Social Choice Considerations in Cloud-Assisted WBAN Architecture for Post-Disaster Healthcare: Data Aggregation and Channelization

Sudip Misra, Subarna Chatterjee
School of Information Technology
Takshashila Building
Indian Institute of Technology, Kharagpur
Kharagpur, West Bengal, 721302, India
Email: smisra@sit.iitkgp.ernet.in, chatterjeesubarna@yahoo.com
Contact number: +91-3222-282338, +919432827620

Abstract

In a cloud-assisted Wireless Body Area Network (WBAN) architecture, for post-disaster medical relief operations the health-data of patients equipped with body sensors are transmitted from the Local Data Processing Units (LDPUs) to the health-cloud through a set of mobile monitoring nodes. In such a scenario, this paper focuses on two fundamental research issues – aggregation of health data from the LDPUs within the monitoring nodes, and data channelization by dynamic selection of cloud gateways. While the existing literature mainly center around data aggregation among body sensors, our focus is on data aggregation among LDPUs at the monitoring nodes, which makes the problem challenging due to reasons attributed to mobility and health data prioritization. The proposed solution generates a preference order for each patient, based on his/her severity and acuteness. For each patient, an *Exigency Factor*, which measures health criticality, is associated. Additionally, a pseudo-cluster based "fair" aggregation policy is proposed on the basis of the *theory of Social Choice*. Cloud gateways are also dynamically allocated to channelize the prioritized and aggregated health data in a "fair" manner. Simulation results illustrate that the proposed pseudo-cluster based aggregation scheme results in improved cloud-assisted WBAN architecture when compared with cluster-based, tree-based, and structure-free aggregation methods, in terms of reliability of node selection, number of packets transmitted, redundancy during transmission, and probability of congestion. The proposed health data channelization scheme also demonstrates that the choice of packets to be transmitted to the cloud is biased on data criticality, and the selection of gateway is biased on gateway capacity and reduced communication cost.

Index Terms

Wireless Body Area Network (WBAN), Cloud Computing, Social Choice, Arrow's Impossibility Theorem, Clustering, Data aggregation

1. Introduction

The world witnesses many indomitable disasters, natural or human-induced, such as floods, tsunamis, and terrorist bomb blasts, which take a deep toll upon the society and mankind. The deadly consequences lead to enormous agony and distress among the victims. Contemporary relief operations are generally active and prompt, but are insufficient in managing large disasters with large numbers of injured or severely wounded victims. It becomes increasingly difficult for the healthcare units to diagnose, monitor, and provide medical facilities to the huge mass of affected patients. The aftermath of a disaster typically leads to utter confusion and mismanagement. One of the main reasons behind such anarchy is that the healthcare units function as individual entities in isolation. A more collaborative and cooperative approach, in respect of both technology and management, would improvise such situations of medical emergency.

In this paper, we look at some of the technological prospectives to mitigate the post-disaster healthcare problems and organize the efforts of the medical teams in a post-disaster scenario. Our work considers disaster management using a cloud-assisted Wireless Body Area Network (WBAN) environment. Specifically, our system model considers WBAN-equipped patients in the lower layer of hierarchy communicating over a multi-tiered architecture with a cloud. The rationale behind the choice of such a model is discussed in the paper.

WBAN is evolving as a promising healthcare technological solution in the recent times [15], [38], [4]. It has found widespread admissibility in many application domains such as battlefield, disaster healthcare, and biomedical applications [23], [21], [8]. Conventionally, in a WBAN, nodes are embedded with sensors that are capable of sensing and monitoring physiological attributes of a human body such as heart rate, body temperature, blood pressure, oxygen saturation level, and respiratory rate. These nodes can be mounted on the human body or implanted within, thereby forming a scattered network topology. Subsequent computations are performed on a local data processing unit (LDPU), which is mapped to several body sensor nodes.

In the proposed cloud-assisted WBAN platform, the on-body sensors of each patient communicate with the LDPU. Generally, LDPUs, in turn, communicate with a base station (BS). A BS might be several hops away from a patient resulting in unwanted transmission delay or communication difficulties. This can subsequently have fatal and distressful implications on the affected victims. However, in emergency scenarios, the process of monitoring and tracking a patient's health should be eminently prioritized, and, hence, the health data transmissions should be well optimized. Moreover, to organize the endeavor of the medical teams in a methodical manner, we design to integrate the health centers with a health-could platform. This facilitates concurrence and cooperative functioning of the individual centers as an extensive healthcare unit, thereby rendering better medical support and service.

It is difficult to have each LDPU communicate with a BS, specially in medical emergency situations. Such type of communication not only increases the cost, but also exhausts the battery of a body sensor and introduces latency in transmissions. It is important to mention that the nodes are responsible for carrying data of post-disaster critically ailing patients. So, while communicating such critical data, we must ensure to minimize cost and delay and also maximize the throughput of healthcare. So, the set of LDPUs of different patients must be partitioned into several subsets. Corresponding to every subset, an aggregated form of data is transmitted to an intermediate entity or a mobile monitoring node. These monitoring nodes communicate with the cloud via cloud gateways. Since these monitoring nodes are mobile, an *additional problem* is to map the cloud gateway a mobile node must communicate with. The data from the cloud may be further transmitted to a board of online doctors or medical experts, who are responsible for the management of post-disaster medical relief operations.

The proposed work focuses on three aspects:

(i) Dynamic formation of subsets of LDPUs for each monitoring node.
(ii) Performing loss-less data aggregation on behalf of every subset.
(iii) Establishing a communication map of every monitoring node to a cloud gateway.

The first two aspects are fundamentally important from a doctor's perspective. Lossless data of patients should be available to the medical expert or doctors to undertake further analysis. It is essential to ascertain the proper choice of LDPUs that will lead to the establishment of a group that communicates with the monitoring node with minimal cost and delay. Moreover, the aggregation mechanism should be sensitive to health criticality. Consequently, the data of a critically injured patient should not be lost as a result of aggregation. The third aspect channelizes the health data and manages the data traffic to alleviate congestion. It tries to maintain an overall optimality in the process of communication. Health data transmission and its subsequent analysis can be overlaid on the cloud environment. In the architecture, the health-cloud renders RAAS (Resource As A Service) by providing dynamic monitoring and medical diagnosis facilities to the patients.

The aforesaid concerns not only insist on making a critical patient's data readily available to medical experts, but also channelizes and improvises the overall network performance. In this paper, we propound the *Body Area Network Data Aggregation Algorithm* (*Banag*)[1] and *Optimal Channelization Algorithm* (*OCA*), which attempts to address the aforesaid issues in a rational way. The proposed algorithms are anchored in the *theory of Social Choice* [7].

A. Motivation

Dynamic data aggregation is a very important issue in the proposed system architecture. Based on the choice of the LDPUs that form a group, application performance can improved. We select (or

[1]Incidentally, the origin of the word Banag is from a village in the Tibet Autonomous Region of China.

ignore) an LDPU to be inside (or outside) a group based on an aggregation function only. A proper aggregation function necessarily needs to be "fair", so that none of the eligible nodes are ignored from consideration. Motivated by the generalized criteria of *fairness* in the social welfare literature [32], the following properties are postulated in this paper:

(i) *Majority*: An aggregation function violates the majority criterion, if some node has a majority of the first place preferences, but eventually after aggregation, it remains overlooked.

(ii) *Condorcet*: An aggregation function violates the condorcet criterion, if some node is preferred constantly against every other candidate, but ends up not being the winner of the selection.

(iii) *Irrelevant Alternatives*: An aggregation function violates the irrelevant alternatives criterion, if having a loser node drop out of the race, changes the winner of the selection.

(iv) *Monotonicity*: An aggregation function violates the monotonicity criterion, if one can transform the winner into a loser by moving the winner up the preference list on some of the individual preferences.

An aggregation function is expected to be "fair", so that the outcome of aggregation is consistent with the individual preference of each LDPU. Furthermore, many LDPUs might get wrongly grouped, if the fairness criteria are not conformed with. Not only aggregation or clustering matters, but so does the method of choosing a winner.

Following the aggregation, the data is transmitted to the cloud platform for subsequent analyses. However, in a post-disaster environment, it is required to monitor patients' health criticality remotely. This includes ambulatory healthcare services where the health status of a patient is examined continuously over the time when the patient is being moved to the emergency healthcare center. Therefore, the gateway through which the health data is transmitted to the cloud changes along with the global position of the patient. It is important to select gateway that not only has a capacity of forwarding the health data but also incurs a minimal energy expenditure while communication with the LDPU of the patient. Herein comes the significance of data channelization. If the data is not properly channelized, it causes some gateways to be over-loaded, and some to remain idle. Data passing through over-loaded gateways may introduce unnecessary delay due to queuing within a gateway and also increase communication cost significantly.

B. Contribution

The importance of social choice is perceptible. We discuss and analyze how Social Choice theory helps to improve fairness, accuracy, and energy efficiency from a system point of view. The contributions of our work are as follows:

- The work focuses on pseudo-cluster formation so that the aggregation is not biased on leader nodes. Each patient is considered to be a member of a democratic society. This provides opportunity to every patient to transmit its data to LDPU. However, data from LDPU is processed and analyzed to a superficial extent before further transmission to the health-cloud.
- The data aggregation among the LDPUs is done in a "fair" way as discussed above in Section A. The health-data packets are organized based on the acuteness or the severity of the disaster affected patients. A health-based priority is established based on which the data is transmitted to the medical teams in an ordered manner.
- The aggregation mechanism also considers mobile centers of aggregation, thereby increasing the flexibility and scalability of the system. Load-sharing with optimal packet transmissions are ensured to speed-up the healthcare operation.
- After data aggregation, dynamic gateways are allocated. Even at this stage, prioritization of packets still continues, i.e., nodes with the most critical data are attended first. Gateways are allocated to nodes by optimizing the communication cost, health-data criticality, and gateway capacity.

C. Organization of the paper

The rest of the paper is organized as follows. Section 2 contains a discussion on the related work. Section 3 contains the details of the architectural design. Sections 4 and Section 5 illustrate the workflow of the proposed algorithms. Section 6 presents some examples that illustrate the proposed solution. Section 7 illustrates the results of simulation. Finally, Section 8 concludes the work with discussions about how it can be enhanced in the future.

2. Related Work

In this Section, we present the related research works in the domain of WBAN-cloud, data aggregation and gateway selection. Health monitoring using WBANs has found widespread application in recent times [3], [11], [9]. Lin et al.[26] proposed a privacy preserving scheme for cloud-assisted health monitoring. It mainly focuses on cryptographic issues of data, and reduces computational complexity in key management. A Telemdedicine Repository based on cloud is proposed in [37]. The system extracts health information from patients, and accordingly renders medical suggestions. However, it does not rank patients based on their health criticality. Such a system cannot be applied in a post-disaster like critical scenario because in such cases non-prioritized transmission of enormous data might lead to network failure. In another work, Rashid et al. [35] proposed a cloud-based platform for ubiquitous monitoring. It allows users to interface through their system, thereby rendering remote healthcare monitoring. It is merely an application development work, not looking into the network difficulties during massive medical-relief operations. An endeavor of building up a mobile health record [13] system wsa made. However, this work does not consider online patient monitoring intricacies during emergency scenarios. Our work implements on a collaborative and cooperative functionality using cloud computing. The health data is transmitted to the cloud, and stored there for online remote monitoring by medical experts.

Data aggregation is a well excavated research sphere, and has found its applications in several discrete domains in the recent past. However, not many research works have focused on data aggregation in the context of WBANs, which have specific requirements. In [14], the authors have proposed a data aggregation mechanism for Body Sensor Networks (BSNs), in which an algorithm aims to aggregate nodal data within the network prior to transmission is designed. The idea thrives on the principle of transmission of a single large aggregated data through longer paths, instead of multiple transmissions involving small data fragments through shorter paths. The algorithm reduces the communication cost, which, in turn, optimizes power consumption. However, in our work, the system aggregates data from the LDPUs, instead of from the body sensors, and involves processing the raw data sensed by the on-body sensors.

