PAPER NAME

Rayuwati_2019_J._Phys.__Conf._Ser._13
64_012024.pdf

AUTHOR

Rayuwati Rayuwati

WORD COUNT

3318 Words

CHARACTER COUNT

17600 Characters

PAGE COUNT

9 Pages

FILE SIZE

938.9KB

SUBMISSION DATE

Jun 26, 2023 4:46 PM GMT+8

REPORT DATE

Jun 26, 2023 4:46 PM GMT+8

● **10% Overall Similarity**

The combined total of all matches, including overlapping sources, for each database.

- 8% Internet database
- Crossref database
- 6% Submitted Works database

- 9% Publications database
- Crossref Posted Content database

● **Excluded from Similarity Report**

- Bibliographic material
- Cited material
- Manually excluded sources

- Quoted material
- Small Matches (Less then 10 words)
- Manually excluded text blocks

PAPER • OPEN ACCESS

Comparison of Encoding and Decoding Methods for Binary Files

To cite this article: Rayuwati *et al* 2019 *J. Phys.: Conf. Ser.* **1364** 012024

View the article online for updates and enhancements.

Comparison of Encoding and Decoding Methods for Binary Files

Rayuwati[1]*, Mahmud Mustapa[2], Askar Taliang[3] and Akbar Iskandar[4]

[1]Universitas Gajah Putih, Takengon, Indonesia
[2]Universitas Negeri Makassar, Makassar, Indonesia
[3]Universitas Patria Artha, Indonesia
[4]Department of Informatics, STMIK AKBA, Makassar, Indonesia

* rayuwati68@gmail.com

Abstract. The main problem to be an obstacle in communication between platforms, networks such as local networks or the Internet is a problem in data transfer. The problem often found in a hardware which is cannot receive data that is processed by other hardware. So to change data between two computer systems, it is necessary to convert data from one form to another. One solution to this problem is transfered data or file, previously converted to the plaintext form first. Conversion to the plaintext form can be supported by most computer systems. In this study, a software was designed using Visual Basic programming language that can convert binary files to plaintext using the Base 64 encoding or decoding method, UUEncode / UUDecode, XXEncode / XXDecode, BinHex 4.0, Hexadecimal, and Quoted-Printable. The results of the design is produce a software that can be used as an alternative program to convert files into text files, besides that the program can also display a benchmark comparison graph (speed) of each method and comparison of overhead (the ratio of the size of the file to the file before encoding).

1. Introduction

Problems in exchanging data or sharing files between different networks or computer platforms may cause several problems. Incompatibility between one system with another system can cause data exchange to be impossible [1]. In addition to these problems other problems [2][3] may be related to sending messages in the form of e-mail on the Internet network. The e-mail delivery system on an Internet network is designed to exchange data only in plain text (ASCII) [4][5].

To exchange the data, encoding techniques are usually used. Encoding comes from the word "to encode" which means encoding, writing or making changes. In other words encoding is the process of encoding a data into symbols, characters or certain codes [6][7][8]. In addition, encoding can also be interpreted as the process of converting a form of data into a form of data in another format. While decoding comes from the word "to decode" which is means changing the encoding process back to its original form[9][10]. The decoding process is opposite of the encoding process.

To overcome problems like the above, data conversion technique was developed known as encoding or decoding technique. If an e-mail usually has an attachment such as graphics, audio, video, and so on, then the attachment file must be changed in plain text form first. When received, the e-mail

containing the attachment will be decoded and reshaped. In addition, binary conversion to ASCII makes it easy to send and receive binary files via e-mail or communication systems based on plain text [11][12]. There are several encoding / encoding methods used to convert data files into plain text such as Base64, UUEncode, XXEncode, USR, BINHEX, BTOA, BOO, and Quoted-Printable.

All of the methods mentioned above have the same function, so that a binary data can be safely transferred via e-mail or through equipment that only supports text transfer, the difference in the algorithm part of each method used. These methods convert 8-bit binary data (EXE, COM, GIF, BMP, etc.) to ASCII 8-bit (plain text) and vice versa through a process known as bit-shifting.

