ISSN:0975-9646

A Novel Lifelog Management Scheme for Improving Health-Related Quality of Life

Myung-Kyu Yi[1], Taeg Keun Whangbo[2]
IT College, Gachon University
Gyeonggi-Do, South Korea

Abstract— **The recent proliferation of digital cameras and camera-equipped smart phones permits an end user to record everyday events and experiences. A vast amount of daily activity records, called lifelogs, represent a valuable health data resource. The advent of lifelogging technology has been changing the way healthcare services are being used, emphasizing the importance of personalizing services. To provide accurate personalized healthcare service, it is necessary to gather as much personal lifelog data as possible about patients. However, as the amount of lifelog data becomes vast, the need for an efficient way of managing and storing such vast information emerges. In this paper, we propose a novele lifelog management scheme based on the characteristics of collected data in a hierarchical lifelog system. In our proposed scheme, the available storage for the preprocessing of lifelog data is selected to optimally utilize computing and storage resources. The analytical results indicate that our approach achieves better performance than the existing lifelog management scheme by selecting the appropriate storage.**

Keywords—**lifelog, lifelog management system, healthcare**

I. Introduction

With the ever-improving performance of smartphone cameras and the universal dissemination of cloud services, users can now record and store the events in their daily lives more easily and conveniently. Users can now not only store pictures, videos, or memorandums but also share them with others via SNS, such as through Twitter or Facebook, by recording or looking up when, where, and whom they met, and for what purpose, while the location information collected via the GPS terminals or GPS sensors in smartphones allows users to systematically track and record diverse information such as location, biometric, and exercise data. Lifelog is the act of recording human life in the form of digital data, or signifies the recordings themselves. Such act of recording is propelled by the users' gut instinct, and is so pervasive in people's daily lives in the form of a diary, whose origin can be traced back to the prehistoric frescos left in many caves. Lifelogging is made possible by certain kinds of technologies, including webcams and other surveillance equipment, wearable computers, streaming video and other data, complimented by the ubiquity of data storage media that can handle large amounts of digital information. The prolific use of social media could even be called lifelogging, since many of these tools serve to update the user's location, purchases, activities and even feelings in real time. The use of Facebook, Twitter, Foursquare and other social media applications probably constitutes the majority of this type of activity, although there are more dedicated lifeloggers using more specific technologies to map their activities in real time. A lifelog produces a dataset consisting of continuous streams of sensor data. Sharing this information has a wide range of advantages for both user and society. By analyzing the lifelog data, we are able to know what has transpired in an individual, as well as where and when it happened. The main objectives of the life log systems are to help people record their experiences in digital format and to provide convenient ways to retrieve and navigate the desired events from recorded experience. If the concept of lifelog is expanded a little further, it may include a food log, which signifies text or photographic recordings of what one eats every day, or the act of recording the titles of the books that one reads, the songs that one listens to, or the movies or TV programs that one watches. One can also keep track of his or her blood pressure, liver somatic indices, and blood type as well as other daily-life data, such as hours of sleep, work hours, credit card transaction data, and consumption patterns. Various biometric data can now also be tracked, such as the movement of the pupils while reading or watching TV. Lifelog started with the simple purpose of recording or sharing mainly data regarding one's personal life, but with the introduction of advanced specialized analytic methods by many corporations, a new type of business based on the lifelog recently emerged, with an aim of improving the quality of people's personal lives. An example of such would be the personalized recommendation service often found in online shopping malls or Internet portals. Such service relies on the analysis of the usage patterns of individual users to come up with recommendations customized for individuals. Amazon, a leading online bookstore, automatically analyzes the purchase and query data of each visitor. The next time the visitor logs onto the website, Amazon shows the visitor "other products related with the recently checked products" or "the other products bought by the other customers who also bought this item." These are all examples of a system's adoption of the lifelog. The advancement of lifelog technologies is forecasted to bring a host of significant changes to people's daily lives. Compared to the incomplete and fragile human memory filled with errors, the lifelog records each piece of information related to a person, no matter how minute, thereby equipping people with a photographic memory. Also, in the event of an accident or a terror attack, for instance, detectives would be able to reconstruct the crime scene digitally by analyzing the lifelog information obtained in the crime scene. Moreover, the biometric data collected in real time can help people manage their health significantly better, and can play a bigger role in protecting the social minorities, such as children and aged people. By analyzing the lifelog information, companies will be able to provide people with diverse customization services or pinpoint target users with behavioral target advertisements supported by the technologies that learn and guess user behavior, activities, or intentions. As such, the service linked with the lifelog data is shifting its focus to individuals. As people will no longer look for what they want but will simply choose from the given offerings, the use of information will become increasingly effective. As a result, it is forecasted that the importance of the classification and selection of information will grow in the future.

