Secured Data Aggregation In WSN Using Genetics

Balaji R. Dande
Department of InformationTechnology
Sinhgad College of Engineering
Pune, Maharashtra, INDIA

Nilesh J. Uke
Department of Information Technology
Sinhgad College of Engineering
Pune, Maharashtra, INDIA

***Abstract*—In wireless sensor networks (WSNs), the cluster-based data aggregation technique increases energy rate and consumes extra amount of energy. Secure data transmission and data aggregation are critical in designing cluster-based sensor networks. More-ever the secured data transmission is critical for enhancing the data authentication and confidentiality. In order to organize secure data communication in typical WSN this project proposes a genetically derived secure cluster-based data aggregation in WSN. The proposed model elects cluster head as data aggregator, the cluster head selection is based on node connectivity. Then genetic algorithm is used clustering process. The cluster members communicated through the aggregator, before organizing data transmission the aggregator needs to check member authentication. The encryption technique isutilized that offers authenticity, confidentiality and integrity.**

***Keywords*—WSN, Data Aggregation, Genetic Algorithm, Energy efficiency, Authentication**

I. INTRODUCTION

The sensor network incorporates little and less cost sensing devices together with remote radio handset for looking at nature. It includes the information gathering also transmitting the data to one or more sink nodes. The primary focal point of this system is that it doesn't require any framework or outer supply for information gathering [1]. The primary utilizations of WSN are wild territory checking, woodland fire recognition, building security observing, military observation etc.

Wireless sensor network are extremely vitality obliged frameworks in which every sensor device has a restricted force and processing capacity [1] [2]. They are basically event based frameworks with a wide scope of ecological sensing applications from vehicle following to living space observing. A sensor organize generally comprises of an extensive number of sensor nodes isolated into groups. Base station gathers and forms the information from bunches. Contingent upon the structural planning, there may be more than one level of cluster heads between sensor hubs and the base station.

The limitations of WSN impact on security, the adversity will take an advantage of sensor network. In addition, the sensor nodes occupied with limited power and limited memory, it will leads disadvantage to carry a data across different nodes [3]. The WSN characteristics will lead to the different security challenges. More ever the limited resources availability can minimizes data transmission it needs an extra bandwidth. The data aggregation increases bandwidth to improve data transmission

The procedure of gathering the data from different sources emulated by excess end accordingly minimizing the transmission consider is termed information collection. This procedure prompts vitality protection. In addition, the inalienable excess in information assembled from the sensor hub can be evacuated through the methodology of in-system information aggregation application particular data. In any case, in threatening environment, the accumulated information needs to be shielded from different assaults for achieving the information privacy, honesty and validation. Subsequentlysecurity plays a major part in information [4][5].

The proposed method composes different security techniques such as data confidentiality, data integrity, data freshness and authentication. The procedure of protecting the transmitted information from inactive assaults compares to data confidentiality. The methodology of securing the information from unlawful client is the most difficult errand. This can be explained by utilizing an encryption procedure such that just the proposed client with suitable key can open and read the data.[6][7][8][9]

The data integrity procedure, which guarantees that the transmission of information from the sender to the collector is not tainted, is said to be information uprightness. This uncovers that the information are gotten with no duplication, insertion, and adjustment and reordering. The data freshness ensures that the transmitted information is crisp and no enemy can replay old messages. That is, it defends the transmitted information from replay assault.

Authentication guarantees the unwavering quality of the message by recognizing its cause. This allows the recipient node to check whether the information has been transmitted by asserted sender or not. Subsequently, unless the enemy has a legitimate verification key, it can't infuse information or partake in any occasion inside the system. Data authentication ensures the received data by verifying data with ordinal data.

II. BACKGROUND

1. DATA AGGREGATION

The data aggregation is a procedure used to tackle the implosion and cover issues in information driven steering. Information originating from different sensor nodes is collected as though they are about the same quality of the sensation when they achieve the same steering node on the route once more to the sink. Data collection is a generally utilized procedure as a part of wireless sensor systems. The security issues, data privacy and trustworthiness, in data collection get to be fundamental when the sensor system is conveyed in an unfriendly environment. Data collection is a procedure of collecting the sensor information utilizing conglomeration approaches. Figure 1 illustrate that information total is the methodology of totaling the sensor information utilizing conglomeration approaches. At that point the calculation utilizes the sensor information from the sensor nodes and afterward totals the information by utilizing some total calculations such as unified methodology.