2. Related Works

The development of computers and communication systems result is the need for tools to protect information in digital form and to provide security services became very important [5][13][14]. At present most communication systems leave some data and information recorded or stored. To avoid the data falling into the unauthorized person, usually the data will be encrypted to ensure the confidentiality of the data itself [15][16]. In addition, one of the obstacles in the process of sending digital data is the absence of a standardized standard for any software and hardware needed for data transmission. Often data must be converted into other forms before being transmitted.

3. Research Methodology

3.1. Calculating Process Speed

The speed of the encoding and decoding process is done by using an API (Application Programming Interface) function from Windows, namely the GetTickCount () function. This function is found in the library "KERNEL32.DLL". This function has no parameters and gives an output value in the long integer data type[17]. The purpose of this function is actually to find out how long the Windows system runs in units of milliseconds. The way to declare it in Visual Basic modules is as follows::

```
Declare Function GetTickCount Lib "kernel32" () As Long
```

For calculating the speed of the process carried out in the following way. Before the encoding and decoding process are run [18], the value of GetTickCount () is first taken and stored in a variable memory. The next step is to run all the encoding and decoding routines, after the process is complete, retrieve the GetTickCount () value and save it to another variable. The difference between the two variable values is the length of the encoding and decoding process that is done. For more details, consider the following code example:

```
StartTime = GetTickCount( )
        Routine Process of encoding / decoding
        ........
        ........
EndTime = GetTickCount( )
ElapsedTime = EndTime – StartTime
```

3.2. Output Summary

After completing the encoding and decoding process, the program will display a description of the process carried out. This information is summarized in the "Summary" Tab. In this Summary there are information such as:

(1) Type of Encoding / Decoding method

This method accordance to the type of encoding / decoding method chosen during the process.

(2) Encoded / Decoded File

Namely the file name and path that will be encoding and decoding.

(3) Encoded / Decoded File Size

The file size will be encoded or decoded in bytes. Obtained by using the FileLen (pathname) function.

(4) Output File

Namely the output file name, has an extension according to the method chosen.

(5) Output File Size

That is the size of the output file generated after the encoding and decoding process, in bytes. Obtained by using the FileLen (pathname) function.

(6) Elapsed Time

The length of time required for the encoding and decoding process in one millisecond. Obtained by using the GetTickCount () API function.

(7) Running Time

Representation of time in units of hours, minutes, seconds and milliseconds.

(8) Overhead

Is the percentage of adding or reducing file size after being encoded and decoded. Obtained by using the formula:

$$Overhead = \frac{Size\ of\ result\ file - Size\ of\ original\ file}{Size\ of\ original\ file} \times 100\%$$

3.3. Algorithm

Algorithms are steps and gradual and specific sequences of problems [19][16][20]. This algorithm translated into programs using certain programming languages. Algorithms are used to analyze and explain the sequence and relationship between the activities to be pursued. In addition the algorithm also functions to solve a problem so that the desired goal was achieved.

Algorithm for encoding strings in Base64:
1. Change the ASCII string character to decimal.
2. Change the decimal value in binary form.
3. Combine all binary values so that they become one single string in the form of a string of bits.
4. Break the string into one group, each consisting of 6-bits.
5. Change the binary value of each group to the decimal form again.
6. Change the decimal value into the Base64 character form by using the Base64 encoding table

Algorithm for decoding strings in Base64 (the opposite of the encoding process):
1. Change the character of the string in the form of a Base64 decimal value.
2. Change the decimal value into binary form.
3. Combine 6-bits in each character into one main string with a binary digit shape.
4. Break the string into one group, each consisting of 8-bits (starting from left to right).
5. Change each 8-bit binary value in decimal.
6. Change the decimal value to ASCII

The data used in this research is secondary data,secondary data obtained. [21]

4. Result and Discussion

4.1. The Design of Main Module

The main module is designed for declaring functions such as checking the existence of files, checking file extensions, retrieving file names, functions of browsing folders. The types of functions or routines contained in the main module can be seen in the following table:

Table 1. Routine Table on the Main Module

Routine Name / Function	Parameter	Return Value	Utility
Function Browse For Folder ( )	*(ByVal lngHwnd As Long, ByVal strPrompt As String)*	*String*	*Browsing folder*
Function GetTickCount ( )	*Tidak Ada*	*Long Integer*	Mengecek lama sistem *Windows* Berjalan
Function FileExist ( )	*(FileName As String)*	*Boolean*	Mengecek keberadaan *file*
Function GetFileName ( )	*(cFileName As String)*	*String*	Mengambil nama *file*
Function GetFileExt ( )	*(cFileName As String)*	*String*	Mengambil ekstensi *file*