Lifelog management is one of the important issues in efficient lifelog sharing and management. To support healthcare services, there is a large amount of data useful for medical purposes within the voluminous and varied data gathered in daily life. All lifelog data should be managed effectively because it is necessary to gather as much health information as possible in order to achieve high-quality screening and treatment. If the data volume and number of processing operations increase in the lifelog system, the processing throughput has to be improved appropriately for efficient life log management. In this paper, we propose a novel lifelog management scheme, called AHLM, which is built based on the characteristics of collected data in hierarchical lifelog system. The rest of the paper is organized as follows. Section 2 introduces an overview of related work. Section 3 illustrates the proposed lifelog management scheme. Section 4 shows the evaluation of the proposed scheme's performance and analysis of results. Finally, our conclusions are presented in Section 5.

II. RELATED WORK

In recent years, several approaches have been used to implement a lifelog system. MyLifeBits is designed to store and manage everything in a person's lifetime that can be captured in digital format[1]. This project includes full-text search, text and audio annotations, and hyperlinks. MyLifeBits initially focused on capturing and storing scanned and encoded archival material. It evolved to having a goal of storing everything that could be captured. Rawassizadeh et al. propose Ubiqlog which is a generic mobile phone-based life-log framework, uses a mobile phone as a device for life logging[2]. Ubiqlog consists of an open architecture and features a data model specifically designed for life logging. UbiqLog is a generic and holistic framework, which can be used for different cases and can be configured based on the user's requirements. Its openness in terms of sensor configuration allows developers to create flexible, multipurpose life-log tools. Experience Explorer is a client-server platform that enables life logging, via mobile context collection, and processes the data so that meaningful higher-level context can be derived[3]. This can then be mashed-up with third party Internet services, in order to provide richer social networking and content associations while keeping privacy and security in mind. Another lifelog system implementation is discussed in [4]. This work focuses on realtime storage and retrieval of lifelog in a ubiquitous environment. The developed system supports semi-automatic activity analysis and provides an intuitive graphical interface for users to browse Mobile Lifelogger their lifelogs that correlates the space and temporal information of the displayed sensory data. Prananto et al. proposes a multi-level abstracted metadata datbase design to retrieve every day experiences in various ways with a user-friendly web interface[5]. The simulated performance with real lifelog media data shows that suggested multi-level metadata database structure can retrieve experience media data around four times more efficient compared with usual single-level annotation-based metadata database in the sense of retrieval time. Chennuru et al. present an ongoing research of using mobile phones to record and index lifelogs using activity language[6]. By converting sensory data such as accelerometer and GPS readings into activity language, we are able to apply statistical natural language processing techniques to index, recognize, segment, cluster, retrieve, and infer high-level semantic meanings of the collected lifelogs. Based on this indexing approach, the proposed lifelog system supports easy retrieval of log segments representing past similar activities and automatic lifelog segmentation for efficient browsing and activity summarization. Kim et al. propose hierarchical structured data logging to support lifelog based personal services and to reduce the processing complexity and storage cost[7]. They present hierarchical lifelog data logging to optimally utilize computing and storage resources. A major drawback to current lifelogging practices is existing lifelog management system do not consider the characteristics of collected lifelog data. Moreover, all lifelog data should be managed effectively because it is necessary to gather as much health information of user as possible in order to achieve high-quality screening and treatment for the healthcare service. Therefore, we propose a novel lifelog management scheme based on the characteristics of the collected data.