Most other existing works on data aggregation consider general terrestrial sensor networks. In general, data aggregation is sensor network can be executed using the standard approaches such as cluster-based [42], tree-based [41] or structure-free [12] algorithms. Cluster-based techniques are the most popular, and widely used. In [1], the authors have propounded a hierarchical cluster-based data aggregation algorithm that analyzes the spatial and temporal behavior of sensor nodes. It identifies the corrupted nodes, and accordingly extracts the data in the raw form from the misbehaving nodes, thereby, improving energy efficiency and bandwidth utilization. Another cluster-based aggregation scheme [39] considers mobility of nodes during cluster formation. The proposed algorithm minimizes the communication delay within a cluster, and, thus, maintains cluster stability. The system has found its applications in urban regions, since it can easily be configured to monitor human movements.

An energy-efficient data aggregation approach is proposed in [34]. However, in this approach, the cluster-heads transmit the data directly to the base station. This is a severely energy exhaustive process, as the cluster-heads might be several hops away from the base station. In [27], cluster-based aggregation in under-water acoustic sensor networks is examined. The proposed algorithm designed ensures an optimal usage of underwater resources such as bandwidth, energy, and network lifetime. Similar approaches are adopted in [29], [31], [30] for acoustic environments. However, none of the above mentioned clustering approaches are socially “fair”. In these clustering approaches, the clusters are governed by one or more cluster-heads. This introduces dictatorship, and, thus, the algorithms fail to act equally among the cluster peers. Also, unanimity is not maintained in a cluster based approach since not all the nodes are allowed to have decisive potential. In our work, a “fair” aggregation algorithm is designed, that considers each LDPU as a member of a democratic society.

Few works such as [22] have addressed the problems of clustering by introducing a hybrid approach. It primarily addresses the problem of handling voluminous data. But the underlying mechanism is suitable for use in the target tracking domain in general wireless sensor networks. [17] focuses on reducing the number of communication edges in tree-based aggregation. It proposes intelligent water drop (IWD) algorithm to generate an optimal tree. However, it has serious loophole in the fact that the destination aggregation node is found with some probability. This may have serious effects on a sensor network with irregular topology. In [5], a tree-based aggregation approach is proposed, in which the root of the tree provides data as and when queried. This prevents flooding of packets and subsequent loss of energy. In [41] as well, a tree-based aggregation scheme is proposed to reduce operational complexity and the number of relay nodes. However, it involves repetitive packet transmissions, thereby deteriorating the overall network performance. In general, tree-based approaches tend to increase the height of the tree to achieve greater modularity. For example, in [18], a tree-based aggregation is followed. Data from the lower level of the tree are aggregated up to the root. Although it results in faster response, it reduces the number of leaf nodes and degrades the system performance, because the leaves are principally responsible for the sensing operation. For WBANs, it should be always remembered the LDPUs are specific to a human. It has to maintain its battery upto threshold level in order to be able to transmit health data of a patient.

In addition to the cluster-based and tree-based approaches, some structure-free approaches also exist. These methods face a major challenge in spatial and temporal convergence. Since these algorithms are not based on an explicit structure, it is quite difficult to transmit packets from multiple sources to a common destination in a time-synchronized manner. In [12], a structure-free aggregation scheme is proposed. Although the algorithm does not involve significant overhead due to formation of a tree or a cluster, it involves several packet broadcasts and transmissions. This reduces the lifetime of the nodes and eventually the network lifetime. So, structure-free aggregation is not suitable for use in our architecture.

Overall, we see that the existing techniques do not conform to the general "fairness" criteria as discussed in Section 1. Our work combines the advantages of the different types of aggregation and produces a pseudo-cluster based method that conforms to the "fairness" criteria of Social Choice.

Another issue that our work focuses on is data channelization. Few works [46], [45] have focused on an appropriate scheduling of cloud service providers. The works follow a blind scheduling algorithm and are independent of the demand and service rate of the service providers. However, for fair and balanced allocation of gateways, it is quite significant to have a prior knowledge of demand and capacity of the cloud-gateways to prevent traffic congestion, overloading of servers, and unwanted communication delays. [40] proposed a new gateway selection scheme to control traffic at the entry of the gateway nodes. It also concentrates on load balancing by designing a cost function that computes the intermediate hops while communication. [19] addresses the same problem of gateway allocation among self-organizing nodes in a mesh networks. It tries to optimize the intermediate number of hops and thus, achieves a gain in throughput of the network. Additional work such as, [6], [28], [20] also exist for gateway selection. But none of them are applicable in WBANs as none of them address the problem of prioritizing data. As we are working with health-data of severely ailing patients, it becomes necessary to transmit a highly critical data before a normal one.

Our work optimizes the overall healthcare operation in a post-disaster scenario by a cloud-assisted WBAN architecture. Urgency of data is introduced in this system. Moreover the algorithms conform to the *theory of Social Choice* and hence exhibit "fairness". Fairness and health-data criticality add a new dimension to the system of data aggregation and channelization.

3. System Architecture

The proposed system architecture is three-tier, as shown in Fig. 1. It has the following actors: *Local Data Processing Unit* (*LDPU*), *Monitoring Node* (*MN*), and *Cloud Gateways* (*G*). Additionally, there are two intermediate processing units I_1 and I_2.

The sensor nodes are mounted on or within a human body. These nodes sense the physiological attributes of a patient and transmit those to a LDPU, which is placed on the body of the patient itself. We assume that the LDPUs are always within the range of multiple *mobile* monitoring nodes. The mobility of these nodes requires that each LDPU is covered by at least a single monitoring node. Several LDPUs of patients choose to form a group. The selection of a monitoring node is performed on I_1. After the selection and

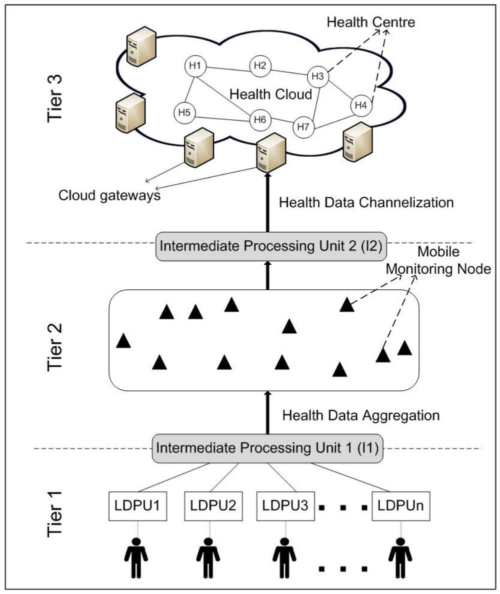


Fig. 1: WBAN Health-cloud based architecture

formation of a group on behalf of every monitoring node, the members of each group transmit their data to the selected monitoring node. This ensures load distribution among all the monitoring nodes. These monitoring nodes aggregate the data obtained from multiple LDPUs. The aggregated data from each monitoring node gets transmitted to the cloud through the cloud gateways. Each monitoring node chooses its own gateway so that a single gateway is not over-loaded.

The traditional problem of data aggregation [14], [43], [10] spans merely between sensors, i.e., at the lowest tier of our architecture. Obviously, our work is also applicable in Tier 1. However, in this paper we focus on aggregation in the middle layer, i.e., Tier 2, as it makes the problem interesting in a post-disaster medical relief scenario. In such a scenario, it is expected that a huge number of victims and patients will try to communicate to the health-cloud at the same time. Communication within Tier 1 is specific to a single patient or a victim. So, chances of congestion and subsequent collisions are more pronounced in Tier 2 than in Tier 1. A large volume of packet transmissions and broadcasts happen in Tier 2. The uncontrolled traffic consequently leads to chaos in the network channels. Such a scenario incurs superfluous delay. Also, the body sensors are severely vulnerable to exhaustion of energy due to redundant packet transmissions and recurring back-offs. In addition to the aforesaid problems, Tier 2 consists of mobile nodes. So, based on the coordinates of the monitoring nodes, they can be dynamically selected for participating in aggregation. As the data aggregation problem at the middle tier is more challenging, the present paper focuses on it only. Further, after proper aggregation, we try to control the traffic by proper channelization through cloud gateways, so that no gateway is over-loaded within health packets resulting in a similar situation as just mentioned.

Following the architectural aspects, we briefly illustrate the conditions that are necessary to be signified while ensuring compliance with Social Choice. The applications of the *theory of Social Choice* primarily necessitates combining individual preferences in a group, so that a collective decision can be reached while ensuring optimal fairness criteria. Integral to the *theory of Social Choice* is *Arrow's Impossibility Theorem* [36], which states the following:

"*If the number of choosers is finite and there are at least three candidates, no aggregation method can simultaneously satisfy universality, transitivity, unanimity, independence and non-dictatorship*".

In the purview of our present work, the above indicates that a good decision, election, clustering, or aggregation algorithm in a cloud-assisted WBAN scenario satisfies the following properties.

(a) *Non-Dictatorship*: The existing clustering algorithms render the cluster-heads to be the dictators, while deciding the set of nodes that will be (need to be) activated [24], [2]. Hence it is not a "democratic" approach in the sense that equality is not maintained among all the nodes and not every node participates in decision making. This is attributed to the hierarchical arrangement of nodes. It may be mentioned at this juncture that in order to maintain non-dictatorship, we take a simple *de tour* around normal cluster formation. This is discussed in Section 4.

(b) *Universality*: The leader (which normally votes/decides) ranks the candidates in order of increasing distance between themselves and the candidates. The preferences of the deciders are then single-peaked. This can introduce inaccuracy while tracking or monitoring because of the inherent inconsistency, and inaccuracy in sensor measurements that are likely to happen in a WBAN. It is evident that mere dependence on sensor readings is insufficient. In general, a group of voters, agents, or deciders have single-peaked preferences over a group of outcomes if: (i) they each have an ideal choice in the set; and (ii) outcomes that deviate from their ideal choice are strictly less preferred. To prevent single-peaked preferences, this work considers the choice of a winner, based on the sensor readings of the society of monitoring nodes.

(c) *Unanimity*: The outcome of the aggregation/clustering method should not contradict the individual heads, when they decide unanimously.

(d) *Transitivity*: The outcome of the aggregation method must always be a complete ranking, possibly with ties. This may not happen all the time, especially if the nodes have a cyclic choice.

(e) *Independence*: The preference between two alternatives should be dependent only on the individual preferences associated with those alternatives and independent of the others.

In addition to using Arrow's Impossibility Theorem for data aggregation, at I_1, we also implement the above points while channelizing data through the gateways, so that fairness in gateway allocation can also be ensured.

4. Banag: The Body Area Network Data Aggregation Algorithm

In this Section, we present the detailed description of the proposed *Banag* algorithm, and the corresponding implementation of the *theory of Social Choice* in this respect.

Banag deals with pseudo-cluster formation and data aggregation. Initially, based on some criteria, a subset of LDPUs forms clusters without Cluster Heads (CHs). Evidently, the formation of a cluster is totally unbiased and non-sovereign. Since no leader exists during cluster formation in *Banag*, the subset of LDPUs do not behave like a typical cluster [16], [25]. Hence, we refer to such a cluster as a "Pseudo-Cluster", the existence of which makes the problem further challenging.

We consider the LDPUs stationed in the individual patient bodies. We have a set of LDPUs (referred to, generally, as nodes) $N = \{0, 1, ..., n\}$ and a set of *Monitoring Nodes* $M = \{m_0, m_1, ..., m_m\}, m \ll n$. We assume a WBAN where every node is capable of estimating the coordinate of a mobile *MN*, once the *MN* is detected. We assume that the time varying position coordinates of each patient is known and stored within the node's local memory, along with the location information of the other existing nodes. So, we also have two symmetric distance matrices X(1..n) and Y(1..n) that contain global coordinates of every patient. The coordinate of a node *i* is given as:

$$C(i) = (X[i], Y[i]) \tag{1}$$

Prior to using the *theory of Social Choice*, it is needed to identify the set of outcomes. In our case, every node has its own Individual Preference String (IPS). Each IPS is an *ordering* of all the *MN*s. Let the number of nodes, and the number of *MN*s, at a particular time instant t, be n and m, respectively. Then, there will be a total of n number of IPSs, each of length m.

