A Review on Overall Analysis of Different Data Dissemination Strategies

Kanak Manjari

Abstract- Wireless Sensor Network (WSN) technology has provided the availability of small and low-cost sensor nodes with capability of sensing various types of physical and environmental conditions, data processing, and wireless communication. But the sensor nodes in WSN have a limited transmission range, and their processing and storage capabilities as well as their energy resources are also limited. Thus an efficient architecture is required for overcoming these issues which strongly depends on Data Dissemination strategy used. Data Dissemination, where each node is able to send data to every other node in the mesh network and each node decides to forward the data using some protocols. So it is an important issue to design suitable Dissemination protocol according to different application scenarios. A variety of different energy efficient Data Dissemination protocols were proposed in recent years. In this paper, we present the comprehensive analysis of various Data Dissemination strategies. This paper provides useful insights for the network designer such as which Data Dissemination protocols scale well, reduce overall energy consumption or improve task completion.

Keywords: Wireless Sensor Network, Data Dissemination, Dissemination strategies.

1. Introduction

A Wireless Sensor Network has spatially distributed autonomous sensors to monitor different physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants and to cooperatively pass their data through the network to a main location [1]. It is also called Wireless Sensor and Actor Network (WSAN) sometimes. The topology of the WSNs can vary from a simple star network to an advanced multi-hop Wireless Mesh Network. Basically WSN is the collection of mobile nodes or static nodes which are capable of communicating with each other in order to collect data accurately. It is a multi-hop and self-organization network which consists of sink nodes and sensor nodes with wireless communication mode. Sink nodes have the responsibility to link to wireless sensor network and external network. Each sensor node is a smart embedded device which includes four units: data collection unit, data process unit, wireless communication unit and battery. Figure 1 shows a schematic diagram of components of sensor node. Wireless sensor network has four basic components: processing unit, sensing unit, radio unit (communication unit), battery (power source). Once deployed across the monitoring region, each of them can complete data sensing and collection independently. Meanwhile, using communication unit, they can coordinate with each other to realize data delivery according to different queries. The main application of wireless sensor networks is to sense the environment and transmits the acquired information to the sink for further processing.

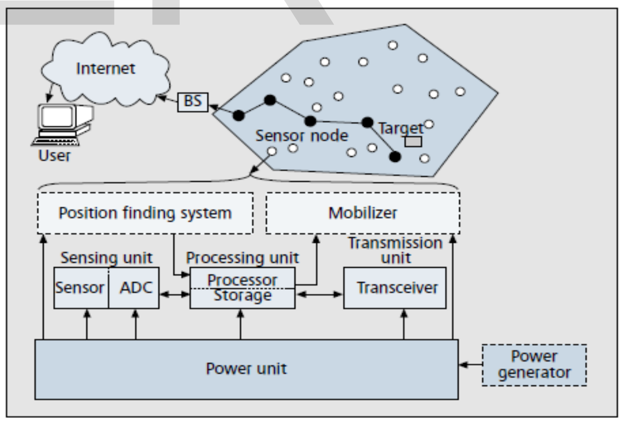


Figure 1: Components of Sensor Nodes

2. Data Dissemination

The process of routing of data and queries in WSN is called data dissemination. Data Dissemination is the process of transferring desired data from active sensor nodes to data collecting nodes (i.e. sink) in the sensor network. ***Source Node*** is a node which generates data and event where event is information to be routed. Node which is interested in data called the ***Sink Node*** and interest is descriptor for some event that node is interested in. Event is transferred from source to sink after source receives an interest message from sink. Data dissemination is two steps process in which in first step interest of nodes is broadcasted in network and in second step nodes after receiving the request sends data having requested data. There are various data dissemination methods: Flooding, Gossiping, Spin etc.