III. ADAPTIVE LIFELOG MANAGEMENT SCHEME

In this section, we describe the proposed scheme, called AHLM, for effective lifelog management. An overview of system architecture of AHLM scheme is shown as Fig. 1. In the proposed scheme, a mobile device uses a GPS log, microphone, temperature, moisture, and light sensor. The mobile device collects personal lifelog data from diverse sensors and user behavior logs such as mail, web history, application usage statistics. Some information must be converted into semantic information. Therefore, the mobile device extracts features by pre-processing the log data and transfers the lifelog data into a stable server or in the cloud[8]. Generally, all raw data are stored in the log server and pre-processed data are transferred into event server to recognize the event. Therefore, all extracted lifelog data are stored in the mobile device to provide a healthcare service. All lifelog data should be managed effectively because the collected lifelog data in real time become large volumes of data from a range of sources and sensors. The complexity of computation will increase as the volumes of lifelog data become larger. All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified. Similar to [7], we proposed the enhanced mechanism of hierarchical structured data logging and caching for lifelog data management.

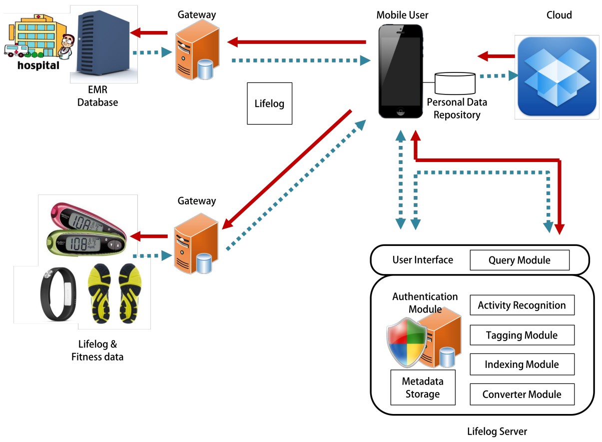

Fig. 1 The Proposed system architecture

Fig. 2 shows the proposed lifelog system with a three-level storage hierarchy. The storage hierarchy consists of three levels : mobile device, server, and cloud. The lifelog data is stored along each of these levels either temporarily or permanently. The aim of the storage hierarchy design is to manage the lifelog system resources more efficiently. The storage hierarchy is able to ensure the adequate storage performance with a combination of a slow high capacity storage and a fast low capacity storage. At the top of the hierarchy is a storage for mobile device. If the data is found at this level, the lifelog data can be loaded quickly. A mobile device, however, has power limitations, which is a finite battery life. Moreover, it is quite a small space compared with a computer's storage capacity even though most mobile device support external SD storage cards. Battery and storage are a mobile device's most important resources in the lifelog system. In order to perform complexity computation, therefore, large volume data should be stored and analyzed on the server or cloud. The lower level of the storage hierarchy is the storage of cloud which provides users with immediate access to a broad range of resources with high capacity and excellent stability. But, it is not easy to access the lifelog data directly because it can come from network latency. Also, it will cost the user too much financially to upload the big volume of data to the server or cloud. In our proposed scheme, the available storage is selected for the preprocessing and storing of lifelog data to reduce performance overhead. Initially, each mobile device temporarily stores personal lifelog data that are sent from diverse sensors. Based on the characteristics of the collected data, the mobile device can send the lifelog data to the server or cloud. In the case of text data with small amounts of data, the collected text data need to be processed and analyzed in mobile devices. This is because the mobile device gives users quick and easy access to the lifelog data by removing the time-delay from the network traffic and the processing overheads. In the case of multimedia data with large amounts of data, otherwise, the collected multimedia data need to be processed and analyzed for extracting lifelog data in server or cloud environment. Therefore, the lifelog data should be transferred into a stable server or cloud when the mobile device is connected to the server or cloud. If all the data was successfully transferred, then they will be removed. A server or cloud can store a lot of lifelog data from multiple mobile device into a lifelog database. If a personalized application needs to analyze or utilize the lifelog data, then it queries the lifelog database to obtain the accumulated lifelog data.