We denote $A_{m_1,t}$ to be the set of nodes monitored by m_1 at time t. Then, the set of outcomes, $A_{m_1,t+1}$, is a non-empty subset of all possible permutations of m monitoring nodes, i.e., $A_{m_1,t+1} \subseteq M$. The elements of M are also termed as alternatives. The alternatives are responsible for turning m_i as the winner.

All IPSs are then mapped to a single collective preference set, i.e., the Pseudo-Clusters, each of which is a collective *ordering* of *MN*s generated by respecting individual node preferences. If a Pseudo-Cluster can be generated over the set M by combining all IPSs, then we can obtain a “fair” outcome. The final surviving members (*MN*s) of a pseudo-cluster will be members of $A_{m_1,t+1}$. In our model, we represent every IPS of an active node *i* as a preference relation which is an *ordering* R_i defined as:

A preference relation R_i of a node *i* is referred to as the ordering over any set *A*, if it satisfies the following properties [33]:

- *Completeness*: For all $x, y \in A$, either xR_iy or xR_iy must hold true.
- *Reflexivity*: For all x $\in A$, xR_ix must hold true.
- *Transitivity*: For all x,y,z $\in A$, xR_iy, $yR_iz \Rightarrow xR_iz$.

The notion or understanding of $v_1R_iv_2$ signifies that an active node *i* weakly prefers v_1 over v_2, i.e., according to node *i*, v_1 is at least as good as v_2. A strict preference relation P_i is also defined and coupled with R_i. P_i is the asymmetric component of R_i. The notion of $v_1P_iv_2$ indicates that only $v_1R_iv_2$ holds true and $v_2R_iv_i$ does not hold strictly, i.e., according to node *i*, v_1 is strictly better than v_2. Another association with R_i is its symmetric component I_i, where $v_1I_iv_2$ holds true *iff* both $v_1R_iv_2$ and $v_2R_iv_i$ hold true, i.e., node *i* is indifferent over v_1 and v_2. For example, say the number of active nodes in a network be 5,and one denoted as v_1,..v_5. The IPS of v_1 is represented as $v_2R_1v_1R_1v_4R_1v_3R_1v_5$. This implies that node v_1 prefers v_2 the most and v_5 the least. This entire preference string for node v_1 at time t is denoted by $R_{1,t}$. We define $R_t = (R_{1,t}, R_{2,t}, ..., R_{n,t})$ as a preference profile over set A_t, which is the set containing IPS of every active node at time *t*. We now theoretically characterize the proposed Social Choice based aggregation scheme.

Definition 4.1. A preference domain, $\mathcal{P}$, is a non-empty set of potential preference profiles. Since m is the number of *MN*s, R_t^m is the set of all preference profiles at time t. So, we have the following:

$$\mathcal{P} \subseteq 2^{R_t^m} \setminus \Phi$$

Definition 4.2. If preferences are complete and transitive, they can be expressed as an *ordering*, or a list of best-to-worst alternatives. However, if the preference list contains indifference between one or more pairs, then the preference can be expressed as a *weak ordering* of alternatives. For example, the ordering and weak ordering over any set A is expressed as $O(A)$ and $WO(A)$, respectively.

Definition 4.3. A social choice function (SCF) is defined as a mapping f : $WO(A)^m \rightarrow 2^A \setminus \Phi$. Thus, n weak orders over set A is aggregated to a non-empty subset of A.

Having defined a preference domain $\mathcal{P}$, an ordering $O(A)$, a weak ordering $WO(A)$, and a social choice function f, we propose a theorem that complies with the notions of the *Theory of Social Choice*.

Theorem 4.1. *Each element of the set R_t is an ordering over set A_t, because $R_{i,t}$ consists of complete and transitive preferences, $\forall i \in A_t$.*

Proof: For the sake of maintaining simplicity in this proof, we use the symbolic notations instead of the alphabetic ones, while indicating strict preferences. Since we need to prove the elements of the set R_t as *ordering*, we consider only the strict preferences (and not the indifferences). We denote the strict preferences by the symbol $\succ$. The notation corresponding to weak preferences remains the same, as discussed earlier.

We have to prove that $R_{i,t}$ has *complete* and *transitive* preferences. The proof of *completeness* is trivial. As mentioned before, every node has to express its preferences about every member of A_t, and the length of every IPS is n_{act}. A preference should always exist between any pair of nodes. So, for every IPS, we have,

$$a \succ b \vee b \succ a, \forall a, b \in A_t \Rightarrow aR_ib \vee bR_ia, \forall a, b \in A_t \tag{2}$$

Hence, $R_{i,t}$ consists of *complete* preferences.

Now, to prove the *transitivity* of preferences, we assume $n = 3$, and $A_t = \{x, y, z\}$. Let us assume that node x prefers itself to y, and y to z, as illustrated in Fig. 2a.

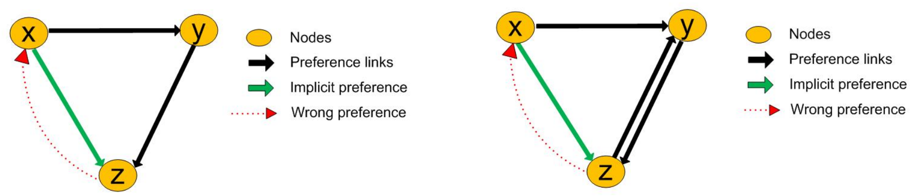
We have,

$$R_{x,t} : x \succ y \succ z \tag{3}$$

Evidently, x prefers z the least, i.e., z is the worst node according to x. Hence, x prefers itself more than z. So, $x \succ z$. If z was preferred to x, it would be a cyclic preference, which is logically unacceptable. In fact, looping within preferences of any node simply prevents us from determining that node's favorite alternative. Hence, we conclude that *transitivity* is inevitable for a logical preference. We note that $R_{i,t}$ consists of *transitive* preferences. This proves that $R_{i,t}$ is an *ordering*.

Theorem 4.2. *Each element of the set R_t is a weak order over set A_t because $R_{i,t}$ may contain indifferences in addition to strict preferences, $\forall i \in A_t$.*

Proof: In the proof of Theorem 4.1, we have already established that elements of R_t are ordering(s) over



(a) Transitivity of preferences for node x

(b) Transitivity of preferences containing indifference

Fig. 2: Transitivity analysis

set A_t. As stated in Definition 4.2, we find that *weak* preferences exist only when there is an indifference between one or more pairs of nodes. We need to prove that $R_{i,t}$ has *complete* and *transitive* preferences, even in the presence of indifferences.

As already mentioned, whenever a node fails to prefer between two or more nodes, i.e., whenever there is a tie, the node is said to posses an indifference for all those nodes. In this case, every node has to express either a preference or an indifference between every possible pair of nodes in A_t. So, for every IPS, we have,

$$a \succeq b \text{ or } b \succeq a, \;\; \forall a, b \in A_t \;\Rightarrow\; aR_ib \text{ or } bR_ia, \;\; \forall a, b \in A_t \tag{4}$$

Hence, $R_{i,t}$ is *complete*.

For the proof of *transitivity*, we assume $n = 3$, $A_t = \{x, y, z\}$. Let node x prefer itself to y and is indifferent between y and z, as illustrated in Fig. 2b. The arrows between y and z indicate that both $y \succ z$ and $z \succ y$ hold true for x. Hence, x is indifferent between y and z.

So, we clearly have,

$$R_{x,t} : x \succ y \succeq z \tag{5}$$

Since x is indifferent to y and z, y and z have the same meaning to node x, i.e., $y \equiv z$. So, it follows that,

$$x \succ y \Rightarrow x \succ z \tag{6}$$

Thus, we conclude that $R_{i,t}$ is *transitive* even in presence of indifferences, $\forall i \in A_t$. Hence, we prove that each element of the set R_t is a weak order over set A_t.

Having said that R_t is the set of weak orders over set A_t, we now define the SCF that is to be followed in this paper as $f : R_t \rightarrow 2^{A_t} \setminus \Phi$, where A_t is already defined as the set of active nodes at time *t*.

The application of the *Theory of Social Choice* will also necessitate some definitions that are depicted below:

Definition 4.4. A node i is defined as a social choice winner at time t, if it is the outcome of a social preference profile R_t. A social choice winner i is mathematically expressed as:

$$\exists i \in N, f(R_{1,t}, R_{2,t}, ..., R_{n,t}) = i \tag{7}$$

Definition 4.5. Any node j is defined as a social choice loser at time t, if it is not a social choice winner. The set of social choice loser L can be expressed as:

$$L = \{\exists j \in N, f(R_{1,t}, R_{2,t}, ..., R_{n,t}) \neq j\} \tag{8}$$

*A. **Computation of IPS***

We study how to find the IPS of an node *i* at time *t*. In our case, a node i estimates the location of an *MN* j at time t as $K_{i,j,t} = (x_{j,t}, y_{j,t})$. Thereafter, each node prepares its own IPS. Following this, an election is executed in a way that renders a node to speak for its own preference or choice of an *MN*.

As previously stated, a node *i* expresses its priority of preference for all *MN*s $j : j \in M$. For this purpose, every node *i* computes the Euclidean distance of every other *MN* j (for the current timestamp) and the estimated location obtained, i.e., $K_{i,j,t}$.

In general, the Euclidean distance of a node *j* from a coordinate $K = (x, y)$ is defined as:

$$\xi(C(j), K) = \sqrt{(x - X[j])^2 + (y - Y[j])^2} \tag{9}$$

So, for all $i \in N$, each node computes the following:

$$\xi(C(i), K_{i,j,t}), \forall j \in M \tag{10}$$

Based on this, P_i, and I_i are respectively defined as:

$$aP_ib, \textit{ iff}, \ \xi(C(i), K_{i,a,t}) < \xi(C(i), K_{i,b,t}) \tag{11}$$

and

$$aI_ib, \text{ \textit{if}, } \xi(C(i), K_{i,a,t}) = \xi(C(i), K_{i,b,t}) \tag{12}$$

Hence, we arrive at the definition of R_i as:

$$aR_ib, \text{ \textit{if}, } \xi(C(i), K_{i,a,t}) \leq \xi(C(i), K_{i,b,t}) \tag{13}$$

After all the IPSs ($R_{i,t}, \forall i \in N$) are prepared, a preference profile R_t gets created at time t. A voting or election algorithm is executed on the preference profile until every node is allocated to an *MN*.

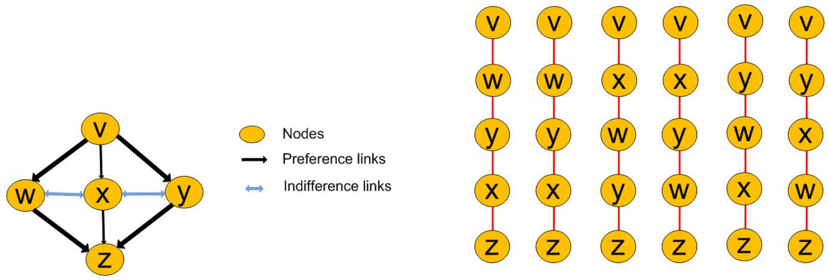
B. ***Computation of Borda scores***

Subsequent to the computation of IPS of every node, it is required to evaluate a node's preference score from the preference profile. Each node transmits its IPS to I_1. At this stage, a voting algorithm is executed. So, relative to every *MN*, there may exist both winner and loser nodes. The loser nodes get assigned to some other *MN*. In our case, the social choice function is entirely based on Borda's Count Method or simply referred to as the Borda method [44], [47], for decision making and aggregation. Following the Borda method, every *MN* m_i has a Borda score on behalf of each LDPU. The Borda score of a monitoring node m_j in an ordering of node i, denoted by $\beta_{m_j}(i)$. We have,

$$\beta_{m_j}(i) = |M| - r(m_j, R_{i,t}) \tag{14}$$

In Equation (14), $r(m_j, R_{i,t})$ is the rank of a monitoring node m_j in an ordering of node i at time t, i.e., the rank of m_j in $R_{i,t}$. A higher rank implies a lower position in preference ordering. For example, the most preferred will have rank 0, the next will have rank 1, and so on. This also implies that the Borda score increases as the preference increases.