ISSN 2229-5518

2.1 FLOODING

In the flooding protocol each node receiving a data or management packet repeats the packet by broadcasting it. Only packets which are destined for the node itself or packets [4] whose hop count has exceeded a preset limit are not forwarded. The main benefit of flooding protocol is that it requires no costly topology maintenance or route discovery. Once sent a packet will follow all possible routes to its destination. If the network topology changes sent packets will simply follow the new routes added. Flooding does however have several problems. One such problem is **implosion**. Implosion is where a sensor node receives duplicate packets from its neighbors". Figure 2 illustrates the implosion problem. Node P broadcasts a data packet ([P]) which is received by all nodes in range (nodes Q and R in this case). These nodes then forward the packet by broadcasting it to all nodes within range (nodes P and S). This results in node S receiving two copies of the packet originally sent by node P. This can result in problems determining if a packet is new or old due to the large volume of duplicate packets generated when flooding. **Overlap** is another problem which occurs when using flooding. If two nodes share the same observation region both nodes will witness an event at the same time and transmit details of this event. This results in nodes receiving several messages containing the same data from different nodes. Figure 3 illustrates the overlap problem. Nodes P and Q both monitor geographic region B.

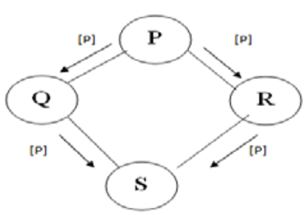


Figure 2: Implosion problem

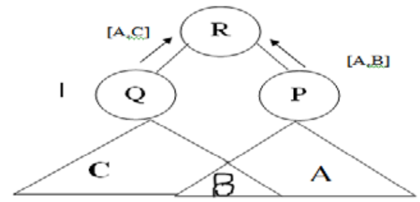


Figure 3: Overlap problem

2.2 GOSSIPING

The Gossiping protocol is based on the flooding protocol [4]. Instead of broadcasting each packet to all the neighbors the packet is sent to a single neighbor chosen at random from a neighbor table. Having received the packet the neighbor chooses another random node to send to. This can include the node which sent the packet. This continues until the packet reaches its destination or the maximum hop count of the packet is exceeded. Gossiping avoids the implosion problem experienced by flooding as only one copy of a packet is in transit at any one time. However the protocol does take a long time to deliver a packet to its destination as the hop count can become quite large due to the protocols random nature.

2.3 SPIN

The SPIN family of protocol is an enhancement of the flooding protocol which is based on data-centric routing [5]. Classic flooding as has three problems: implosion, overlap and resource blindness. In order to overcome the problems of implosion and overlap the SPIN family of protocol use a 3 way negotiation or 3 way Handshaking Protocol before sending data. When a node detects an event it advertises (ADV) the event by transmitting a description of the event. This avoids transmitting the full details of the event. This advertisement is picked up by neighboring nodes and if they are interested in the data they reply requesting the data (REQ). When the original node receives a request it sends the data to the requesting node. The receiving node will then repeat the process by advertising the data. This prevents nodes from receiving duplicate packets as data is only sent when requested. Also as data is

ISSN 2229-5518

described in the advertisement message the problem of overlap can be overcome by checking to see if the node has already received similar data relating to that event. The protocol described above is SPIN-1. SPIN-2 is an extension of SPIN-1 which attempts to overcome the resource blindness problem. Before taking part in the above protocol nodes poll their resources. If their resources fall below a threshold the node will not send or relay data packets.

3. DIFFERENT DATA DISSEMINATION STRATEGIES

Authors in [6] have classified Data Dissemination strategies into four major categories based on basis of operations: Push-based strategy, On-demand (or pull-based) strategy, hybrid strategy and data allocation over multiple broadcast channels.

3.1 Push Based Strategy: In this strategy sensory data is pushed from source nodes to sink nodes following multi-hop routing [7]. And thus query result can be retrieved from sink nodes without communication cost. The push-based Data Dissemination protocol is very efficient when the query rate is relatively higher compared to data acquisition frequency as it can reduce the query routing cost to zero. However, this protocol has some disadvantages: (1)The method trades off communication cost in storage phase to guarantee efficient query in the later steps. (2) All sensory data must be broadcasted to sink node using multi-hop routing [8]. Naturally, the neighbor nodes of sink nodes will undertake more data delivery task than other sensor nodes, and then the hotspots will be formed. These nodes must lose their effectiveness in advance because of energy load unbalance. Therefore system robustness and stability will be hard to be effected.

Flat Broadcast- Flat broadcast is the simplest scheme for data scheduling. Here all data items are broadcast in a round robin manner. The access time for every data item is the same, i.e., half of the broadcast cycle. This scheme is simple, but its performance is poor in terms of average access time when data access probabilities are skewed.