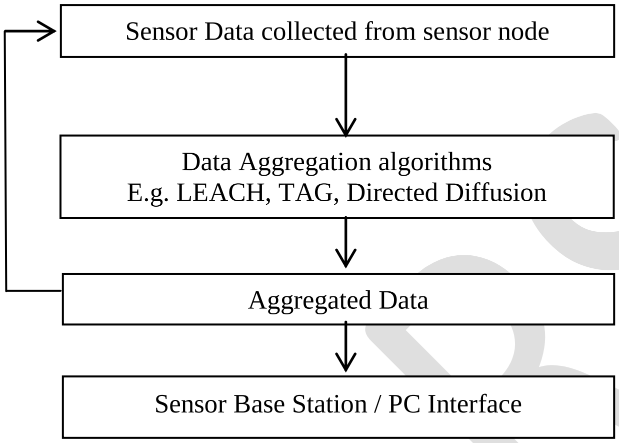

Sensor Data collected from sensor node
Data Aggregation algorithms
E.g. LEACH, TAG, Directed Diffusion
Aggregated Data
Sensor Base Station / PC Interface

Fig. 1 Architecture of Data aggregation

1.1 DATA AGGREGATION BASED NETWORK

Flat networks: In flat networks, each sensor node assumes the same part and is furnished with roughly the same battery power. In such systems, information accumulation is fulfilled by information driven directing where the sink normally transmits a question message to the sensors, for instance, through flooding and sensors which have information matching the inquiry send reaction messages once again to the sink. The decision of a specific correspondence convention relies on upon the particular application nearby.

Diffusion: Directed diffusion (DD)[10] may be a famous data conglomeration ideal model for remote gadget systems. It’s an information driven and application mindful standard, inside the sense that each one data produced by sensor hubs is called by property estimation sets. Such a plan joins the data returning from entirely unexpected sources in transit to the sink by killing excess and minimizing the measure of transmissions. Amid this implies it spares the vitality utilization and will expand the system lifespan of WSNs. Amid this subject normally base station telecast the message to the intrigued supply hub. Thusly every hub gets interest. These hobbies plot the quality worth like name of item. Each hub get the investment will reserve it for later utilize. Since the investment is shown by the system bounce by jump, inclination square measure setups to draw data fulfilling the inquiry at the asking for hub. An inclination may be an answer connection to the closer from that the investment was gotten.

SPIN: The sensor convention for information through negotiation [11] the gazing hub that has new information publicizes the information to the nearby hubs inside the system utilizing the Meta information. A nearby hub that is keen on this sort of data sends approaching to the pioneer hub for information. The pioneer hub reacts and sends information to the sinks each hub has an asset overseeing ability to stays informed regarding its vitality use inside the sensing component system. Each hub surveys its assets like battery power before information transmission. SPIN is additionally appropriate for situations with portable sensors, since the sending choices are in light of local neighborhood information.

Hierarchical Networks: In which information conglomeration information must be carried out at unique nodes, with the assistance of these unique hub we can diminish the quantity of number of information bundle transmitted to the sink. So with this system enhances the vitality productivity of the entire system.

1.2 CLUSTER-BASED NETWORKS FOR DATA AGGREGATION:

These Wireless sensor system is asset imperative that is the reason sensor can't straightforwardly transmit information to the base station. In which all general sensors can send information parcel to a bunch head (neighborhood aggregator) which totals information parcel from all the consistent sensors in its bunch and sends the succinct summary to the base station. With the assistance of the plan we spare the vitality of the sensors. Drain [12]: Low vitality versatile bunching has been proposed to compose a sensor system into a set of groups so that the vitality utilization can be occasion disseminated among all the sensor hubs.

1.3BASED NETWORKS FOR DATA AGGREGATION:

In which every sensor sends information to the closer neighbor. Power- Efficient Data-Gathering Protocol for Sensor Information Systems (PEGASIS) is kind of chain based information collection. In PEGASIS [13], all sensors are organized into a direct chain for information total. The nodes can structure a chain by utilizing a ravenous calculation or the sink can choose the chain in an incorporated way. In the Greedy chain development accepts that all sensors have comprehensive information of the system. The most remote node from the sink starts chain arrangement and, at each one stage, the closest neighbor of a node is chosen as its successor in the chain. In every information assembling round, a hub gets information bundle from one of its neighbors, totals the information with its own, and sends the totals information bundle to its other neighbor along the chain. Inevitably, the pioneer node in there is like group head sends the accumulated information to the base station. Figure-2 underneath demonstrates the chain based information conglomeration strategy in PEGASIS.