4.2.1 ***Modified Borda Score for Weak Ordering****:* Equation (14) simply computes the Borda score of a node in a preference profile by merely assigning points to it. So, it can be used to find the Borda score $S(a)$ of a node a, only if preferences are *ordering* with no ties, i.e., none of the preferences is a weak ordering. In case of a tie or an indifference, Equation (14) fails to take care of the indifferent nodes. Hence, we intend to modify Equation (14) to prevent the loss of generality so that weak orderings can also be addressed.



(a) Weak order preference (b) Linear extensions

Fig. 3: Weak order analysis

In order to understand weak orderings more meticulously, we present some intuitive explanations of a weak ordered set. A weak order can be conceptualized as a set of *buckets*. We consider a set of weak preferences, partitioned into buckets, such that inter-bucket elements possess preference among one another and intra-bucket elements enjoy indifference. We can think of an IPS as set of k buckets B_0, B_1,..., B_{k-1}. These buckets form partitions over the set A_t. So, for every node i, $R_{i,t}$ can be expressed as a set of k buckets. For example, let us consider the preferences of node x in Fig. 3a. So, the weak ordering for node x is given below.

$$R_{x,y} = v \; P_x \; w \; I_x \; x \; I_x \; y \; P_x \; z \tag{15}$$

We partition the ordering into 3 ($k = 3$) buckets. $B_2 = \{v\}$, $B_1 = \{w, x, y\}$, $B_0 = \{z\}$. B_2 and B_0 are the most and least preferred ones. We denote the length of each bucket i as L_i, i.e., $L_i = |B_i|$, and a node at the i^{th} bucket of an ordering by a_i.

We introduce a new term Linear Extension (LE) [33] of a weak ordering (or simply an ordering) $R_{i,t}$, denoted by $\mathcal{L}_{R_{i,t}}$. It is defined as an ordering of all possible logical permutations of preferred and indifferent nodes, taken in the correct order. So, referring to Fig 3, we have the following LE of $R_{x,t}$, expressed as a Hasse diagram in Fig. 3b.

In the LE, we see that nodes v and z are at the top and bottom, respectively. Since nodes w, x and y are indifferent to another, all possible permutations of these nodes are taken.

We use LE to find the Borda score of an alternative a_i in the i^{th} bucket of a weak ordering of a node v, and denote it by $LE_{a_i}(v)$.

*4.2.2 **Computation of** $LE_{a_i}(v)$:* Let there be n active nodes in the network, and let B_0, B_1,.., B_{k-1} be the bucket partitions of a weak ordering of node v. We denote the elements of $WO(v)$ such that

$\{a_1, a_2, ..a_a\} \in B_{k-1}$, $\{a_{a+1}, a_{a+2}, ..a_{a+b}\} \in B_{k-2}$, ..., and $\{a_{a+b+...+1}, a_{a+b+...+2}, ..a_{a+b+...+\sigma}\} \in B_0$, where $a+b+...+\sigma = n$. Let $\beta_1, \beta_2, ...\beta_n$ be the Borda scores of nodes $a_1, a_2, .., a_{a+b+\sigma}$, such that $\beta_n < \beta_{n-1} < ... < \beta_1$. Under strict preference, every bucket contains exactly one element. Hence, the rank of the i^{th} node in an ordering will be i itself (e.g., rank of the 0^{th} node is 0, rank of the 1^{st} node is 1, and so on). So, according to Equation (14), $\beta_i = n - i$. Let $L_{i+1} = \lambda$, and $|B_i| = L_i = \rho$. We also have, $L_{k-1} = a, L_{k-2} = b, ...,$ and $L_0 = \sigma$. So,

$$L_{k-1} + L_{k-2} + ... + L_{i+1} = a + b + ... + \lambda = \delta \tag{16}$$

Therefore, $\{a_{\delta+1}, a_{\delta+2}, ..., a_{\delta+\rho}\} \in B_i$. Since, alternatives within a bucket are equivalent, there are $L_i!$ number of ways in which ρ elements of B_i can be linearly ordered. Moreover, each alternative $a_i \in B_i$ is assigned a single Borda score for all possible permutations of the other equivalent alternatives $(L_i - 1)!$ number of times. The LE method computes the Borda Score assigned to each alternative over all possible linear extensions of a weak ordering, averages it, and eventually assigns this value as an alternative's Borda Score in the weak ordered preference. Hence, we have [33],

$$LE_{a_i}(v) = \frac{(L_i - 1)!\beta_{\delta+1} + ... + (L_i - 1)!\beta_{\delta+\rho}}{L_i!} \tag{17}$$

Simplifying, we get,

$$\begin{aligned} LE_{a_i}(v) &= \frac{[n - (\delta + 1)] + ... + [n - (\delta + \rho)]}{L_i} \qquad (18) \\ &= n - L_{k-1} - L_{k-2} - ... - L_{i+1} - \frac{L_{i+1}}{2} \qquad (19) \end{aligned}$$

Hence, Equation (14) reduces to:

$$\beta_a(i) = LE_a(v_i) \tag{20}$$

C. Formation of Pseudo-Cluster

After Borda scores are successfully assigned to every alternative, as described in Section B, the pseudo-clusters (PC) are formed. Each PC, on behalf of an *MN* will be formed based on the *weak-ordering* of nodes, placed in the sequence of best-to-worst preference. The PC for an *MN* i is denoted by $A_{m_i,t+1}$. Every time the election algorithm executes, an *MN* emerges as a winner as per Definition 4.4. The corresponding node (LDPU) gets appended to $A_{m_i,t+1}$, and it remains active for the next timestamp as well. The winner is removed entirely from the preference profile to avoid having redundant winners, and the election algorithm is repeated. In this manner, the proposed algorithm selects the nodes that will be monitored by their corresponding *MN*s at timestamp t+1.

A node i is allocated to the PC of an *MN* m_j ($A_{m_i,t+1}$), if the following inequality is satisfied.

$$\beta_{m_j}(i) \geq \beta_{m_j}(k), \forall k \in M \tag{21}$$

Consequently *MN* m_j emerges as a social winner and subsequently, node i is added to $A_{m_j,t+1}$. The main reason for the success of m_j is that it has obtained the maximum Borda score from the society of nodes. In return of its success, m_j agrees to provide monitoring service to node i. Hence, i is decided to be a member of the PC of m_j.

*D. **Data Aggregation from Society of Nodes***

Once a cluster is formed, the next step is to aggregate the data before transmitting it to the health-cloud. This aggregation depends on the social choice of nodes that forms the pseudo-cluster PC for every *MN*. During aggregation, it is intended to prioritize patients based on their health criticality. So, for every LDPU, we introduce a new metric named *Exigency Factor* (*EF*). *EF* of an LDPU i is denoted by Γ_i. *EF* estimates the deviation of the patient's health data from a threshold value. Let $y_{i,t}$ and y_{norm} be the measured and standard values of a human, respectively. So, we have,

$$\Gamma_i = \frac{y_{i,t} - y_{norm}}{y_{norm}} \tag{22}$$

The data of each node i is fused based on the absolute health data value $y_{i,t}$, as well as the *Exigency Factor* Γ_i. Thus, the locally aggregated data of an LDPU is abstracted as a two-dimensional vector. The *Exigency Factor* contributes to the horizontal component and the absolute health data D_i contributes to the vertical component of it.

$$\Gamma_i = x_{i,t}, D_i = y_{i,t} \tag{23}$$

The Borda score $\beta_{m_j}(i)$ of the *MN* m_j on behalf of each LDPU i is a multiplicative factor. Since, health data from multiple LDPUs are aggregated in a single packet, the aggregated data of the pseudo-cluster will be a one dimensional vector containing the criticality-based data. Data of an LDPU i, i.e., D_i, is fused with its corresponding *Exigency Factor* to a composite data $\mathcal{D}_i$.

$$\mathcal{D}_i = {\Gamma_i}^2 \times D_i \tag{24}$$

Thus, the weight of the output data of each LDPU is directly proportional to the square of the *Exigency Factor* or the acuteness of the patient. This gain in the data weight eventually accounts for a faster response from the medical teams, once the packet reaches the health-cloud. After obtaining the severity-based weighted data from each LDPU of a pseudo-cluster, the aggregated data is merged in a uni-directional vector $\mathcal{V}$. $\mathcal{V}_{m_i}$ of a *MN* m_i is depicted in Fig. 4.

$LDPU_1$	$LDPU_2$	...	$LDPU_k$	$\Delta(\vec{m_i})$
$\mathcal{D}_1$	$\mathcal{D}_2$	...	$\mathcal{D}_k$	$\Delta_x(\vec{m_i})\hat{i} + \Delta_y(\vec{m_i})\hat{j}$

Fig. 4: Aggregated data format

We also define a *contribution* vector $\Delta(\vec{k})$ for node k as:

$$\Delta(\vec{k}) = \beta(k) \times \vec{K_{k,t}} \tag{25}$$

Resolving $\Delta(\vec{k})$ into its component vectors, we get,

$$\Delta_x(\vec{k}) = \beta(k)(x_{k,t}\hat{i}) = \beta(k)(\Gamma_k\hat{i}), \Delta_y(\vec{k}) = \beta(k)(y_{k,t}\hat{j}) = \beta(k)(D_k\hat{j}) \tag{26}$$

For each m_j, we have,

$$\Delta_x(\vec{\cdot}) = \frac{\sum\limits_{\forall i \in A_{m_j,t+1}} \Delta_x(i)\hat{i}}{\sum\limits_{\forall i \in A_{m_j t+1}} \beta(i)} \tag{27}$$

$$\vec{\Delta_y}(\cdot) = \frac{\sum\limits_{\forall i \in A_{m_j t+1}} \Delta_y(i)\hat{j}}{\sum\limits_{\forall i \in A_{m_j t+1}} S(i)} \tag{28}$$

where $\vec{\Delta_x}(\cdot)$ and $\vec{\Delta_y}(\cdot)$ are the aggregated horizontal and vertical vector components of the PC, respectively. After obtaining the fused *contribution* components of the cluster, we now obtain the aggregated data as:

$$\vec{\Delta(\cdot)} = \vec{\Delta_x}(\cdot)\hat{i} + \vec{\Delta_y}(\cdot)\hat{j} \tag{29}$$

where $\vec{\Delta(\cdot)}$ is the collective health data value of LDPUs, under a particular *MN* in 2D format. This aggregated information can be represented as a 1D vector $\mathcal{B}_1[1..m]$, $\mathcal{B}_1[i] = \vec{\Delta_{m_i}}(\cdot)$. $\vec{\Delta(\cdot)}$ plays a significant role while the aggregated data gets channelized through cloud gateways. We discuss this, in Section 5. The entire procedure of a PC-based data aggregation is encapsulated in Algorithm 1. Having discussed the functional details, we now finally design the goal function of *Banag*, denoted as $\mathcal{F}_1 : N \times M \rightarrow \mathcal{B}_1$. $\mathcal{F}_1$ is mathematically expressed as

$$\mathcal{F}_1(N, M) = \alpha(f(WO(N)^m), M), 1 \leq i \leq |M| \tag{30}$$

subject to the following constraints for "fairness".

$$\nexists i | \forall j, j \in \begin{cases} A_{m_1} & , i \in N \\ A_{m_2} & , i \notin N \end{cases}, m_1 \neq m_2 \tag{31}$$

$$\forall m_j \in M, \sum_{i \in A_{m_j}} \beta_{m_j}(i) \geq \sum_{i \in 2^n} \beta_{m_j}(i) \tag{32}$$

$$\nexists i \in A_{m_j} | \beta_{m_j}(i) < \beta_{m_k}(i) \tag{33}$$

$$\forall i \in N, R_i = P_i \cup I_i \tag{34}$$

where α denotes the aggregation function. Equations (49) through (34) denote the constraints due to non-dictatorship, universality, unanimity, and transitivity, respectively. Evidently, the outcome of *Social*

Choice is aggregated based on health priority and subsequently transmitted. We analyze the computational complexity of our function below.

