ISOLATION CONSERVES MULTI –KEYWORD GRADED UP ENCODE CLOUD DATA

Chaitra c[1], Dr. M.V vijaykumar[2]

[1] *M. Tech Student, Department of Computer Science & Engineering, Dr. Ambedkar institute of Technology,Bangalore-560056*

[2] *Professor and co-coordinator, Department of Computer Science & Engineering, Dr. Ambedkar institute of Technology,Bangalore-560056*

---***---

Abstract - *Security is the important aspect in cloud computing, there could be multiple users outsourcing there data to the cloud, so it is important to maintain the security of the outsourced data. Value and the relative index of the attribute in orders are needed to be hided for confidentiality. Preserving the privacy for multi-keyword search for the ciphered data stored on the cloud. For secured utilization of data, strict privacy prerequisites have to be formed, to make search efficient much possible matches will be listed in the document for the particular query. We additionally implement similarity under inner product to measure the comparability.*

***Key Words*: Cloud computing, Privacy preservation, Ranked keyword search, Multi owners, Ciphered text.**

1. INTRODUCTION

Providing service of better quality from intense and registering is possible by cloud computing. Into the cloud the data proprietor outsources there information and this is one of the application of the cloud computing called database as service. To maintain the security of the information to be stored on the cloud it has been converted to ciphered data. Secured searching of the ciphered text is the main concept of research [3], defining and solving the secured searching of the content in the ciphered text is the first task. The concept of searching the ciphered text is proposed where in keyword-based is performed on the ciphered text. Searching in the ciphered data is further developed by [4] [6].

Ranked multi-keyword searching is the improved technique where computation and storage cost is more. Searching data in ciphered form is first defined in [9] and further developed by [16] [13].we propose an effective ordering technique to bolster quicker inquiry assessment than the trifling straight sweep way. As of not long ago, the inquiry answer is figured by assessing the condition for each encoded information point e (pi) in the database. This straight sweep is not adequate for substantial databases. We receive the ball tree file to recover such focuses while pruning however many information focuses as could be expected under the circumstances. Quickly, a ball tree is a parallel tree with the end goal that each non-leaf hub speaks to a ball and has two youngster hubs. Information direct having a place toward a parent goes to the kid ball whose middle is nearer to the information point. All information focuses are just put away at the leaf hubs. To construct such a ball tree, we could continue isolating the information point space into two parcels (left and right tyke) recursively until the quantity of information focuses in some segment is beneath a predefined edge and we make this segment as a leaf hub. We call this edge as “max leaf size” and we will test it in our trial section. The point by point calculation to build the ball tree (i.e., how to scrambled records into balls) could be found and we don't try to bring the precisely same calculation here.

From one perspective, to meet the successful information recovery require, the vast measure of reports request the cloud server to perform result pertinence positioning, rather than returning undifferentiated outcomes. Such positioned look framework empowers information clients to locate the most applicable data rapidly, as opposed to burdensomely dealing with each match in the substance accumulation. Positioned pursuit can likewise carefully take out pointless system activity by sending back just the most important information, which is very alluring in the "pay-as-you-utilize" cloud worldview.

For security assurance, such positioning operation, be that as it may, ought not to release any catchphrase related data. Then again, to enhance the query output precision and in addition to upgrade the client looking knowledge, it is additionally essential for such positioning framework to bolster different catchphrases seek, as single watchword inquiry regularly yields extremely coarse outcomes.

As a typical practice demonstrated by today's web search tools, information clients may have a tendency to give an arrangement of catchphrases rather than just a single as the marker of their pursuit enthusiasm to recover the most important information. What's more, every watchword in the hunt demand can help limit the query output assist. "Organize coordinating", i.e., whatever number matches as could reasonably be expected, is a productive likeness measure among such multi-watchword semantics to refine the outcome significance, and has been broadly utilized as a part of the plaintext data recovery (IR) people group. In any case, how to apply it in the scrambled cloud information look framework remains an exceptionally difficult errand due to intrinsic security and protection snags, including different strict prerequisites like the information protection, the file protection, the catchphrase security, and numerous others.

2. CONTRIBUTION

Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching," i.e., as many matches as possible, to capture the relevance of data documents to the search query. Specifically, we use "inner product similarity", i.e., the number of query keywords appearing in a document, to quantitatively evaluate such similarity measure of that document to the search query. During the index construction, each document is associated with a binary vector as a subindex where each bit represents whether corresponding keyword is contained in the document.