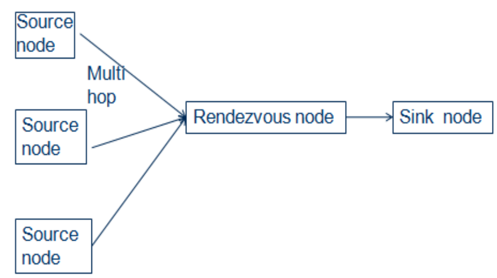
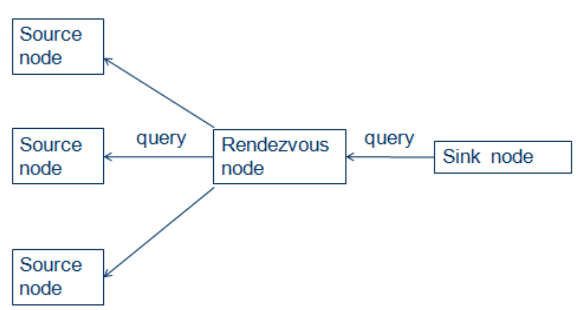


Figure 4: Push Based Strategy

Broadcast Disks- Data items are assigned to different *logical* disks so that data items in the same range of access probabilities are grouped on the same disk. Data items are then selected from the disks for broadcast according to the relative broadcast frequencies assigned to the disks.

3.2 Pull Based Strategy: The pull-based Strategy is opposite to push-based Strategy. The source nodes will never deliver the sensor data voluntarily. Instead, they store data at home and wait for query passively. On the contrary, the consumer nodes broadcast query demands to source nodes throughout the network on their own initiatives. Obviously, communication cost takes place only when it is needed, so the method can be applied to the situation where the data production rate is higher than data query rate. The disadvantage is that even though some source nodes have no related data with query, they have to participate into data delivery.



3.3 Push Pull (Hybrid) Strategy: Another new approach called hybrid broadcast, is to combine push-based and on-demand techniques so that they can complement each other. It introduces the combination between consumer node and source node. In the first phase, source nodes get the storage location and then transfer sensory data to rendezvous nodes closest to location. Then, consumer nodes can directly transmit query to rendezvous node using same regulations. In this way, queries flooding problem can be avoided efficiently.

4. Comparison of Data Dissemination Protocols

This table briefly compares some of the important Data Dissemination protocols [2].

Table 1 – Comparison of Data Dissemination Protocols

Protocol	Full Form	Introduction	Advantage	Disadvantage
DD	Directed Diffusion	Data centric and application aware paradigm	saves network energy & increase in network lifetime	Not suitable for continuous data delivery
LOHD	Location Oblivious Hybrid Data Dissemination	incorporates enhanced PUSH and PULL	Sink node location not required	Increased overhead due to flooding
TTDD	Two- Tier Data Dissemination	Based on decentralized architecture and uses grid structure.	Suitable for multiple mobile links	Unexpected dissemination node failure
BPP	Balancing Push Pull	combines push & pull for information dissemination and gathering	Suitable for continuous & query based data delivery	No search algorithm used
SOM	Solution Mapping on Broadcast and On-demand Channels	composite and generic algorithm	Bandwidth Efficient	End-to-End delay not considered
CORD	*Core Based Reliable Data Dissemination*	is motivated by the goal of reducing the energy consumption	Highly energy Efficient	No bandwidth Saving
RTDD	*Real Time Data Dissemination*	consists of 3 steps: normal routing, projection routing & overhearing of MS	Effective for Real Time Data Dissemination	Fails for fast varying mobile sinks

ISSN 2229-5518

5. Existing work for efficient Data Dissemination

5.1 Tree based approach

Suraj et al. [9, 24] proposed an energy-efficient data dissemination protocol which generates a tree T from the sensor network. It can be represented as a graph $G(V, E)$ where V are the sensor nodes and E are the links between them. The tree construction is independent of the sink position. This method reduces the traffic and prolongs the lifetime of the network. There are two categories of the nodes in the network; one is relay node (RN) and the other is non relay node ($non\text{-}RN$). The relay node is responsible to relay the data from the nodes to its next relay node. The non relay node can only communicate its data to a relay node. So it is a unidirectional communication between non relay and relay nodes. However, the communication is bidirectional between two relay nodes. The tree topology changes when the role of the node changes from relay to non relay or from non relay to relay node. To rotate the responsibility of the relay node, each node's residual energy is considered.