Lemma 4.3. *The asymptotic computational complexity of the SCF (f), is $T_f(m,n) = O(\max(m,n) \times n)$, where m and n are the number of mobile monitoring nodes, and the number of LDPUs in a single pseudo-cluster, respectively.*

Proof: To compute the computational complexity of f, we initially determine the complexity to determine the Borda scores (β) of every LDPU from the linear ordering. An LDPU creates a preference ordering (IPS) for all the m monitoring nodes of the society in $O(m)$ time. The total computation complexity, $C_1(m,n)$, for generation of IPSs is governed by the Equation below.

$$\begin{aligned} C_1(n) &= C_1(n-1) + O(m) + \Theta(1), C_1(1) = c_1 \\ &\Rightarrow O(mn) \end{aligned}$$

where $\Theta(1)$ is the time required for decision making and linear ordering. For tie-breaking, Equation (20) is implemented. Assuming an *indifference* among all m monitoring nodes, we have the computational complexity for tie-breaking ($C_2(n)$) modeled as,

$$\begin{aligned} C_2(n) &= C_2(n-1) + O(n) + \Theta(c), C_2(1) = c_2 \\ &\Rightarrow O(n^2) \end{aligned}$$

The overall computational complexity, is equated as,

$$\begin{aligned} T_f(m,n) &= O(mn) + O(n^2) \\ &= O(\max(m,n) \times n) \text{ (according to asymptotic algebra)} \end{aligned}$$

■

Proposition 4.1. The worst case asymptotic computational complexity of Banag, T(m,n), is $O(n \times \max(m,n) + m)$, where m and n are the number of mobile monitoring nodes, and the number of LDPUs in a single pseudo-cluster, respectively.

Proof: As obtained from Lemma 4.1, $T_f(m,n) = O(\max(m,n) \times n)$. We need to determine the computational complexity to perform data aggregation i.e., to complexity of α. We denote the complexity by T_α. The maximum number of LDPUs, possible under a single $A_{m_i}, \forall m_i \in M$, is $k = \lfloor \frac{n}{m} \rfloor$. Thus generation of $\vec{\Delta_{m_i}}(\cdot)$ takes $O(k)$ time. Therefore, we have,

$$T_\alpha(m) = T_\alpha(m-1) + O(k) + c_3 \simeq O(m)$$

Finally, we obtain,

$$T(m,n) = O(\max(m,n) \times n) + O(m) \simeq O(n \times \max(m,n) + m)$$

■

We now present a theoretical analysis of the performance of the proposed algorithm.

Theorem 4.4. *The probability of selection of the most suitable monitoring node from a set of size n is always less than that of a pseudo-cluster of the same size i.e., $P(R/P) > P(R/C)$.*

Proof: Let C and P denote the events of aggregation under the conventional and proposed approaches, respectively. Let R be the event of choosing the best nodes. Let n be the size of the set on which aggregation is performed.

Conventional approach:

A node is a loser if any of the $(n-1)$ nodes is a winner, instead.

$$P(\overline{R/C}) = \frac{n-1}{n} \tag{35}$$

So, the probability that the "best" node is selected is:

$$P(R/C) = 1 - \frac{n-1}{n} = \frac{1}{n} \tag{36}$$

Proposed approach: A node is voted by n nodes. A node i emerges as a winner *iff*, $\nexists j : \beta(j) > \beta(i), \forall j \neq i$. This means node i must receive more than 50% of votes in its favor. So, we have,

$$
\begin{aligned}
P(R/P) &> \frac{1}{2} && (37)\\
&\geq \frac{\frac{n}{2}+1}{n} && (38)\\
&\geq \frac{n+2}{2n} && (39)\\
&> P(R/C) && (40)
\end{aligned}
$$

Hence, it is proved that $P(R/P) > P(R/C)$.

■

Corollary 4.1. *The process of health criticality-based best node selection in Banag is deterministic, unlike usual aggregation.*

Proof: As already discussed in Section 2, usual aggregation algorithms do not consider health criticality. In Banag, each LDPU i possess Γ_i at a particular time t. At any point of time, we must have,

$$\forall i, \exists j, \Gamma_j \geq \Gamma_i, \forall i, j \in N \tag{41}$$

Thus, we deterministically choose node j. ■

Corollary 4.2. *The "winner" node chosen for monitoring the LDPUs is independent of dictatorship.*

Proof: Banag is Pseudo-Cluster centric. Prior to aggregation, the data from each LDPU is processed to generate a preference profile from the society of LDPUs. The collective decision of choosing a monitoring node that is finally reached is based on equal contribution of the individual preference of each LDPU. Thus, the winner node is chosen democratically.

■

5. OCA: The Optimal Channelization Algorithm

After performing data aggregation, the aggregated data is transmitted to the health-cloud interfaced by the cloud gateways. The *OCA* focuses on channelizing health data from the *monitoring nodes* to the cloud gateways, thereby achieving a "fair" distribution of traffic load. The load balancing among various

Algorithm 1 The Body Area Network Data Aggregation Algorithm (Banag)

Input:

- Set of LDPUs at time *t*: N
- Set of monitoring nodes *MN* at time *t*: M

Output: A set of pseudo-clusters $A_{m_i,t+1}, \forall m_i \in M$

for all $i \in N$ **do**
  for all $j \in M$ **do**
    Compute $\xi(C(i), K_{i,j,t})$
  end for
  Create the IPS of node i at time t: $R_{i,t}$
end for
count = $|N|$
while each node $i \in N$ is not assigned an *MN* **do**
  for $k = 1$ to count **do**
    Find m_k, such that $\exists j \in N \wedge \beta_j(m_k)$ is maximum
    $j \leftarrow A_{m_k,t+1}$
    Remove j from R_t
  end for
end while
for all $m_j \in M$ **do**
  for all $i \in A_{m_j,t+1}$ **do**
    Compute $\Delta_{\vec{x}}(i)$, $\Delta_{\vec{y}}(j)$
  end for
  Compute $\Delta_{\vec{x}}(\cdot)$, $\Delta_{\vec{y}}(\cdot)$
  Compute $\vec{\Delta}(\cdot)$
end for

gateways also reduces unwanted delays due to buffering and transmission. In this Section, we illustrate the *OCA* and its implementation.

Let there be g number of cloud gateways. The maximum demand of each gateway g_i is $D_{max_{g_i}}$. The current intake of a gateway is considered as the total number of monitoring nodes that are allocated to a particular gateway g_i, and is denoted by $D_{cur_{g_i}}$.

$$D_{cur_{g_i}} = \sum_{\forall m_i \in M \cap A_{g_i}} m_i \tag{42}$$

A cloud gateway g_i is assigned monitoring nodes, iff, the gateway is not an over-loaded. An over-loaded gateway is defined in Definition 5.1.

Definition 5.1. A gateway g_i is said to be over-loaded, if its current intake is less than its maximum demand. So, we have,

$$D_{cur_{g_i}} < D_{max_{g_i}} \tag{43}$$

The set A_{g_i} contains the set of *MN*s that are already allocated to gateway g_i. We apply social choice to decide the channelization of the data from the *MN*s to gateways. Initially, a dynamic cost matrix C_t is built for every sensor gateway pair. $C[i][j]$ is the communication cost of a mobile node m_i to a gateway g_j. C_t is expressed as:

$$C_t = \begin{bmatrix} C_{1,1} & C_{1,2} & .. & C_{1,g} \\ C_{2,1} & C_{2,2} & .. & C_{2,g} \\ . & . & .. & . \\ . & . & .. & . \\ C_{m,1} & C_{m,2} & .. & C_{m,g} \end{bmatrix}$$

We also introduce the term named *Capacity Factor* (CF). The CF of a gateway g_i, denoted as CF_{g_i}, is expressed as:

$$CF_{g_i} = \frac{D_{cur_{g_i}} - D_{max_{g_i}}}{D_{max_{g_i}}} \tag{44}$$

So, *CF* of a gateway has a fixed range, i.e., $0 \leq CF_{g_i} \leq 1$. We define a utility function $u(\cdot,\cdot)$ for each *MN* gateway pair.

$$u(m_i, g_j) = \frac{1}{C_t[i][j]} \times CF_{g_i} \tag{45}$$

We design a maximization function for the allocation of *MN*s to gateways. We choose an *MN* m_j such that the following is satisfied:

$$|\Delta(\vec{m}_j)| \geq |\Delta(\vec{m}_k)|, \forall m_k \in M \tag{46}$$

Thus, the monitoring node containing the most critical aggregated data is prioritized. For that m_j, we try to allocate a gateway in an optimal manner. A gateway g_i is allocated to m_j if the following condition is satisfied.

$$u(m_j, g_i) = \max(u(m_j, g_k)), \forall g_k \in G \tag{47}$$

This implies that m_j is added to set A_{g_i}. In this manner, the *MN*s are picked up in prioritized fashion, and are assigned to cloud gateways. An *MN* is allocated only if its utility value u for a particular gateway is greater than that corresponding to the other *MN*s for that gateway. However, after each assignment, the current intake $D_{cur_{g_i}}$ of a particular gateway increases, and, hence, its *CF* decreases. The goal function of *OCA* is represented as $\mathcal{F}_2 : M \times G \rightarrow \mathcal{B}_2$, where $\mathcal{B}_2[1..m]$ is the allocation matrix for every monitoring node. We mathematically express $\mathcal{F}_2$ as,

$$\mathcal{F}_2(M, G) = \mathcal{B}_2, \mathcal{B}_2[i] = \max_{g_j \in G} u(m_i, g_j) \tag{48}$$

subject to the following constraints for "fairness".

$$\nexists m_i | \forall m_j, j \in \begin{cases} A_{g_1} & , m_i \in M \\ A_{g_2} & , m_i \notin M \end{cases}, g_1 \neq g_2 \tag{49}$$

$$\forall g_j \in G, \sum_{m_i \in A_{g_j}} u(m_i, g_j) \geq \sum_{m_i \in 2^m} u(m_i, g_j) \tag{50}$$

$$\nexists m_i \in A_{g_j} | u(m_i, g_j) < u(m_i, g_k) \tag{51}$$

Equations (49) through (34) denote the constraints due to non-dictatorship, universality, and unanimity respectively. The complexity analysis of *OCA* is shown below and the algorithm is presented in Algorithm 2.

Proposition 5.1. The asymptotic computational complexity of OCA is $T_{OCA} = O(mg)$, where m and g are the number of monitoring nodes, and the number of cloud gateways, respectively.

Proof: Initially, to set the gateways parameters, $O(g)$ time is required. Followed by the initialization, time complexity ($T'(m, g)$) for allocation of gateways to monitoring nodes is modeled by a recurrence relation, using lines 7 to 15 of Algorithm 2. We have,

$$T'(m, g) = T'(m-1, g) + O(g) + c' \tag{52}$$

where, $O(g)$ amount of time is required at every iteration to schedule the gateways in a fair manner. c' is the constant amount of time for processing and analysis. Using Master Theorem, we obtain, $T' = O(mg)$. We have, $T_{OCA} = O(g) + T' \simeq O(mg)$. This completes the proof.

■

We now discuss some of the theoretical characteristics of *OCA*.

Theorem 5.1. *As the number of gateways increases, the probability of the correct gateway to be the social choice winner converges to a constant value.*

Proof: Let there be n number of cloud gateways. At a particular time, only a single gateway emerges as the winner. Let us assume that *success* X denotes the event of choosing the correct gateway. The *failure* signifies choosing the wrong gateway. We have,

$$P(X) = \frac{1}{n}, P(\overline{X}) = \frac{n-1}{n} \tag{53}$$

Let S denote the event of a gateways emerging as a social winner. Following binomial distribution, we have,

$$\begin{aligned} P(S) &= P(X \geq 1) \\ &= 1 - ({}^nC_1 + {}^nC_2 + .. + {}^nC_n) \\ &= 1 - {}^nC_0\left(\tfrac{n-1}{n}\right)^n \\ &= 1 - \left(\tfrac{n-1}{n}\right)^n \end{aligned}$$

For large values of n, we have, $\frac{n-1}{n} \approx 1$. Experimentally, it has been established that,

$$\left(\frac{n-1}{n}\right)^n \approx 0.37 \tag{54}$$

$$\text{Therefore, } P(S) \approx 0.73 \tag{55}$$

This proves the theorem.