4.2. Xceed Binary Encoding Library

For the purposes of the encoding and decoding process in this program, the Xceed library encoding (www.xceedsoft.com) is used. Xceed Binary Encoding Library is a library for developers of Windows operating systems which allows an application to encode binary data and files into text or vice versa. This library can do the process of encoding and decoding buffer memory, strings, BLOBs, streaming data or single files. This library supports various types of encoding / decoding methods such as UUEncode, XXEncode, Base64, BinHex 4.0 and Quoted-Printable. [22]

This library also supports all 32-bit programming languages that use ActiveX controls, including Visual Basic, Visual C ++, Delphi, C ++ Builder, Visual FoxPro, Visual Interdev, Office 97. To be able to use the Xceed Binary Encoding Library in Visual Basic it must be created a reference to the Xceed Binary Encoding Library into Visual Basic. The trick is to use the Project → References command, then a list of references from the library that can be used in Visual Basic will be displayed.

The library reference list is activated by selecting the library from the Xceed Binary Encoding Library with the name: Xceed Binary Encoding Library v1.0 by marking the check mark on the library then pressing the "OK" button. If it is not in the type library reference for the Xceed Binary Encoding Library, then install the library (the library can be downloaded on the website: www.xceedsoft.com). Then you can do your own search by pressing the "Browse" button and looking in the / Program Files / Xceed Components / Xceed Binary Encoding Library / Bin folder for files with the name "XCEEDBEN.DLL" (without quotes) then select the file to use inside Visual Basic. [23]
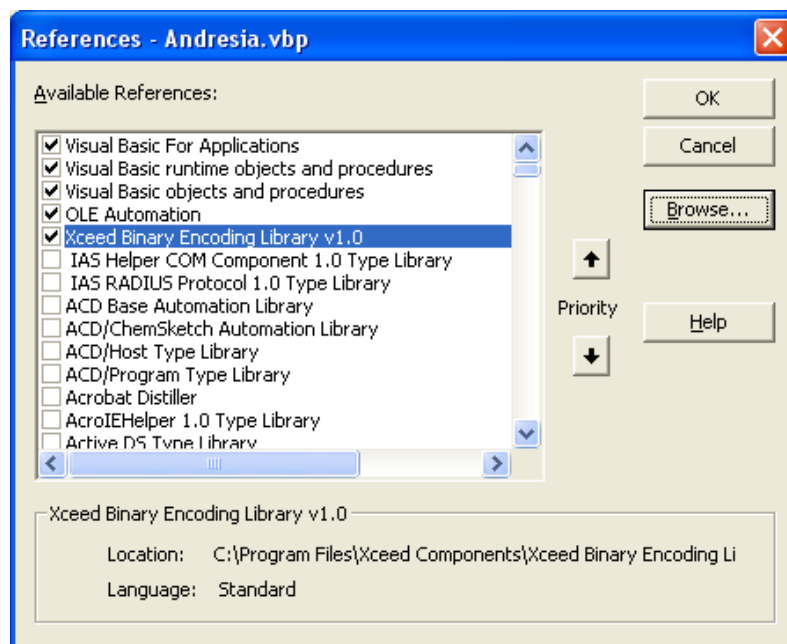
Figure 1. Selected of Xceed Binary Encoding Library V 1.0 for Visual Basic