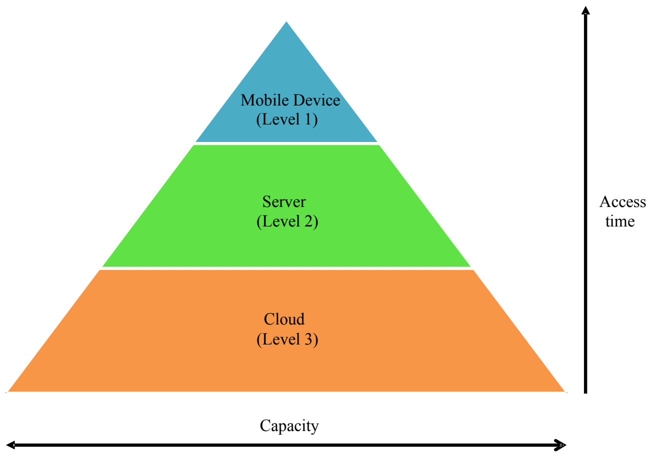

Fig. 2 Storage hierarchy

IV. Performance Analysis

A. Signalling Cost Function

In this section, we compare our proposed scheme called AHLM with the existing hierarchical lifelog management (HLM) scheme[7]. The performance metric is the total signaling cost, which is the sum of the collecting, processing, storing and utilization costs for lifelog management. We define the costs and parameters used for the performance evaluation of location update and packet delivery costs as follows:

- N : The size of collected data per unit time
- λ_e : The occurred event rate per unit time
- λ_r : The retrieval rate per unit time.
- B : The constant weighting factor of transmission delay.
- S : The constant weighting factor of storage overhead.
- C_C : The cost of unit data collection.
- C_{Pi} : The cost of unit data processing i-th level.
- C_{Si} : The cost of unit data storing i-th level.
- C_U : The cost of unit data utilization.
- $T_{C\text{-}Pi}$: Unit data transfer cost from collecting to processing in i-th level.
- $T_{Pi\text{-}Sj}$: Unit data transfer cost from processing in i-th level to storing in j-th level.
- $T_{Si\text{-}U}$: Unit data transfer cost from storing i-th level to utilization.

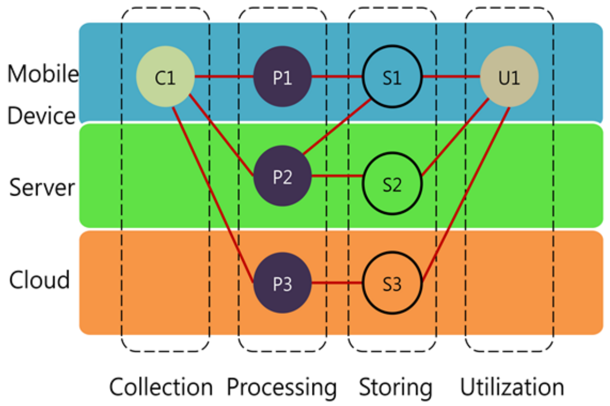

Fig. 3 A hierarchical memory level model

Similar to [4], we define λ_e as the occurred event rate per unit time. It indicates how often raw data are extracted from all lifelog data in the lifelog management system. We define λ_r as the retrieval rate per unit time. Also, we assume that λ_e and λ_r follow the exponential distributions with means $1/\lambda_e$ and $1/\lambda_r$. We define B and S as the constant weighting factors of transmission delay and storage overhead, respectively. We define N as the size of collected data per unit time. The total signal cost in the HLM scheme can be defined as follows:

$$C_T = \begin{bmatrix} \lambda_e \cdot N \cdot (C_c + B \cdot T_{C-P_i} + C_{P_i} + B \cdot T_{P_i-S_i} + S \cdot C_{S_i}) \\ + (1-\lambda_e) \cdot N \cdot (C_c + B \cdot T_{C-P_i} + C_{P_i}) \end{bmatrix} \\ +[\lambda_r \cdot \{ e^{\frac{-\lambda_e}{\lambda_r}} + (1 - e^{\frac{-\lambda_e}{\lambda_r}})(S \cdot C_{S_i} + B \cdot T_{S_i-U} + C_U)\}] \tag{1}$$