Theorem 5.2. *In OCA, an over-loaded gateway is virtually a social choice loser.*

Proof: A gateway g_i is over-loaded if the following condition is satisfied.

$$Dcur_{g_i} = Dmax_{g_i} \tag{56}$$

Therefore, $CF_{g_i} = 0$. This implies, $u(\cdot, g_i) = 0$. Hence, for every other gateway g_j, we have,

$$u(\cdot, g_i) > u(\cdot, g_j), \forall g_j \in G, g_j \neq g_i \tag{57}$$

Therefore, $A_{g_i} = \phi$. This signifies that no monitoring node is allocated to cloud gateway g_i. This can happen only when g_i is a social loser. This completes the proof.

Algorithm 2 Optimal Channelization Algorithm (OCA)

Input:

- Set of monitoring nodes *MN* at time *t*: M
- Set of cloud gateways at time *t*: G

Output: A channelization table containing the mapping of *MN*s to gateways

for all $m_i \in M, g_j \in G$ **do**
 Compute $C_t[m_i][g_j]$
end for
for all $g_i \in G$ **do**
 Compute $D_{cur_{g_i}}$, *Capacity Factor* CF_{g_i}
end for
for all $m_i \in M$ **do**
 Compute m_l with the maximum priority
 for all $g_j \in G$ **do**
  Compute $u(m_l, g_j)$
  $max \leftarrow \max(u(m_l, g_j))$
 end for
 $A_{g_h} \leftarrow m_l | u(m_l, g_h) = max$
 Update $D_{cur_{g_h}}$, Update *Capacity Factor* CF_{g_h}
end for

6. Examples

In this Section, we illustrate those situations where the normal voting or deciding algorithms fail, and Borda's voting strategy contributes. Similar to [7], we discuss the flaws in each, and, finally, we establish the superiority of Borda's count over the other scenarios.

To keep the notions simple, we denote the number of LDPUs, and the number of monitoring nodes as *n* and m, respectively. The linear and weak order preferences are denoted by $\succ$ and $\succeq$, respectively.

Case 1:

Let $n = m = 100$, and the preferences be: 1^{st}: $m_1 \succ m_2 \succ .... \succ m_{100}$, 2^{nd}: $m_1 \succ m_2 \succ .... \succ m_{100}$, .. , 51^{st} : $m_1 \succ m_2 \succ .... \succ m_{100}$, 52^{nd}: $m_{100} \succ m_2 \succ .... \succ m_1$, .. , 100^{th}: $m_{100} \succ m_2 \succ .... \succ m_1$. Clearly, node m_1 wins. After judging the suitability of node m_1, we conclude that nearly half of the nodes have the worst preference for m_1. On the other hand, node m_2 is a better selection.

Case 2:

The most common mechanism of choosing a winner that is widely followed is *plurality voting*. Assuming $m = 21$, let the preferences be: 10 nodes have $m_1 \succ m_2 \succ ... \succ m_{21}$, 6 nodes have $m_2 \succ m_3 \succ m_1 \succ m_4 \succ m_5 \succ ... \succ m_{21}$, and 5 nodes have $m_3 \succ m_2 \succ m_1 \succ m_4 \succ m_5 \succ ... \succ m_{21}$. Since m_1 has maximum 10 votes,it wins. But the majority actually preferred something else. Although m_1 receives 10 votes for itself, but 11 out of 21 nodes preferred either m_2 to m_1 or m_3 to m_1.

Case 3:

From cases 1 and 2, we conclude that choosing a winner node by absolute majority of votes, may lead to an erroneous result. So, a two-stage voting system is helpful. The top two winners at stage 1 are m_1 and m_2, with 10 and 6 votes, respectively. Then, we check the number of preferences. Naturally, m_2 wins, because m_2 is preferred to m_1 in 11 cases. But in this two stage system, there arises a severe flaw. Let us consider the preferences as follows: 4 nodes have $m_1 \succ m_2 \succ m_3 \succ ... \succ m_{11}$, 4 nodes have $m_3 \succ m_2 \succ m_1 \succ m_4 \succ m_5 \succ ... \succ m_{11}$, and 3 nodes have $m_2 \succ m_3 \succ m_1 \succ m_4 \succ m_5 ... \succ m_{11}$. So, node m_3 wins. If 2 nodes, among the first 4, decide not to vote, the less preferred node m_3 wins. So, the resulting system has the following preferences: 2 nodes have $m_1 \succ m_2 \succ m_3 \succ ... \succ m_9$, 4 nodes have $m_3 \succ m_2 \succ m_1 \succ m_4 \succ m_5 \succ ... \succ m_9$, and 3 nodes have $m_2 \succ m_3 \succ m_1 \succ m_4 \succ m_5 ... \succ m_9$.

Surprisingly, in this scenario, node m_2 emerges as the winner. Hence, this two-stage voting system does not inspire nodes to participate in election.

Case 4:

Another drawback of the two-stage system in the case of *separability* [7]. Suppose we have $n = 26$, and voting is performed in two sections, each consisting of 13 nodes. Let the preferences be: 4 nodes have $m_1 \succ m_2 \succ m_3 \succ ... \succ m_{26}$, 3 nodes have $m_2 \succ m_1 \succ m_3 \succ m_4 \succ m_5 \succ ... \succ m_{26}$, 3 nodes have

$m_3 \succ m_1 \succ m_2 \succ m_4 \succ m_5 \succ ... \succ m_{26}$, and 3 nodes have $m_3 \succ m_2 \succ m_1 \succ m_4 \succ m_5 \succ ... \succ m_{26}$. Let the other half be organized as follows: 4 nodes have $m_1 \succ m_2 \succ m_3 \succ ... \succ m_{26}$, 3 nodes have $m_3 \succ m_1 \succ m_2 \succ m_4 \succ m_5 \succ ... \succ m_{26}$, 3 nodes have $m_2 \succ m_3 \succ m_1 \succ m_4 \succ m_5 \succ ... \succ m_{26}$, and 3 nodes have $m_2 \succ m_1 \succ m_3 \succ m_4 \succ m_5 \succ ... \succ m_{26}$.

Following two-stage voting, m_1 wins in both sections. However, if the voting was done on the entire set of 26 nodes, m_1 would have been eliminated in the first stage. Hence, this method is non-separable.

Case 5:

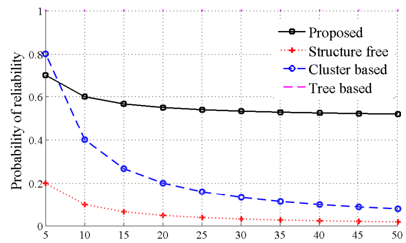
This is a special case, where we discuss about the *agenda* in voting. Suppose we have $n = 3$, and the preferences are $m_1 \succ m_2 \succ m_3$, $m_2 \succ m_3 \succ m_1$ and $m_3 \succ m_1 \succ m_2$. Such preferences will not have a definite winner. If we set the *agenda* by taking b and c first, and then node a, a wins the election. Hence, any node can be a winner arbitrarily, based on how we set the agenda. Consequently, we can say that this method is not "neutral". It exhibits agenda-based bias.

Borda Count Analysis :

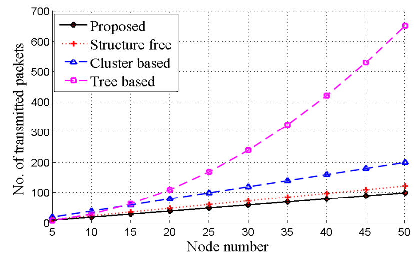
We analyze the Borda count method of social choice voting. We refer to the above-mentioned cases, and, thereby, understand the performance of the Borda count method. In Case 1, we find $S(m_1) = 5149, S(m_2) = 9900$, $S(m_3) = 9800$, ... and, $S(m_{100}) = 4951$. So, the optimally preferred node m_2 emerges to be the winner with the maximum Borda score. Again, in Case 2, we have $S(m_1) = 419$, $S(m_2) = 426$, $S(m_3) = 415$, and, so on. Evidently, m_2 emerges as the socially chosen winner. In Case 3, we have $S(m_1) = 107$, $S(m_2) = 113$, and $S(m_3) = 110$. The other nodes have lower Borda scores naturally, because of their low ranks. So, m_2 is the socially preferred node. Likewise, the other cases can also be analyzed using the Borda count strategy of social choice voting.

7. Simulation Results

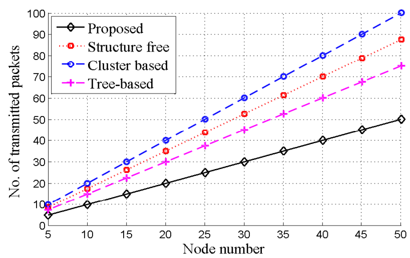
In this Section, we discuss the results of simulations of both the proposed data aggregation and channelization solution, *Banag* and *OCA*, respectively. We also present the results of comparison of *Banag* with the existing cluster-based [42], tree-based [41] and structure-free [12] data aggregation methods. It may be noted here that while the existing solutions were proposed for sensor data aggregation in the upstream node, *Banag* aggregates data collected from the LDPUs. We also examine the correctness of *OCA* in the later part of the section.



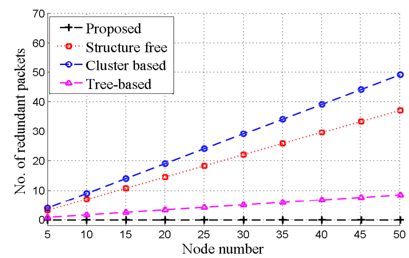
(a) Comparison of Probability of reliability



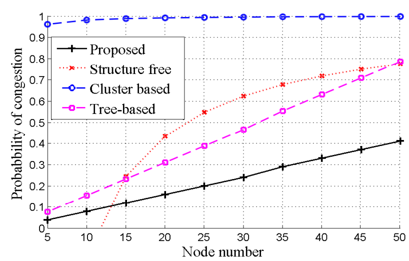
(b) Comparison of packets transmitted in the worst case



(c) Comparison of packets transmitted in the average case



(d) Comparison of redundant packets



(e) Comparison of probability of congestion

Fig. 5: Evaluation results

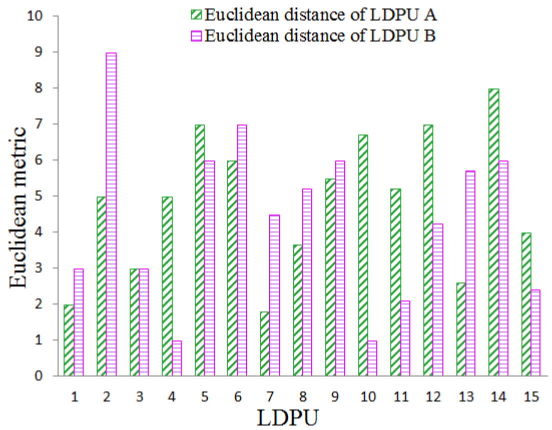
A. *Simulation setup*

The simulation was carried over uniformly and randomly placed 50 LDPUs (transmission range = 30m), 15 mobile monitoring nodes (transmission range = 50m), and 3 cloud gateways for analysis of the network parameters, as illustrated by Fig 5. The size of a pseudo-cluster is assumed to be 5 and the density of the monitoring nodes as 3 per km^2. However, to enhance the understandability of pseudo-cluster formation and generation of preferences, as illustrated in Fig 6, the comparative analysis is performed over 2 LDPUs. In Figs 7, and 8, the performance of *OCA* is evaluated using 3 cloud-gateways with variable computing capacity.