With the addition of the library into Visual Basic it can start programming with the Xceed Binary Encoding Library. To find out more about the routines and objects provided by the library, you can press the F2 key or click on View → Object Browser in Visual Basic IDE (Integrated Development Environment) to view the library contents of the Xceed Binary Encoding Library. The partial form of the Xceed Binary Encoding Library objects is shown in the following figure. In addition, the library into Visual Basic can start programming with the Xceed Binary Encoding Library. To find out more about the library contents of the Xceed Binary Encoding Library, you can press the F2 key or click on View. The partial form of the Xceed Binary Encoding Library objects is shown in the following figure.
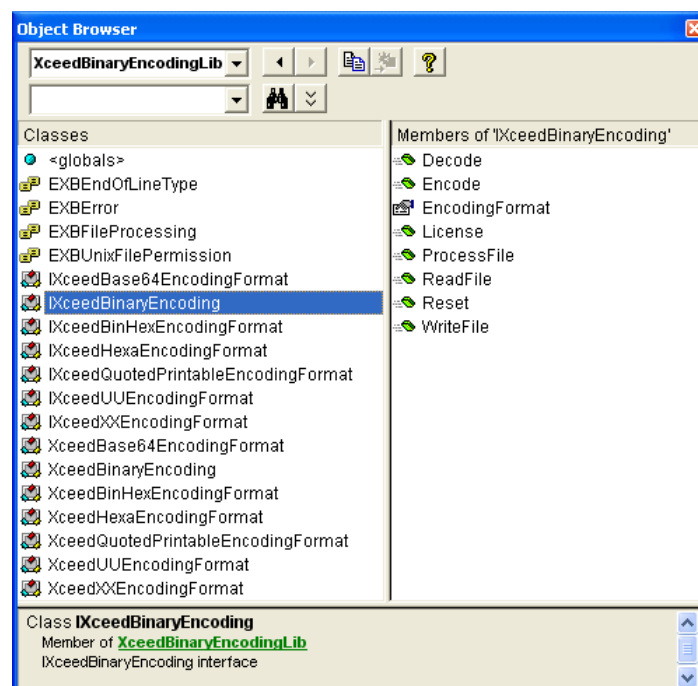

Figure 2. Object Browser Visual Basic

4.3. Run The Program

To use this program, run the executable file (Andresia.exe) from the location where the file was copied or if the program is installed, the program shortcut icon can be run from the Start button → Windows program and the display will appear as shown in Figure 3. This program only consists of one single form. All encoding and decoding processes are carried out in this form.

To do the encoding and decoding process, the first thing to do is to select the file that will be processed first by clicking on the picture file button, then the path name and file will appear in the "Source File:" and rich text box according to the method the selected encoding or decoding will contain the contents of the file displayed in hexadecimal form.
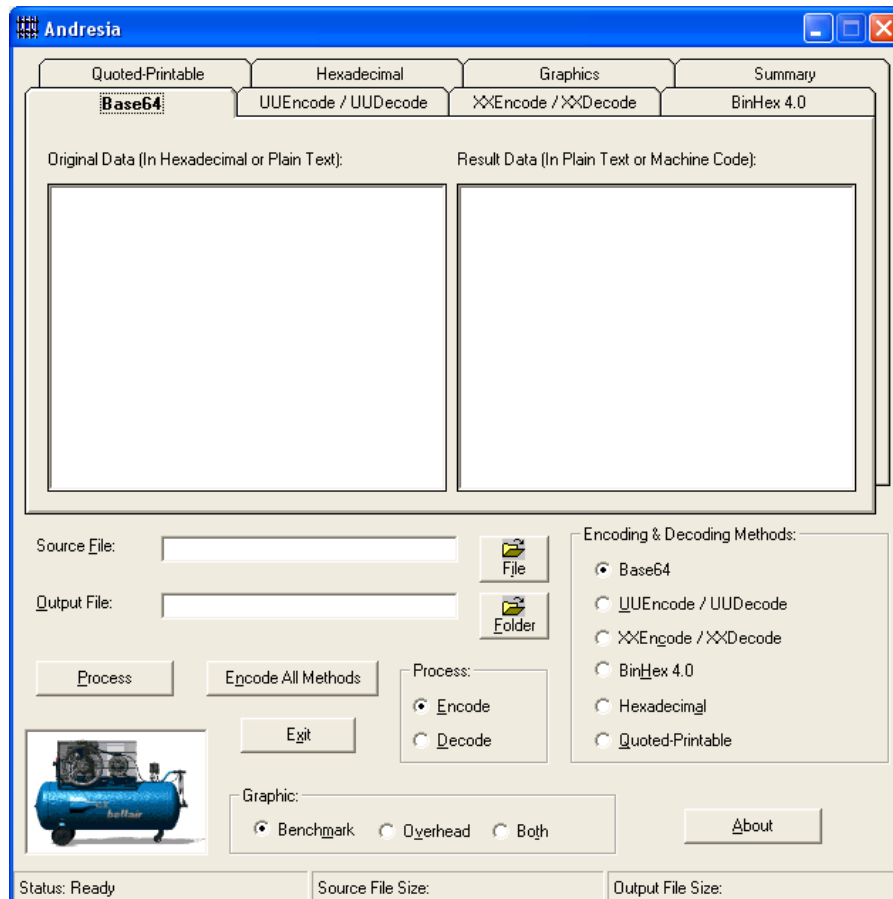