In the HLM scheme, all data are processed and stored when an event occurs. In Eq.(1), therefore, the first term is costs for collecting, processing, and storing. The second term is costs for collecting and processing when event does not occur in the HLM scheme. The last term is costs for retrieval as occurrence of retrieval. A hierarchical memory level model is used to analyze the performance of the lifelog system. We describe the data flow including collection(C), processing(P), storing(S) and utilization(U) as depicted in Fig. 3. In the AHLM scheme, the sensed raw data are processed in cloud environments rather than mobile devices because of the large data size. We denote k as the level of hierarchical structure in proposed scheme. Also, we denote T_s as the size of collected lifelog data per unit time in our proposed scheme. For simplicity, we assume that the lifelog data are considered as types of multimedia data when the size of raw data exceeds threshold value threshold T_h Otherwise, lifelog data are considered as the types of text data. To get the value of k, we define the cost difference function from Eq.(1) as follows:

$$\Delta(\lambda_e, \lambda_r, N, T_h) = C_T \tag{2}$$

Given Δ, the algorithm to find the optimal value of k is defined as follows:

$$k = \begin{cases} 1 & , \text{ if } T_s < T_h \\ \text{minimum}(\Delta(\lambda_e, \lambda_r, T_h, N)), & \text{otherwise} \end{cases} \tag{3}$$

The optimal value of k is a designed value. It is computed prior to communications based on the value of λ_r , λ_e , T_h, and N. From Eq.(3), we can define the total signal cost in the AHLM scheme as follows:

$$C'_T = \begin{bmatrix} \lambda_e \cdot N \cdot (C_c + B \cdot T_{C-P_k} + C_{P_k} + B \cdot T_{P_k-S_k} + S \cdot C_{S_k}) \\ + (1-\lambda_e) \cdot N \cdot (C_c + B \cdot T_{C-P_k} + C_{P_k}) \end{bmatrix} \\ +[\lambda_r \cdot \{ e^{\frac{-\lambda_e}{\lambda_r}} + (1 - e^{\frac{-\lambda_e}{\lambda_r}})(S \cdot C_{S_k} + B \cdot T_{S_k-U} + C_U)\}] \tag{4}$$

B. *Numerical Resuls*

In this section, we compare our proposed scheme called AHLM with the

TABLE I
PERFORMANCE ANALYSIS PARAMETERS

Parameter	Value	Parameter	Value
λ_e	0.001~0.1	λ_r	0.1~0.01
N	1~200	C_C	1~10
B	1~100	S	1~100
C_U	1~10	C_P	1~100
C_S	1~100	$T_{C\text{-}P}$	1~10
$T_{P\text{-}S}$	1~1	$T_{S\text{-}U}$	1~10

In this section, we will demonstrate some numerical results. Table I shows some of the parameters used in our performance analysis discussed in [7-12]. There are various types of cases to calculate the total signalling cost for the HLM scheme depending on different computing resources level. Both the HLM and AHLM schemes, the collecting and utilization are only performed by the mobile device. The first case is C_1-P_1-S_1-U_1, in which the preprocessing and storing of lifelog data are performed by the mobile device. The second case is C_1-P_2-S_2-U_1, in which the server performs the intermediate processing and storing. The third case is C_3-P_1-S_3-U_1, in which the cloud performs the intermediate processing and storing. The fourth case is C_1-P_2-S_1-U_1, in which it operates in exactly the same manner as in the first case except that the server performs the intermediate processing. Finally, the final fifth case is C_1-P_3-S_1-U_1, in which it operates in exactly the same manner as in the first case except that the cloud performs the intermediate processing. Fig. 4 shows the optimal level k with increasing of the size of collected data per unit time N for λ_r = 0.1, and λ_e = 0.5. This figure shows the size of collected data per unit time N increases, the optimal level k also increases.