1.4 TREE BASED NETWORKS FOR DATA AGGREGATION:

In which all nodes are composed in manifestation of tree means progressive, with the help of halfway node we can perform information accumulation procedure and information transmit leaf node to root node. Tree based information total is suitable for applications which include in network information total. An illustration application is radiation-level observing in an atomic plant where the most extreme worth gives the most helpful data to the wellbeing of the plant. One of the primary parts of tree-based systems is the development of a vitality productive information conglomeration tree.

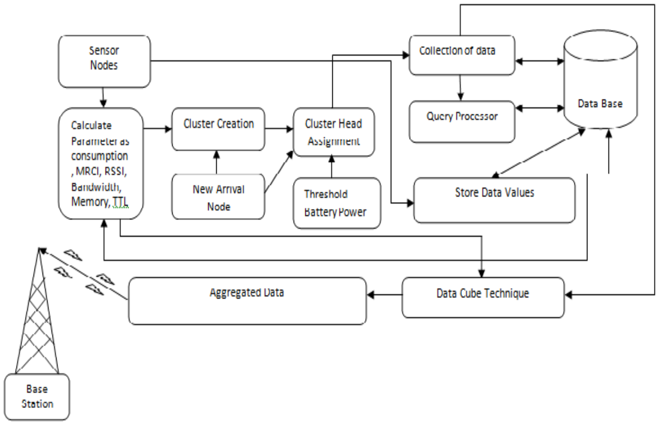


Fig. 2 Architecture of Data Collection and aggregation for WSN

III. PROPOSED WORK

Initially the CHs are picked in view of the hub network, which goes about as data aggregator (DAG). At that point, the bunching procedure is executed utilizing the hereditary calculation. This procedure exceptionally minimizes vitality utilization and consequently upgrading the system lifetime. At the point when the bunch part needs to transmit the information to the aggregator, an information encryption procedure is used. The crypto module (CyM) used offers classified to the information parcel (DP), along these lines guaranteeing the legitimacy and respectability of the sensed information.

Genetic Algorithm

A versatile system that aides in understanding pursuit and streamlining issues relates to GAs. It is in view of the hereditary techniques performed by natural living beings.

The GA is shown in the following [14]

Step- 1: Start

Step2: Initialize n-number of nodes

Step3: Compute node efficiency by considering node fitness Status.

Step 4: Repeat this step to all nodes until find best efficient Node.

Step 5: Select parent node by considering above step

Step 6: Execute crossover to produce offsprings

Step 7: Perform mutations

Step 8: Compute node energy of each individual

Step 9: Replace the parents by the corresponding offsprings in new generation.

End if

End if

3.1 CLUSTER FORMATION BASED ON GA

Here each node in a network represented with a bit of chromosome. The nodes in a cluster is divided as cluster member nodes and Cluster head CH and with are represent as 0 and 1s respectively

Step1: Create nodes in a network and elect a cluster head CH by considering its connectivity nature, which node have a highest energy it chooses as a CH node

Step2: To allocate a sensor nodes to the selected cluster head CH by using cluster function F. The nodes allocation is bases on these parameters such as distribution factor (α), transmission cost factor (β) and energy factor (γ).

Step 3: Each individual within the cluster is evaluated based On fitness function f_i, this function is estimated based on cluster distance, transmission range and node distance

Step4: Based on fitness function f the population make over into the future generation

Step5: Utilizing Roulette–Wheel determination strategy, the chromosomes with higher fi when contrasted and other chromosome is chosen for producing another offspring

Step6: Replace lower fitness chromosomes with a new offspring from among selected chromosomes

Step 7: The idea of elitism is likewise incorporated that ensures that the current best individual at every era dependably made due to the cutting edge.

Step 8: Emulating the chromosome choice, single-point hybrid procedure is connected

Step 9: When normal node turn as a cluster head CH, then all other nodes verifies whether the elected node near to all the other nodes, if it satisfies then normal node become as a new cluster head CH.