5.2 Based on the need for robustness & scalability

Intanagonwiwat et al. [10] says in favor of designing localized algorithms and present directed diffusion as a set of abstractions that describe the communication patterns underlying such algorithms. The design features differ from traditional wireless networks and are data-centric and application-specific. The authors propose "Directed Diffusion" to be used as an abstraction to model the communication patterns of localized algorithms. The data that each sensor generates is characterized by a number of attributes. Other sensors that are interested in a certain type of data, disseminate this interest to the network (in the form of attributes and degree of interest).

5.3 Based on Geographic Routing

Zorzi and Rao [14] proposed Geographic Random Forwarding (GeRaF) protocol, which uses geographic routing where a sensor acting as relay is not known *a* priori by a sender. There is no guarantee that a sender will always be able to forward the message toward its ultimate destination, that is, the sink. This is the reason that GeRaF is said to be best-effort forwarding. GeRaF assumes that all sensors are aware of their physical locations, as well as that of the sink.

5.4 Based on global energy dissipation

W.R. Heinzelman et al. [15] proposed a 2-level hierarchical routing protocol (LEACH) which attempts to minimize global energy dissipation and distribute energy consumption evenly across all nodes. This is achieved by the formation of clusters with localized coordination, by rotating the high-energy cluster heads and by locally compressing data.

5.5 Based on the need of global coordinate system

D. Braginsky et al. [16] proposed Rumor Routing as a logical compromise between query and event flooding. With Rumor Routing paths (possibly multiple and non-optimal) are created leading to each event. Whenever a query is generated it is sent on a random walk until it crosses one of the paths leading to the event of interest. It is possible that the query will never cross such a path, in which case query flooding can be used as a last resort. The authors use the heuristic of two lines intersecting in a bounded rectangular region to indicate the plausibility of their solution. The main focus of this paper is the method for setting up paths to an event.

5.6 Based on mobile network environments & nomadic clients

Askoy et al. [17] have presented a large-scale low overhead on-demand broadcasting model called R*x*W (Requests time Wait). In R*x*W, at each broadcast tick, the server chooses an item with the highest value of (R * W) where R is the number of outstanding requests and W is the waiting time for the first request. The entry for this data item is then removed from the queue that keeps track of the number of requests and earliest request time for data. The algorithm makes no assumptions regarding access probabilities of data items. However, the size of the queue is equal to the size of the database, therefore, large databases will require a significant overhead in terms of time to find the highest value of R * W and space to store frequency and time.

5.7 Based on mobile network environments & nomadic clients

Xuan et al. [18] have proposed two on demand broadcast scheduling strategies, which consider deadlines attached to requests to decide the next item to be broadcast. In the first strategy, the server always broadcasts a request with the earliest deadline first (EDF) and every request is scheduled once no matter how frequently the same request is encountered. The second strategy is called EDF-batch, in which the server broadcasts an item according to EDF but after broadcasting, it removes the other entries for the same request.

5.8 Based on mobile network environments & nomadic clients

Datta et al. [22] have proposed the protocols that dynamically change the contents of broadcast according to client requests. In these protocols the broadcast data and index are organized using the (1,m) indexing strategy [20]. The server decides the data item to be included in the broadcast on the basis of priority, which is given by (IFN*PF), where IF is Ignore Factor, PF is the Popularity Factor, and N is an adaptive scaling factor. PF makes sure that most popular data items are included in the broadcast and IF makes sure that less popular long neglected data items are also broadcast. In the Constant Broadcast Size protocol, the broadcast size is fixed, and after each broadcast cycle, the server calculates the priority of data items, sorts them in descending order of priority and adds them to the broadcast until it is full. In the Variable Broadcast Size protocol, all the items with PF > 0 are added to the broadcast set.