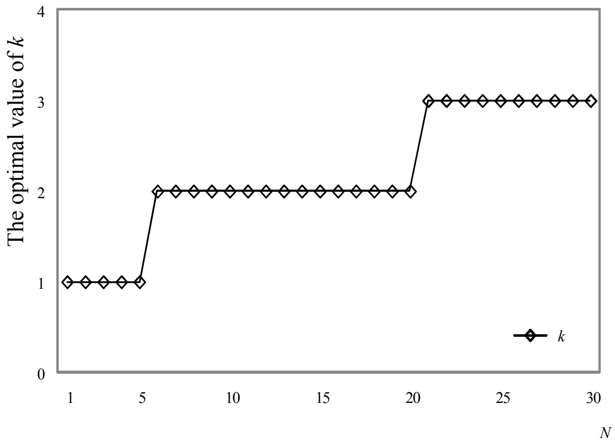

Fig 4. Effect of N on the optimal value of k

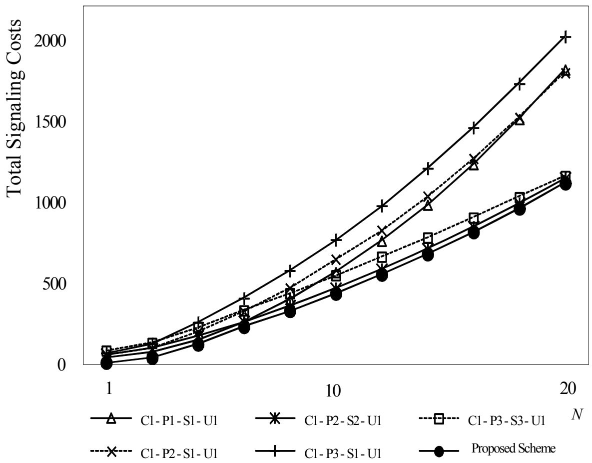

Fig 5. Effect of N on the total signalling cost

This is because when N is large, the processing and storing costs are high. Using the analytical result as shown in Fig. 4, it is possible to get the necessary optimal level k when N is given. For example, when N is between 6 and 20, the optimal level k is 2. We demonstrate the performance comparison between the HLM and AHLM schemes. Fig. 5 shows the effect of the size of collected data per unit time N for $\lambda_r = 0.5$, $T_h = 6$, and $\lambda_e = 0.1$. As shown the Fig 5, the total signalling cost increases as the size of collected data per unit time N increases. It is obvious that the whole performance of the AHLM scheme is the best. For large values of N, the performance of the AHLM scheme is better than that of the HLM scheme. This can be explained by the fact that the AHLM scheme reduces the total signalling cost efficiently by selecting the usage storage for the processing and storing. Fig. 6 shows the effect of the occurred event rate per unit time λ_e for $\lambda_r = 0.5$, $T_h = 6$, and $N = 100$. As shown the Fig. 6, the total signalling cost increases as the occurred event rate per unit time λ_e increases. In this figure, we observe that the AHLM scheme always performs better than the HLM scheme. Fig. 7 shows the effect of the retrieval rate per unit time λ_r for $\lambda_e = 0.5$, $T_h = 6$, and $N = 20$. As shown in Fig. 7, the total signalling cost for the HLM and AHLM schemes have a linear relationship with the effect of the retrieval rate per unit time λ_r. It can be observed that the total signalling cost of the AHLM scheme can be much lower than that of the HLM scheme.