Step 10: After the hybrid, every bit of an individual is connected over the change administrator. In this methodology, when the bit esteem is 0, it is changed to 1 and the other way around. This strategy evades duplication of people. This guarantees hereditary differences inside the populace.

Step 11: Reconstruct the clusters by considering cluster heads CHs.

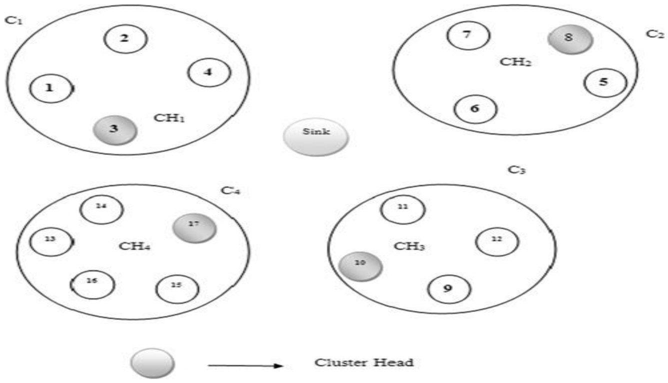


Fig. 3 Cluster Formation

IV. Mathematical Model

Fitness function of the chromosome is calculated as [14]

$$f_i = \frac{1}{E_T} + (D - Dc) + (n - n_c)$$

$$E_T = Energy\ requried\ for\ data\ transmission$$

D= sum of the distance from all nodes to sink

Dc = Sum of the Distance from all nodes to cluster and cluster to sink

Distribution factor calculated as[15]

$$\propto = \frac{N}{C}$$

N = Total number of sensor nodes

C= Total number of cluster heads

Cost Factor: It estimated by considering distance among nodes

$$\mathrm{b} = \mathrm{z} + 10\rho \log 10(\frac{\mathrm{D}}{\mathrm{D_{ref}}}) + \mathrm{g}$$

Energy Factor: The energy factor is calculated by considering sum of transmission energy, reception energy and idle energy state. Energy of CH sensing state is calculated as

$$\gamma = E_{tx} + E_{rx} + E_{idle} + E_{sen}$$

Residual energy is calculated as [16]

$$E_r = E_i - (E_{tx} + E_{rx} + E_{idle} + E_{sen})$$

4.1Energy Efficient Encryption Technique

This strategy gives the safe correspondence structure that checks the DP and drops the false DPs from malignant hubs. The way to encryption plan shifts as a capacity of leftover vitality of the hubs, along these lines keeping the rekeying necessities.

Let consider initial residual energy of nodes. Initialization vector represents vector predistributed to cluster members.

This system obliges every group part and DAG_i to store some data in its reserve that incorporates current virtual vitality and node ID.At the point when a source hub S needs to transmit a DP to its DAG_i, it at first creates a dynamic key (K_i) in light of its starting virtual remaining energy. On the off chance that, the source has effectively performed a secured information transmission, then amid resulting information transmission, the element key is produced as capacity of the past key. S encrypts DP with generated previous keyand it is passed to CyM. CyM takes the key and DP fields as inputs and produces the change code (PC) utilizing a RC5 encryption procedure. This encoded DP is encoded with the PC alongside plain content hub ID and sent to the DAG_i. DAG_ion accepting the message checks the related current E_{ri} put away for the sending hub and concentrates it to determine the key. The message verification is confirmed by deciphering the message and contrasting the plaintext hub ID and the encoded hub ID.In case, DAG_i needs to transmit gathered DP (CDP) from its group individuals to the base station (BS), it encodes CDP with its dynamic key (KCH_i) and encode it utilizing the Rivest Cipher 5 (RC5) calculation and send it to BS safely. BS on accepting CDP translates and separates the information like the past steps.