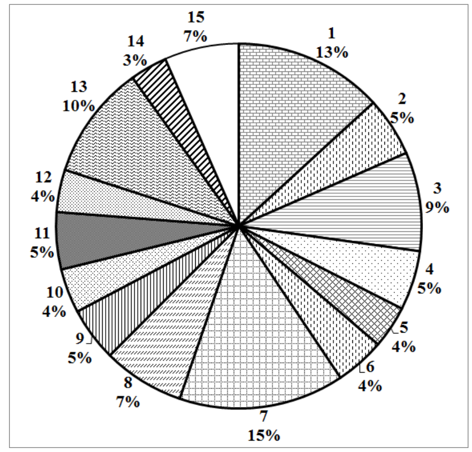
B. *Simulation parameters*

We discuss the simulation parameters below.

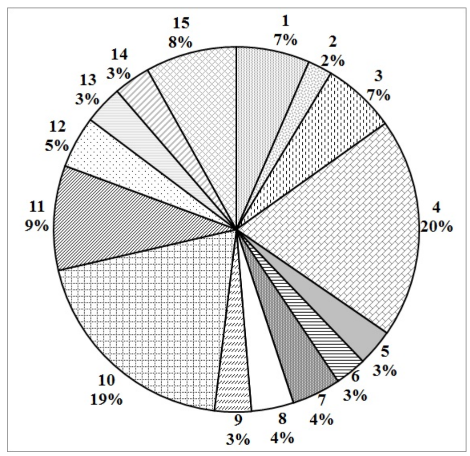
- *Reliability* - Reliability is highly significant for packets containing health data. In this work, the term



(a) Demonstration of nodes vs Euclidean metric



(b) Individual preference of LDPU A



(c) Individual preference of LDPU B

Fig. 6: Social preference analysis

reliability (r) refers to the ratio of the number of "fair" nodes to the total available nodes in an aggregation process. Thus,

$$r = \frac{\sum_{k_i} 1}{n}, \exists k_i \in N, m_j \in M | k_i \in (N \cap A_{m_j \cdot \cdot}),$$

- *No. of transmitted packets* - The total number of packets transmitted to the data aggregation center in the best (n_b), and the worst case (n_w) are analyzed, assuming a zero packet drop rate. We find that for *Banag* $n_b = \max_{m_j \in M}(|A_{m_j \cdot \cdot}|)$, and $n_w = \min_{m_j \in M}(|A_{m_j \cdot \cdot}|)$. For structure free methods, $n_b = \frac{2 \times n}{k_1}$, and $n_w = \frac{2 \times n}{k_2}$, where k_1, and k_2 are random numbers denoting to the minimum, and maximum out-degree of a senor node in an adhoc environment, respectively. For tree based methods, $n_b = \min_{p_i \in P}(p_i)$, and $n_w = \max_{p_i \in P}(p_i)$, where P is the set of all distinct paths connecting from the root to a leaf of the topology. For clustered approaches, $n_b = c \prod_{m_j \in M}(|A_{m_j \cdot \cdot}|)$, where c is the number of distinct clusters in the topology. The worst case occurs, when the network is too big, i.e., c is too large.

- *No. of redundant packets* - Redundancy arises due to unnecessary packet transmission to the center of aggregation. This metric is computed as the total number of nodes (g) that tend to transmit redundant information due to close proximity within the network. g is expressed as,

$$g = \sum_{\exists n_i, n_j \in N} 1, \xi(C(n_i, n_j) \leq d_{th}$$

where, d_{th} is a pre-negotiated threshold distance value to determine proximity.

- *Probability of congestion* - The probability of congestion (P_c) is determined as,

$$P_c = \lceil \frac{n_w}{C_{max}} \rceil$$

where, C_{max} is denoted as the maximum channel capacity in terms of number of packets.

Fig. 5(a) shows the variation of reliability metric in different aggregation techniques. Compared to the structure-free and cluster-base aggregation, the proposed algorithm *Banag* performs remarkably better. However, this metric is not much relevant to tree-based methods, because in such cases each node forms a hierarchical tree structure. So, every node of the network is at some hierarchy of the tree, and is a part of the aggregation. Hence, each node, be it "fair" or not, is considered in the aggregation process. Since every available node is considered, the probability of the reliability metric is always unity for tree-based approaches. Figs. 5(b) and 5(c) illustrate the count of packets transmitted for the purpose of aggregation in the worst and best case scenarios, respectively. In the worst case, we see that the proposed algorithm, *Banag*, outperforms the tree-based and cluster-based approaches, whereas it marginally surpasses the structure free method. Fig. 5(c) shows that Banag performs better than the other three by a reasonable margin, thereby indicating its energy-efficiency as well. As the energy consumption is highly affected by packet communication, a reduction in the count of transmitted packets impacts the battery lifetime also.

Fig. 5(d) illustrates the approximate number of redundant packets that might be sent. In Banag, redundancy is negligible because every node's opinion is distinctly considered with that of the other nodes during aggregation. Moreover, each node generates its unique ordering of preferences.

Fig. 5(e) reveals the likelihood of congestion in the different data aggregation techniques. In this context as well, as the proposed aggregation algorithm performs better than the existing ones. This manifests its improvisation on the aspects of traffic control.

We now analyze the policies of generating IPS, and, subsequently, generating social preference from those. For the sake of simplicity, and for the purpose of illustration of preference profile, we considered a hypothetical scenario of 2 LDPUs and 15 monitoring nodes. Fig. 6(a) shows the Euclidean relation between the LDPUs and the monitoring nodes. Based on the metric, LDPU A prepares an IPS, as shown in Fig. 6(b). Similarly, Figs. 6(c) shows the IPS of LDPU B. Fig. 6(b) and 6(c) clearly illustrate the effect of the Euclidean metric on the IPS of the nodes. After aggregation, LDPUs A and B are monitored by m_7 and m_4, respectively.

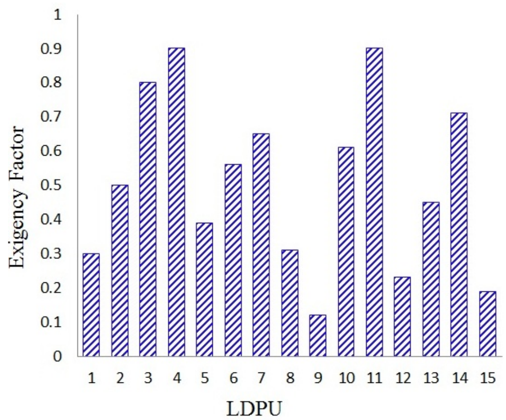
So far we have considered the impact of pseudo-cluster formation of *Banag*. Now we focus on its data aggregation aspect. The aggregation is totally based on the acuteness of health of patients. Fig. 7(a) illustrates randomized *Exigency Factor* of 15 monitoring nodes. Fig. 7(b) shows the sensed data value recorded by the body sensors. This raw data is further processed under the influence of the *Exigency Factor*. The health severity-based composite data finally produced is the one that is in the aggregated form, the graph of which is shown in Fig. 7(c). It shows that patients with high *Exigency Factor* eventually generate a high magnitude of composite data. This magnitude-based weight of the data enables the medical teams to prioritize their service to patients with more health criticality.

Having discussed about the communication advantages of using Banag, now we analyze some results that follow from the *OCA*. *OCA* hugely impacts the allocation of gateways to the monitoring nodes. Fig. 8(a) initially describes the current intake $D_{cur_{g_i}}$ and the maximum demand $D_{max_{g_i}}$ of each gateway g_i. In our experiment we considered $|G| = 15$. Based on the values of $D_{cur_{g_i}}$ and $D_{max_{g_i}}$, the *Capacity Factor* CF_{g_i} is computed, and is shown in Fig. 8(b). A comparative evaluation of Figs. fig:subfigure6 and 8(b) illustrate that the most socially preferred gateways are the ones that have a higher difference between $D_{cur_{g_i}}$ and $D_{max_{g_i}}$, i.e., these gateways can currently serve nodes more efficiently than the others.

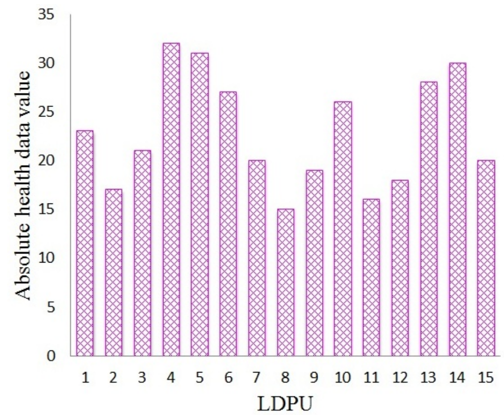
From Fig. 8(c), it follows that based on the *utility* value (u) of a *MN* gateway pair, each node will be allocated to a gateway by following Algorithm 2. Fig. 8(c) illustrates the results of gateway allocation with ten *MN*s and three gateways. Using *OCA*, we get,

$$\sum_{\forall m_i \in C_{t_A}} u(m_i, g_A) = 48.9, \sum_{\forall m_i \in C_{t_B}} u(m_i, g_B) = 45.8, \sum_{\forall m_i \in C_{t_C}} u(m_i, g_C) = 52$$

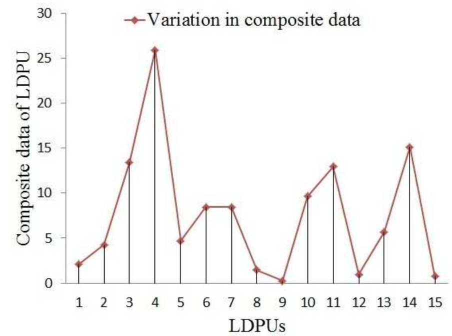
Fig. 8(c). suggests the following allocation of gateways to mobile nodes.



(a) Evaluation of *Exigency Factor* of LDPUs



(b) Demonstration of absolute health data value



(c) Demonstration of composite health data value

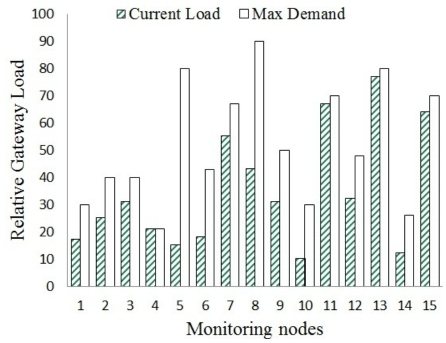
Fig. 7: Impact analysis of *Exigency Factor*

$$A_{g_A} = \{m_1, m_4, m_5\}, A_{g_B} = \{m_6, m_7, m_9\}, A_{g_C} = \{m_2, m_3, m_8, m_{10}\}$$

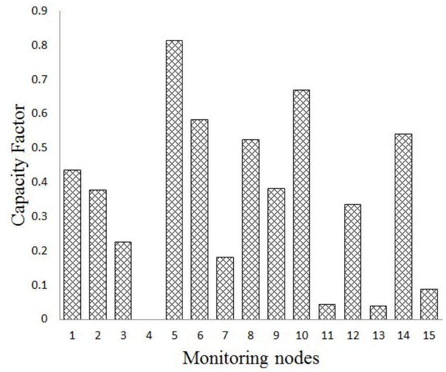
Thus, we find that the summation of the utility values of the allocated nodes for different gateways are differ negligibly, with a standard deviation of 2.5. This suggests that the proposed algorithm is unbiased to gateways and maintains uniformity in allocation.