5.9 Based on multiple servers and multiple clients

Le Gruenwald et al. [21] proposed Adaptive Broadcast Scheduling Algorithms. These algorithms consider three issues in scheduling broadcasts: energy consumed by mobile clients, energy consumed by mobile servers, and real-time constraints on client requests. Simulation experiments conducted to study the performance of the proposed algorithms show that the popularity-based algorithms provide better broadcast hit ratio, access time, and client energy consumption, but more energy consumption than the EWMA based algorithms. The addition algorithms provide better broadcast hit ratio and access time but they are not energy-efficient for servers and clients. The replacement algorithms provide better energy consumption for clients and servers but do not give good broadcast hit ratio and access time.

The following table summarizes the work done by researchers to achieve efficient way of data Dissemination.

Table 2 - Comparative Study of Existing work

S.No.	Paper Title	Author	Publishing Year	Features	Simulator	Limitations
1	Data Dissemination Protocol for Mobile Sink in Wireless Sensor Networks	Suraj Sharma and Sanjay Kumar Jena	2014	TEDD is an energy-efficient data dissemination protocol with mobile sink. Initially, it creates the tree in the network with a root node	NS-2	The higher sink speed increases the frequency of link failure, which causes data loss.
2	Directed Diffusion: A Scalable and Robust Communication	C. Intanagonwiwat and R. Govindan and D. Estrin	2000	This approach decouples data from the sensor that produced it and unique identification of nodes is of secondary Importance	Simple Simulator	design is difficult & localized algorithms tend to be sensitive in the choice of parameter values.
3	Energy-efficient communication protocol for wireless microsensor networks	W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan	2000	attempts to minimize global energy dissipation & distribute energy consumption evenly across all nodes.	Matlab	Lossy compression is used to reduce overall energy dissipation of the system.

ISSN 2229-5518

4	Rumor routing algorithm for sensor networks	D. Braginsky, D. Estrin	2002	It is a logical compromise between query and event flooding. The goal of the algorithm is for the latter case to be rare	LecsSim	Efficient method for allowing far away nodes to efficiently query events, not closed-by nodes.
5	Scheduling for Large Scale On-Demand Data Broadcasting	Askoy. D. and Franklin M.	1998	large-scale low overhead on-demand broadcasting model(RxW). It makes no assumptions regarding access probabilities of data items.	Simple simulator	Could eventually become bottleneck for a very large applications.
6	Energy-Efficient Data Broadcasting in Mobile Ad-Hoc Networks	Le Gruenwald, Muhammad Javed, Meng Gu	2002	applicable to the case where multiple servers and multiple clients exist in a certain area. The (1, m) indexing scheme is used in this algorithm	Awesim	replacement algorithms provide better energy consumption for clients and servers but do not give good broadcast hit ratio and access time.

6. Conclusion

In this paper Data Dissemination strategies are analyzed and discussed briefly. Also the comparison between various Data Dissemination protocols is done which has been proposed by various researchers in past few years but none of them ensure overall efficiency. Each of the protocols discussed performed well in some cases, but showed certain drawbacks in others. It is very difficult to find a perfect Data Dissemination strategy suitable for all application requirements. Their application effectiveness also needs further evaluation on the hardware platform of sensor nodes.

7. About the Author

Kanak Manjari is currently pursuing masters degree program in Computer Science & Engineering from Birla Institute of Technology, Mesra (Noida extension) , Noida – 201301,U.P.,India.
Email:- kanakmanjari@gmail.com

8. References

1. www.wikipedia.org/wiki/Wireless sensor network Access on 20th June 2012.

2. Aayushi Mahajan , Suhaib Ahmed , Swastik Gupta, Ashish Suri . "Comparative Analysis of Different Data Dissemination Strategies in Wireless Sensor Networks"

International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 4, April 2014.

3. C.Shiva Ram Murthy and B.S.Manoj. "Ad Hoc Wireless Networks".

4. Kemal Akkaya , Mohamed Younis ,"A survey on routing protocols for wireless sensor", networks *Ad Hoc Networks 3* (2005), pp 325–349.

5. Ankit Joshi, Lakshmi Priya.M. "A Survey of Hierarchical Routing Protocols in Wireless Sensor Network." In Second International Conference on Information Systems and Technology, pp. 67.

