Mapping with APIs, Online Maps with APIs and Mapservices (M. P. Peterson, ed.), Springer, pp. 3-12.

[4] Kremera, D and Schliedera, C., “Less is more: empirical design criteria for a tourist place recommendation service which decelerates the visiting experience”, Journal of Location Based Services, Vol. 8, No. 4, 2014, pp. 268-284.

[5] S, Hu, 2012. Multimedia Mapping on the Internet Using Commercial APIs, Online Maps with APIs and Mapservices (M. P. Peterson, ed.), Springer, pp. 61-71.