8. Conclusions

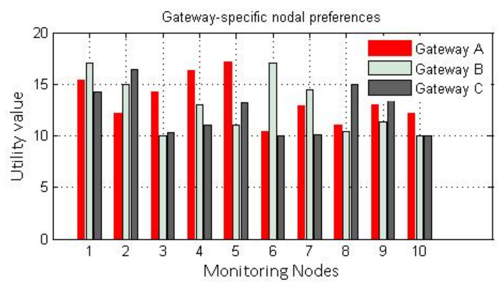
In this paper, we proposed a cloud-assisted WBAN based architecture for aggregating data from LDPUs embedded within patients, and analyzed some of the social choice issues in it. We also proposed an algorithm for channelizing data through dynamic gateway allocation. In the process of aggregation and channelization, we focused on the acuteness of a patient and also expressed the health-criticality as a metric of the transmitted packet. This also enables the medical teams to develop an elementary idea of each patient from the data packet itself. Since our work is based on the *theory of Social Choice*, “fairness” is incorporated while data aggregation and channelization. Health data from LDPUs are aggregated within



(a) Assessment of current capacity and maximum demand of gateways



(b) Projection of relative Capacity Factor of gateways to the monitoring nodes



(c) Evaluation of gateway specific nodal preferences

Fig. 8: Preference analysis of gateways

mobile monitoring nodes. The selection of the center of aggregation is performed by selection the Social Choice winner from the individual preference profiles of the LDPUs. Data from the mobile monitoring nodes are further channelized through cloud gateways. Gateways are selected to reduce communication cost and optimally transmit data. Aggregated data are also fairly distributed among gateways to prevent over burdening of gateways and subsequent delaying that might result in a slow responsive system.

In future we plan to extend our work by considering heterogeneous data while aggregation. A composition of health data collected from heterogeneous body sensors can be fused to make the data more meaningful and informative to the medical experts. In this paper we have considered aggregation of data from on-body sensor nodes only. However, with due considerations, data from sensors implanted within the human body, can be integrated with data from the on-body sensors to produce a better report of the physiological status of a patient. This work can be also extended for object recognition and tracking in surveillance-like scenarios.

References

[1] Y. Alemu, J. bin Koh, and D. K. Kim, “A hierarchical multi-parent cluster-based data aggregation framework for WSNs,” in *International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)*, February 2013, pp. 549 – 554.

[2] H. Alipour, M. Abbaspour, M. Esmaeili, H. Mousavi, and H. Shahhoseini, "DACA: Dynamic advanced clustering algorithm for sensor networks," in 14^{th} *IEEE International Conference on Electronics, Circuits and Systems*, December 2007, pp. 518 – 525.

[3] D. Anzai, K. Katsu, R. Chavez-Santiago, Q. Wang, D. Plettemeier, J. Wang, and I. Balasingham, "Experimental evaluation of implant UWB-IR transmission with living animal for body area networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 1, pp. 183 – 192, January 2014.

[4] J. Bae, K. Song, H. Lee, H. Cho, and H. J. Yoo, "A low-energy crystal-less double-FSK sensor node transceiver for wireless body-area network," *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 2678 – 2692, November 2012.

[5] T. Banerjee, K. Chowdhury, and D. P. Agrawal, "Tree based data aggregation in sensor networks using polynoial regression," in 8^{th} *International Conference on Information Fusion*, vol. 2, July 2005.

[6] S. H. Bouk and I. Sasase, "Multiple end-to-end qos metrics gateway selection scheme in mobile ad hoc networks," in *International Conference on Emerging Technologies, 2009. ICET*, October 2009, pp. 446 – 451.

[7] D. Bouyssou, T. Marchant, and P. Perny, *Decision-making Process: Concepts and Methods*. Wiley, January 2010.

[8] Y. Cai and J. Tan, "Secure group communication in body area networks," in *International Conference on Information and Automation*, June 2008, pp. 555 – 559.

[9] Y.-S. Chen, C.-S. Hsu, and H.-K. Lee, "An enhanced group mobility protocol for 6LoWPAN-based wireless body area networks," *IEEE Sensors Journal*, vol. 14, no. 3, pp. 797 – 807, March 2014.

[10] C.-T. Cheng, H. Leung, and P. Maupin, "A delay-aware network structure for wireless sensor networks with in-network data fusion," *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1622 – 1631, May 2013.

[11] M. N. Deylami and E. Jovanov, "A distributed scheme to manage the dynamic coexistence of IEEE 802.15.4-based health-monitoring WBANs," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 1, pp. 327 – 334, January 2014.

[12] K. W. Fan, S. Liu, and P. Sinha, "Structure-free data aggregation in sensor networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 8, pp. 929–942, August 2007.

[13] G. Fernndez, I. de la Torre-Dez, and J. J. P. C. Rodrigues, "Analysis of the cloud computing paradigm on mobile health records systems," in *Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, July 2012, pp. 927 – 932.

[14] H. Ghasemzadeh and R. Jafari, "Data aggregation in body sensor networks: A power optimization technique for collaborative signal processing," in 7^{th} *Annual IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON)*, June 2010, pp. 1 – 9.

[15] S. H. Han and S. K. Park, "Performance analysis of wireless body area network in indoor off-body communication," *IEEE Transactions on Consumer Electronics*, vol. 57, pp. 335 – 338, May 2011.

[16] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the* 33^{rd} *Hawaii International Conference on System Sciences*, 2000.

[17] D. C. Hoang, R. Kumar, and S. K. Panda, "Optimal data aggregation tree in wireless sensor networks based on intelligent water drops algorithm," *IET Wireless Sensor Systems*, vol. 2, pp. 282 – 292, September 2012.

[18] O. D. Incel and B. Krishnamachari, "Enhancing the data collection rate of tree-based aggregation in wireless sensor networks," June 2008, pp. 569 – 577.

[19] M. Iqbal, X. Wang, D. Wertheim, and X. Zhou, "Loadbalanced multiple gateway support in wireless mesh networks for broadband services," in 18^{th} *Annual Wireless and Optical Communications Conference, 2009. WOCC*, May 2009, pp. 1–5.

[20] S. M. A. Iqbal and M. H. Kabir, "An improved internet gateway discovery and selection scheme in mobile ad hoc network," in *International Conference on Electrical and Computer Engineering (ICECE)*, December 2010, pp. 372 – 375.

[21] E. Jovanov, A. O. Lords, D. Raskovic, P. G. Cox, R. Adhami, , and F. Andrasik, “Stress monitoring using a distributed wireless intelligent sensor system,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 22, pp. 49 – 55, July 2003.

[22] W. S. Jung, K. W. Lim, Y. B. Ko, and S. J. Park, “A hybrid approach for clustering-based data aggregation in wireless sensor networks,” in 3^{rd} *International Conference on Digital Society*, February 2009, pp. 112 – 117.

[23] N. A. Khan, N. Javaid, Z. A. Khan, M. Jaffar, U. Rafiq, and A. Bibi, “Ubiquitous healthcare in wireless body area networks,” in *IEEE* 11^{th} *International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, June 2012, pp. 1960 – 1967.

[24] J. S. Kim, S. Y. Choi, S. J. Han, J. H. Choi, J. H. Lee, and K. W. Rim, “Alternative cluster head selection protocol for energy efficiency in wireless sensor networks,” in *International Conference on Software Technologies for Future Dependable Distributed Systems*, March 2009, pp. 159 – 163.

[25] Y. Liao, H. Qi, and W. Li, “Load-balanced clustering algorithm with distributed self-organization for wireless sensor networks,” *IEEE Sensors Journal*, vol. 13, no. 5, May 2013.

[26] H. Lin, J. Shao, C. Zhang, and Y. Fang, “Cam: Cloud-assisted privacy preserving mobile health monitoring,” *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 985 – 997, June 2013.

[27] R. B. Manjula and S. S. Manvi, “Cluster based data aggregation in underwater acoustic sensor networks,” in *Annual IEEE India Conference (INDICON)*, December 2012, pp. 104 – 109.

[28] R. Manoharan and S. Mohanalakshmie, “A trust based gateway selection scheme for integration of manet with internet,” in *International Conference on Recent Trends in Information Technology (ICRTIT)*, June 2011, pp. 543 – 548.

[29] D. Mantri, N. R. Prasad, R. Prasad, and S. Ohmori, “Two tier cluster based data aggregation (TTCDA) in wireless sensor network,” in *IEEE International Conference on Advanced Networks and Telecommuncations Systems (ANTS)*, December 2012, pp. 117 – 122.

[30] L. Meng, H. Zhang, and Y. Zou, “A data aggregation transfer protocol based on clustering and data prediction in wireless sensor networks,” in 7^{th} *International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, 2011, pp. 1 – 5.

[31] P. Nie and B. Li, “A cluster-based data aggregation architecture in WSN for structural health monitoring,” in 7^{th} *International Wireless Communications and Mobile Computing Conference (IWCMC)*, July 2011, pp. 546 – 552.

[32] P. Norris, “Choosing electoral systems: Proportional, majoritarian and mixed systems,” Special issue of the International Political Science Review, pp. 297–312, July 1997.

[33] D. Polett and A. O. Hudson, “Empowering the voter: A mathematical analysis of borda count elections with non-linear preferences,” The Division of Science, Mathematics, and Computing of Bard College, Tech. Rep., May 2010.

[34] S. S. Ranjani, S. R. Krishnan, and C. Thangaraj, “Energy-efficient cluster based data aggregation for wireless sensor networks,” in *International Conference on Recent Advances in Computing and Software Systems (RACSS)*, April 2012, pp. 174 – 179.

[35] Z. Rashid, U. Farooq, J.-K. Jang, and S.-H. Park, “Cloud computing aware ubiquitous health care system,” in *E-Health and Bioengineering Conference (EHB)*, November 2011, pp. 1–4.

[36] R. Ravindran, “Arrow’s impossibility theorem,” *Resonance*, pp. 1–26, November 2005.

[37] D. Reddy, D. T. Kumar, and G. Ramu, “An efficient cloud framework for health care monitoring system,” in *International Symposium on Cloud and Services Computing (ISCOS)*, December 2012, pp. 113 – 117.

[38] E. Reusens, W. Joseph, B. Latre, B. Braem, G. U. Vermeeren, E. Tanghe, L. Martens, I. Moerman, and C. Blondia, “Characterization of on-body communication channel and energy efficient topology design for wireless body area networks,” *IEEE Transactions Information Technology in Biomedicine*, vol. 13, pp. 933 – 945, November 2009.

[39] M. B. Shah, P. P. Verma, S. N. Merchant, and U. B. Desai, “Human mobility based stable clustering for data aggregation in singlehop

cell phone based wireless sensor network," in *IEEE International Conference on Advanced Information Networking and Applications (AINA)*, March 2011, pp. 427 – 434.

[40] K. Tada and M. Yamamoto, "Load-balancing gateway selection method in multi-hop wireless networks," in *IEEE Global Telecommunications Conference, 2009. Globecom*, December 2009, pp. 1 – 6.

[41] C. H. Tsai, H. Y. Huang, C. W. Hung, and Y. H. Wang, "TDAM: A tree-based data aggregation mechanism in wireless sensor networks," in *IEEE International Symposium on Intelligent Signal Processing and Communication Systems*, November 2012, pp. 4–7.

[42] O. Wu, W. Hu, S. J. Maybank, M. Zhu, and B. Li, "Efficient clustering aggregation based on data fragments," *IEEE Transactions on Systems Man and Cybernetics*, vol. 2, no. 3, pp. 913–926, June 2012.

[43] L. Xiang, J. Luo, and C. Rosenberg, "Compressed data aggregation: Energy-efficient and high-fidelity data collection," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1722 – 1735, December 2013.

[44] L. Yong and W. Zulin, "Rank aggregation performance analysis for borda and local search algorithm," in *International Symposium on Information Science and Engineering (ISISE)*, December 2010, pp. 125 – 128.

[45] L. Zhou and H. Wang, "Toward blind scheduling in mobile media cloud: Fairness, simplicity, and asymptotic optimality," *IEEE Transactions on Multimedia*, vol. 15, no. 4, June 2013.

[46] L. Zhou, Z. Yang, and J. J. P. C. Rodrigues, "Exploring blind online scheduling for mobilecloud multimedia services," *IEEE Wireless Communications*, vol. 20, no. 3, pp. 54 – 61, June 2013.

[47] K. Zou, "Borda method of fuzzy decision making," in *International Conference on Computer Science and Electronics Engineering (ICCSEE)*, vol. 3, March 2012, pp. 403 – 405.