Figure 3. Display Program

5. Conclusion

Based on the discussion that has been done regarding the program of encoding and decoding binary files into plain text, the authors took several conclusions as follows:

1. The encoding process is the process of converting binary files into pure plain text or ASCII files. The goal is that files can be sent over a network that only supports transmission in 7-bit data or in interfaces that only support text like e-mail. While the decoding process is the opposite of the encoding process, which is to change back the plain text file back to its original binary file. This is done when the file has been transmitted.

2. From the test results obtained the fastest average method in the event that the encoding process is Base64 and which has the smallest average overhead is the BinHex 4.0 method. This happened because the method applies compression with run length encoding. In some tests carried out on different file types, the overhead of the BinHex 4.0 method were negative, this means that the output file is smaller than the original file. In terms of transmission speed, it is certain that the BinHex method is transmitted faster than the other methods for the same file size.

3. The method that has the slowest decoding process is Quoted-Printable. Whereas in terms of overhead, the method that has the biggest overhead is also the Quoted-Printable method. From the results of testing this method has an average overhead of above 100%, meaning that the file size of the results becomes more than twofold.

4. From the program output it is known that the average overhead of the method made is above the average of 30%. This means the resulting file size is 1/3 times larger than the size of the original file before encoding. Swollen file sizes can waste storage media space and files will be transmitted much slower 1/3 times as well.

5. In the case of algorithm methods that have the simplest algorithm is the Hexadecimal method compared to the other six methods. While the most complicated algorithm among the six methods is the BinHex 4.0 method.

References

[1] B. R. Setiadi, S. Subagyo, A. B. Johan, M. Nurtanto, S. Sugiyono, and H. Nurdiyanto, "Mobile pocketbook of the 4Cs skills-oriented inform of Quick Response Code," in *Proceedings of the International Conference of Social Science*, 2019.

[2] M. Nurtanto, S. Nurhaji, D. Widjanarko, M. B. R. Wijaya, and H. Sofyan, "Comparison of Scientific Literacy in Engine Tune-up Competencies through Guided Problem-Based Learning and Non-Integrated Problem-Based Learning in Vocational Education," in *Journal of Physics: Conference Series*, 2018, vol. 1114, no. 1.

[3] R. Rahim *et al.*, "Hashing Variable Length Application For Message Security Communication," *ARPN Journal of Engineering and Applied Sciences*, vol. 14, no. 1, pp. 259–264, 2019.

[4] M. Syafrizal, *Pengantar jaringan komputer*. Penerbit Andi, 2005.

[5] W. Supriyanto and A. Muhsin, *Teknologi informasi perpustakaan*. Kanisius, 2008.

[6] B. R. Setiadi, Setuju, and M. Nurtanto, "Installation of Closed Circuit Television To Produce a Quality Assessment Process in Vocational Learning," Dec. 2018.

[7] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, no. 6, pp. 520–541, 1987.

[8] M. Nelson and J.-L. Gailly, *The data compression book*. M & t Books New York, 1996.

[9] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient erasure correcting codes," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 569–584, 2001.

[10] A. Moffat, R. M. Neal, and I. H. Witten, "Arithmetic coding revisited," *ACM Transactions on Information Systems (TOIS)*, vol. 16, no. 3, pp. 256–294, 1998.

[11] E. V. Haryanto and U. P. Utama, *Jaringan Komputer*. Penerbit Andi, 2012.

[12] A. Iskandar, Rismawati, and R. Rahim, "Designing Application for Performance Assessment to Measure Employee Profesionalism in Goverment," in *Joint Workshop KO2PI and The 1st International Conference on Advance & Scientific Innovation*, 2018, pp. 154–161.

[13] W. Stallings, *Cryptography and Network Security, 4/E*. Pearson Education India, 2006.