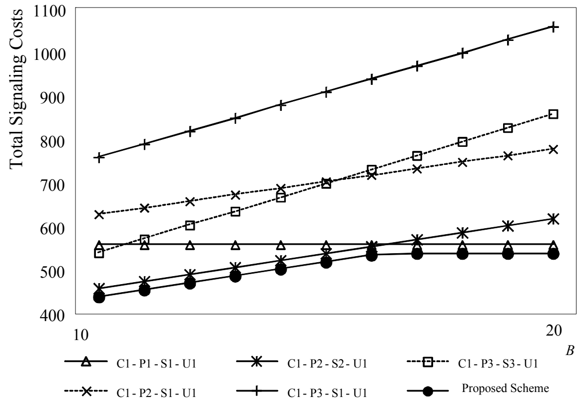

Fig 6. Effect of λ_e on the total signalling cost

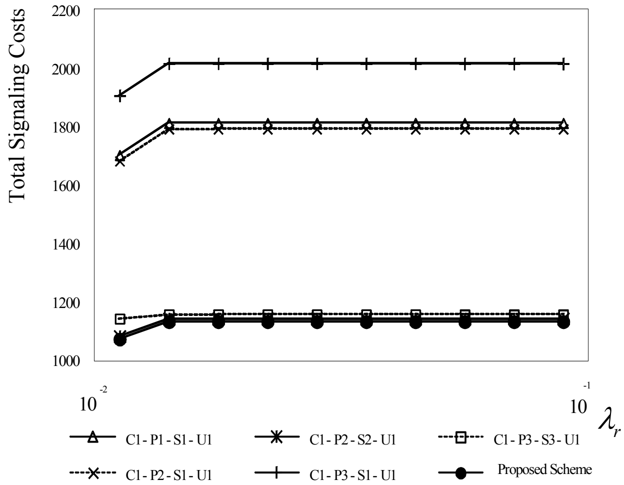

Fig 7. Effect of λ_r on the total signalling cost

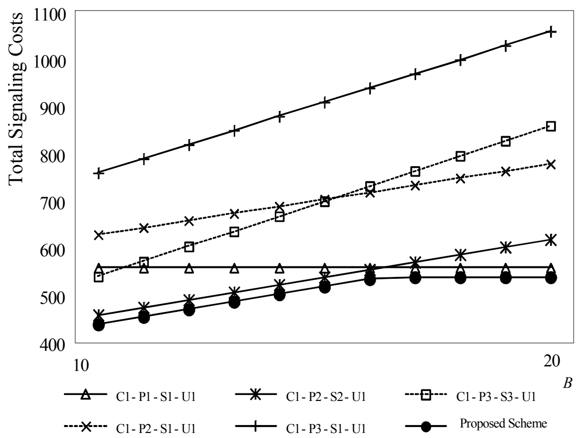

Fig 8. Effect of B on the total signalling cost

Fig. 8 shows the effect of the transmission delay B for $\lambda_r = 0.1$, $\lambda_e = 0.5$, $T_h = 6$, and $N = 10$. As shown the Fig. 8, the total signaling cost increases as the transmission delay B increases. We can see in this figure that our AHLM scheme generates lower signalling cost for lifelog management than the HLM scheme. Fig. 9 shows how the total signalling cost are affected by the storage cost S for $\lambda_r = 0.1$, $\lambda_e = 0.5$, $T_h = 6$, and $N = 10$. We can observe that our AHLM scheme can reduce the signalling cost. Particularly, when the value of S is high, the benefit of the AHLM scheme is significant. Based on the above analysis, we find that our proposed scheme can achieve superior performance, much better than that of the existing HLM scheme.

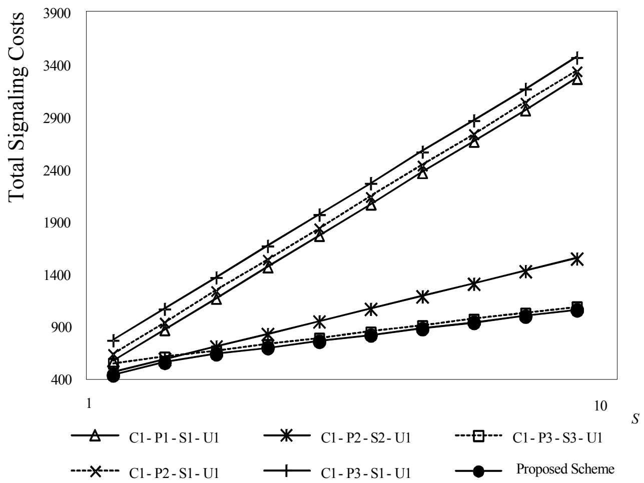

Fig 9. Effect of S on the total signalling cost

V. CONCLUSIONS

In order to achieve high-quality screening and treatment for the healthcare service sector, all lifelog data should be managed effectively because it is necessary to gather as much health information on users as possible. To address these problems, we propose a novel lifelog management scheme based on the characteristics of collected data. In our proposed scheme, the available storage is selected for the preprocessing and storing of lifelog data to reduce performance overhead. The cost analysis presented in this paper shows that our proposed scheme is superior to the existing lifelog scheme in terms of performance by selecting the storage resources.