V. EXPERIMENT RESULTS

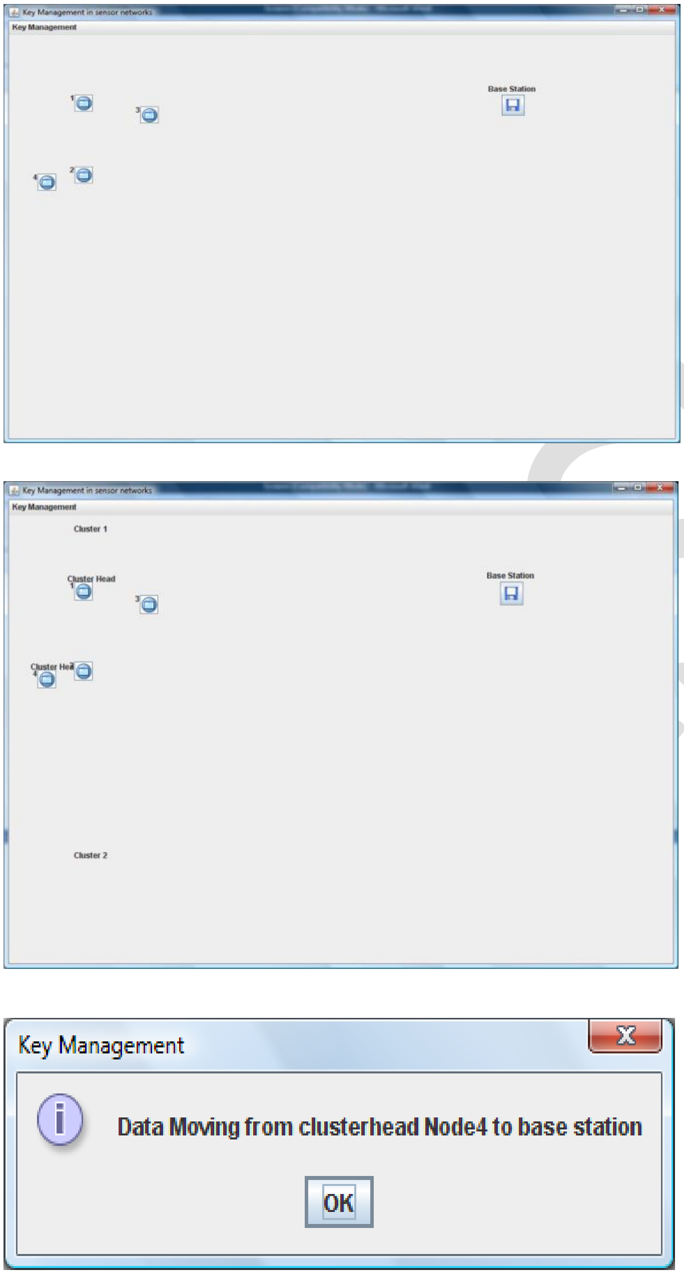
The performance of proposed GDSDA evaluated through Java. A network is design with a range of 1000*1000 sq. we configure a nodes by placing all these nodes on selected region. Here we consider different number of nodes; we assign different energy rate level and transmission range level. We consider bandwidth as 1Mbps and we have evaluated the proposed approach to evaluate results. Here we create four different modules to organize communication

Base station: A processing center, which acts like a gateway in typical networks, it allows to access to another network by providing accessing point.

Sensor Node: Sensor nodes communicate other nodes in a WSN, which exchange a data by communicating with other nodes. Each sensor node have unique ID, it communicates with base station (B) by using encrypted data. The base station assigns a unique ID to the sensor node before it deploying the main key Kshare. The bases station represents Kshare and node ID. The sensor nodes will grouped as a cluster and it will allocated to a base station by specifying initial key Kinit.

Clusters: Cluster is a group of nodes, which is a combination of cluster head and sensor node numbers. It has unique cluster ID and cluster key (Kc). The data exchange is occurred only in between cluster head and cluster nodes rather than sensor nodes for better scalability and group key management according to the Exclusion Basis System framework.

Cluster Head: Cluster head contains better resources and which manage local traffic during communicating with base station. The cluster head is responsible to manage data of cluster nodes during communicating with base station, the kind of communication in between cluster head and cluster node uses authentication key K_c . The data should be encrypted.



Data Moving from clusterhead Node4 to base station

Fig. 4 Cluster creation

Description: After selecting number of nodes which are shown in the above screen shots, now these nodes are ready to communicate with cluster head node, here node-1 and node-4 is cluster head in cluster -1. Now we need to select a data where to move from node -1 to node-4 in a given cluster -1. The cluster head -4 is ready to communicate with base station, by sharing with source node. It's generate a window to select a data file, now we are selecting a file it's ready to move a data from source node 3 to cluster head node -4, after that it moves cluster head node-4 to base station in secured manner.