6. Fang Hongping, Fang Kangling, "overview of Data Dissemination strategies in wireless sensor networks", in International conference on E- Health Networking, Digital Ecosystem & Technologies,2010.

7. Xin Liu, Qingfeng Huang and Ying Zhang, "Balancing Push and Pull for Efficient Information Discovery in Large-Scale Sensor Networks", IEEE Transactions on Mobile Computing, vol. 6, No. 3, March 2007.

8. Jiun-Long Huang, Wen-Chih Peng and Ming-Syan Chen, "SOM: Dynamic Push-Pull Channel Allocation Framework for Mobile Data Broadcasting", IEEE Transactions on Mobile Computing, vol. 5, No. 8, August 2006.

9. Suraj Sharma and Sanjay Kumar Jena, "Data Dissemination Protocol for Mobile Sink in Wireless Sensor Networks", Published 27 April 2014.

10. C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks. In *ACM MobiCom*, 2000.

11.O.Kasten,"EnergyConsumption",*www.infethz.ch/kasten/researchlbathtub/energyconsumption.html*.

12. G. Xing, C. Lu, R. Pless, and Q. Huang, "On greedy geographic routing algorithms in sensing-covered networks", Proceedings ACM MobiHoc'04, Tokyo, Japan, May 2004, pp. 31-42.

13. Jamal Al-Karaki, and Ahmed E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey", IEEE Communications Magazine, vol 11, no. 6, Dec. 2004, pp. 6-28.

14. M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Mutlihop performance", *IEEE Transactions on mobile Computing*, vol. 2, no. 4, Oct.-Dec. 2003, pp. 337-348.

15. W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *IEEE Hawaii International Conference on System Sciences*,2000.

16. D. Braginsky and D. Estrin. Rumor routing algorthim for sensor networks. In *ACM International Workshop on Wireless Sensor Networks and Applications*, 2002.

17. Askoy, D. and Franklin, M., "Scheduling for Large Scale On-Demand Data Broadcasting", Proceedings of the 12th International Conference on Information Networking, pp. 656- 659, Jan 1998.

18. Xuan, P., Sen, S., Gonzalez, O., Fernandez, J., and Ramamritham K., "Broadcast on Demand: Efficient and Timely Dissemination of Data in Mobile Environments", 3rd IEEE Real-Time Technology Application Symposium, 1997.

19. Imielinski, T., Viswanathan, S., and Badrinath, B. R., "Power Efficient Filtering of Data on Air", Proc of 4th Intl Conference on Extending Database Technology, Cambridge, March 1994, pp. 245-258.

20. Imielinski, T., Viswanathan, S., and Badrinath, B. R., "Data on Air: Organization and Access", IEEE Transactions on Knowledge and Data Engineering, Vol. 9, No. 3, May 1997, pp. 353-372.

21. Le Gruenwald , Muhammad Javed, Meng Gu, "Energy-Efficient Data Broadcasting in Mobile Ad-Hoc Networks", The University of Oklahoma, School of Computer Science, Norman, OK 73019.

22. Datta, A., VanderMeer, D. E., Kim, J., Celik, A., and Kumar, V., "Adaptive Broadcast Protocols to Support Efficient and Energy Conserving Retrieval from Databases in Mobile Computing Environments", A TimeCenter Technical Report, University of Arizona, April, 1997.

23. http://123doc.org/document/929287-tai-lieu-so-tay-cua-cac-mang-khong-day-va-dien-toan-di-dong-p11-ppt.htm

24. "Data Dissemination Protocol for Mobile Sink in Wireless Sensor Networks", http://www.hindawi.com/journals/jcengi/2014/560675/

ISSN 2229-5518
25. "Research Issues for Data Communication in Mobile Ad-Hoc Network Database Systems", http://www.researchgate.net/publication/2864883_Research_Issues_for_Data_Communication_in_Mobile_Ad-Hoc_Network_Database_Systems

26.www.ijarcce.com, http://www.ijarcce.com/upload/2014/april/Comparative%20Analysis%20of%20Different%20Data%20Dissemination%20Strategies%20in%20Wireless%20Sensor%20Networks.html

ISSN 2229-5518