Acknowledgement

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.R0166-15-1007, Development of the Standards at Integrated Personal Health Information and Interoperability Technology)

REFERENCES

[1] Jim Gemmellm, Gordon Bell, Roger Lueder, "MyLifeBits: a personal database for everything', Magazine Communications of the ACM- Personal information management, Volume 49 Issue 1, pp. 88—95, 2006.

[2] Reza Rawassizadeh, Martin Tomitsch, Katarzyna Wac, A. Min Tjoa, "UbiqLog: a generic mobile phone-based life-log framework. Personal and Ubiquitous Computing 17(4), pp. 621-637, 2013

[3] Belimpasakis. P. , Roimela. K. , Yu You: Experience Explorer" A Life-Logging PlatformBased on Mobile Context Collection", Third International Conference on Next Generation Mobile Applications, Services and Technologies, pp. 77-82, 2009

[4] Snehal Chennuru, Peng-Wen Chen, Jiang Zhu, Joy Ying Zhang "Mobile Lifelogger – Recording, Indexing, and Understanding a Mobile User's Life, ", Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Volume 76, pp. 263-281, 2012

[5] Prananto, B.H, Ig-Jae Kim, Hyoung-Gon Kim. 2007. Multi-level Experience Retrieval for the Personal Lifelog Media System. Third International IEEE Conference on Signal-Image Technologies and Internet-Based System : 175-182

[6] Snehal Chennuru, Peng-Wen Chen, Jiang Zhu, Joy Ying Zhang, " Mobile Lifelogger : Recording, Indexing, and Understanding a Mobile User's Life. Lecture Notes of the Institute for Computer Sciences, Vol. 76, pp. 263-281, 2012

[7] Minkyung Kim, Dong-Wook Lee, Kangseok Kim, Jai-Hoon Kim," Hierarchical structured data logging system for effective lifelog management in ubiquitous environment", Multimedia Tools and Applications, 2013.

[8] Svebor Karaman, Jenny Benois-Pineau, Vladislavs Dovgalecs, Remi Megret, Julien Pinquier, Regine Andre-Obrecht, Yann Gaestel, Jean-Franeois Dartigues, "Hierarchical Hidden Markov Model in detecting activities of daily living in wearable videos for studies of dementia", Multimedia Tools and Applications 69(3), pp. 743-771, 2012

[9] Ig-Jae Kim, Sang Chul Ahn, Hyoung-Gon Kim, "Personalized lifelog media system in ubiquitous environment. Lecture Notes in Computer Science Volume (4412), pp. 20-29, 2007

[10] Ovgu Ozturk, Tomoaki Matsunami, Yasuhiro Suzuki, Toshihiko Yamasaki, Kiyoharu Aizawa. "Real-time tracking of humans and visualization of their future footsteps in public indoorenvironments", Multimedia Tools and Applications 59(1), pp. 65-88, 2011

[11] Ovgu Ozturk, Tomoaki Matsunami, Yasuhiro Suzuki, Toshihiko Yamasaki, Kiyoharu Aizawa. "Real-time tracking of humans and visualization of their future footsteps in public indoorenvironments", Multimedia Tools and Applications 59(1), pp. 65-88, 2011

[12] Chang-Woo Song, Daesung Lee, Kyung-Yong Chung, Kee-Wook Rim, Jung-Hyun Lee, "Interactive middleware architecture for lifelog based context awareness", Multimedia Tools and Applications 71(2), pp. 813-826, 2013