VI. Conclusion

In this paper,proposed a GDSDA in Wireless Sensor Network at first the CHs are picked in view of the hub integration, which goes about as a DAG. At that point, the bunching methodology is executed utilizing the hereditary calculation. This procedure very minimizes vitality utilization and consequently improving the system lifetime. At the point when the bunch part needs to transmit information to the aggregator, an information encryption procedure is used. The CyM used offers privacy to the DP, subsequently guaranteeing the validness and honesty of the sensed information. By reproduction results, we demonstrate that the proposed procedure minimizes the vitality utilization, guarantees information security and lessens the transmission overhead.

References

[1] Patil, N.S., Patil, P.R.: 'Data aggregation in wireless sensor network'.IEEE Int. Conf. Computational Intelligence and Computing Research,2010

[2] Roy, S., Conti, M., Setia, S., Jajodia, S.: 'Secure data aggregation inwireless sensor networks', IEEE Trans. Inf. Forensics Sec., 2012, 7, (3)

[3] Bhoopathy, V., Parvathi, R.M.S.: 'Securing node capture attacks forhierarchical data aggregation in wireless sensor networks', Int. J. Eng.Res. Appl., 2012, 2, (2), pp. 466–474

[4] Jha, M.K., Sharma, T.P.: 'A new approach to secure data aggregationprotocol for wireless sensor network', Int. J. Comput. Sci. Eng.(IJCSE), 2010, 02, (05), pp. 1539–1543

[5] Makin, B.A., Padha, D.A.: 'A trust-based secure data aggregationprotocol for wireless sensor networks', IUP J. Inf. Technol., 2010, VI,(3), pp. 7

[6] Sen, J.: 'A survey on wireless sensor network security', Int. J. Commun.Netw. Inf. Secur. (IJCNIS), 2009, 1, (2), pp. 55–78

[7] Ozdemir, S., Xiao, Y.: 'Secure data aggregation in wireless sensornetworks: A comprehensive overview', Comput. Netw., 2009, 53,(12), pp. 2022–2037

[8] Alzaid, H., Foo, E., Nieto, J.G.: 'Secure data aggregation in wirelesssensor network: a survey'. Proc. Sixth Australasian InformationSecurity Conf. (AISC), Australian Computer Society, 2008, pp. 93–105

[9] Jha, M.K., Sharma, T.P.: 'Secure data aggregation in wireless sensornetwork: a survey', Int. J. Eng. Sci. Technol. (IJEST), 2011, 3, (3),pp. 2013–2019

[10] KiranMaraiya, Kamal Kant, Nitin Gupta "Architectural Based DataAggregation Techniques in Wireless Sensor Network: A Comparative Study",International Journal on Computer Science and Engineering (IJCSE), Vol. 3No. 3 Mar 2011

[11] VaibhavPandey, AmarjeetKaur and Narottam Chand "A review on dataaggregation techniques in wireless sensor network", Journal of

Electronic andElectrical Engineering, ISSN: 0976–8106 & E-ISSN: 0976–8114, Vol. 1,Issue 2, 2010

[12] Wendi RabinerHeinzelman, AnanthaChandrakasan, and HariBalakrishnan "Energy-Efficient Communication Protocol for Wireless Microsensor Networks", Proceedings of the 33rd Hawaii International Conference on System Sciences IEEE, 2000.

[13] Kwon, T., Hong, J.: 'Secure and efficient broadcast authentication inwireless sensor networks', IEEE Trans. Comput., 2010, 59, (8),pp. 1120–1133

[14] Heidari, E., Movaghar, A.: 'An efficient method based on geneticAlgorithms to solve Sensor network optimization problem',Int. J. Appl. Graph Theory Wirel. Ad Hoc Netw.Sens. Netw.(GRAPH-HOC), 2011, 3, (1), pp. 18–33

[15] Zahmatkesh, A., Yaghmaee, M.H.: 'A genetic algorithm-based approachfor energy- efficient clustering of wireless sensor networks'. Int. Conf.Network Communication and Computer (ICNCC), 2011

[16] Quang, V.T., Miyoshi, T.: 'Adaptive routing protocol with energyefficient and event clustering for wireless sensor networks', IEICETrans., 2008, E 91-B, (9), pp. 2795–2805