[14] R. Ratnadewi, D. Sudrajat, A. E. Sari, S. U. Ady, and D. Rianita, "Security application using data encryption standard algorithm," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 10, no. 7 Special Issue, pp. 1879–1882, 2018.

[15] A. Iskandar *et al.*, "Web based testing application security system using semantic comparison method," in *IOP Conference Series: Materials Science and Engineering*, 2018, pp. 1–6.

[16] A. Iskandar, E. Virma, and A. S. Ahmar, "Implementing DMZ in improving network security of web testing in STMIK AKBA," *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 2.3 Special Issue 3, 2018.

[17] B. W. Kernighan and D. M. Ritchie, *The C programming language*. 2006.

[18] D. LaBerge and S. J. Samuels, "Toward a theory of automatic information processing in reading," *Cognitive psychology*, vol. 6, no. 2, pp. 293–323, 1974.

[19] M. Rivki and A. M. Bachtiar, "Implementasi Algoritma K-Nearest Neighbor Dalam Pengklasifikasian Follower Twitter Yang Menggunakan Bahasa Indonesia," *Jurnal Sistem Informasi*, vol. 13, no. 1, pp. 31–37, 2017.

[20] D. Sudrajat *et al.*, "Expert system application for identifying formalin and borax in foods using the certainty factor method," *Eurasian Journal of Analytical Chemistry*, vol. 13, no. 6, pp. 321–325, 2018.

[21] S. G. Fashoto, O. Amaonwu, and A. Afolorunsho, "Development of A Decision Support System on Employee Performance Appraisal using AHP Model," *JOIV : International Journal on Informatics Visualization*, vol. 2, no. 4, pp. 262–267, Aug. 2018.

[22] M. Mesran *et al.*, "Combination Base64 and Hashing Variable Length for Securing Data," *J. Phys.: Conf. Ser.*, vol. 1028, p. 012056, Jun. 2018.

[23] D. Abdullah *et al.*, "Super-Encryption Cryptography with IDEA and WAKE Algorithm," *J. Phys.: Conf. Ser.*, vol. 1019, p. 012039, Jun. 2018.

9 Kundharu Saddhono, Erwinsyah Satria, Alfian Erwinsyah, Dahlan Abdull... <1%
Crossref

10 daoning.com <1%
Internet

🔷 turnitin

● **Excluded from Similarity Report**

- Bibliographic material
- Cited material
- Manually excluded sources

- Quoted material
- Small Matches (Less then 10 words)
- Manually excluded text blocks

EXCLUDED SOURCES

Rayuwati, Mahmud Mustapa, Askar Taliang, Akbar Iskandar. "Comparison of ...   **73%**
Crossref

researchgate.net   **11%**
Internet

Universitas Putera Indonesia YPTK Padang on 2023-02-23   **8%**
Submitted works

Politeknik Negeri Bandung on 2020-07-15   **8%**
Submitted works

Politeknik Negeri Bandung on 2019-12-18   **8%**
Submitted works

Politeknik Negeri Bandung on 2019-12-18   **8%**
Submitted works

Universitas Sebelas Maret on 2021-04-21   **8%**
Submitted works

Politeknik Negeri Bandung on 2023-05-25   **8%**
Submitted works

Politeknik Negeri Bandung on 2023-05-24   **8%**
Submitted works

Politeknik Negeri Bandung on 2023-05-23

Submitted works

8%

Politeknik Negeri Bandung on 2023-05-23

Submitted works

8%

Nopriadi, Akbar Iskandar, Listia Utami, Okma Yandri. "Utility Software Design t...

Crossref

8%

repository.unimal.ac.id

Internet

5%

repository.unitomo.ac.id

Internet

5%

backend.orbit.dtu.dk

Internet

5%

karyailmiah.trisakti.ac.id

Internet

2%

repository.uin-suska.ac.id

Internet

2%

EXCLUDED TEXT BLOCKS

Journal of Physics: Conference SeriesPAPER • OPEN ACCESSYou may also like

University of Surrey on 2021-10-20

and TechnologyJournal of Physics: Conference SeriesIOP Publishing

A M H Pardede, Yusdiana Br Sembiring, Akbar Iskandar, Dyah Retno Pitasari et al. " Implementation of Data ...