The search query is also described as a binary vector where each bit means whether corresponding keyword appears in this search request, so the similarity could be exactly measured by the inner product of the query vector with the data vector. However, directly outsourcing the data vector or the query vector will violate the index privacy or the search privacy. To meet the challenge of supporting such multikeyword semantic without privacy breaches, we propose a basic idea for the MRSE using secure inner product computation, which is adapted from a secure k-nearest neighbor (kNN) technique and then give two significantly improved MRSE schemes in a step-by-step manner to achieve various stringent privacy requirements in two threat models with increased attack capabilities.

Our contributions are summarized as follows:

1. For the first time, we explore the problem of multikeyword ranked search over encrypted cloud data, and establish a set of strict privacy requirements for such a secure cloud data utilization system.
2. We propose two MRSE schemes based on the similarity measure of "coordinate matching" while meeting different privacy requirements in two different threat models.
3. We investigate some further enhancements of our ranked search mechanism to support more search semantics and dynamic data operations.
4. Thorough analysis investigating privacy and efficiency guarantees of the proposed schemes is given, and experiments on the real-world data set further show the proposed schemes indeed introduce low overhead on computation and communication

3. PROBLEM FORMULATION

3.1 Algorithm

AES is a symmetric square figure. This implies it utilizes a similar key for both encryption and decoding. In any case, AES is very not quite the same as DES in various ways. The calculation Rijndael takes into account an assortment of piece and key sizes
What's more, not quite recently the 64 and 56 bits of DES' square and key size.

The square and key can actually be picked autonomously from 128, 160, 192, 224, 256 bits and need not be the same. In any case, the AES standard expresses that the calculation can just acknowledge a piece size of 128 bits and a decision of three keys - 128, 192, 256 bits. Contingent upon which form is utilized, the name of the standard is changed to AES-128, AES-192 or AES-256 individually.

And also these distinctions AES contrasts from DES in that it is not a feistel structure. Review that in a feistel structure, half of the information piece is utilized to adjust the other portion of the information square and afterward the parts are swapped.
Various AES parameters rely on upon the key length. For instance, if the key size utilized is 128 then the quantity of rounds is 10 though it is 12 and 14 for 192 and 256 bits separately. At present the most widely recognized key size prone to be utilized is the 128 piece key. This portrayal of the AES calculation subsequently depicts this specific execution.

Rijndael was intended to have the accompanying attributes:
- Resistance against every known assault.
- Speed and code minimization on an extensive variety of stages.
- Design Simplicity.

AES works on a 4×4 section real request grid of bytes named the state, albeit a few variants of Rijndael have a bigger square size and have extra segments in the state. Most AES figuring's done in an uncommon limited field.
The key size utilized for an AES figure indicates the quantity of redundancies of change adjusts that change over the info, called the plaintext, into the last yield, called the cipher text.

There are 4 main modules which helps us to solve us the problem easier.
3.2 MODULES
Binary data generation
Data ciphering
Data user access control
Data user query

1. Binary data generation

Information proprietor select the information and make the bit vector for that information. Utilizing that bit vector of the information the parallel information is created. The paired information is the file for the information in the information proprietor. The bit vector is the bytes type of the information in the information proprietor. The bit vector is changed over into the paired information. These bit vector and the twofold information are prepared for the information figuring.

2. Data ciphering

In this segment, we show a formal depiction for the objective issue in this paper. We initially characterize a framework show and a comparing danger demonstrate
.Then the data owner have to encrypt the original data and send it to server. And then encrypt the binary data or the index and send it to server. Service provider did not know about the original content in the data owner. These index are used to refer the data in the service provider. It gives more security in tr he server side, so that the attackers can't use the data. Our system must prevent Server from learning any additional correspondence between plaintext values and ciphertext values except those obtained by prior knowledge. That is, we must protect the plaintext values for any encrypted records or queries from being disclosed to Server.

3. DATA USER ACCESS CONTROL

The client needs information from the server. The client have distinctive options and the client send the inquiry to the server or specialist co-op. Before that the client get the entrance from the information proprietor. For that the client send the insights about him or her to the information proprietor. At that point just the information proprietor gets the data from customer and prepared to send the unscrambling key. the get to control mechanismis utilized to oversee decoding abilities given to clients and the information accumulation can be refreshed regarding embeddings new archives, refreshing records, and erasing reports.
This is a distributed setting where Server is on the remote side and not trusted. More specifically, Owner collects and owns the data R and has all rights to upload, query and encrypt data, and may also grant the query right to authorized users with access control keys. Users is a group of users authorized to post a query and receive the answers. Owner encrypts the data and then uploads it to the Server. Proxy serves a bridge between Users and Server.

4.DATA USER QUERY

The information client inquiry is prepared by the specialist co-op. The specialist organization creates the bit vector for the question from the customer. At that point the specialist co-op changes over the bit vector into parallel information. Specialist organization finds thecomparable information from the list. What's more, send the encoded information to the information proprietor. At that point the customer decodes the got information by the key from the information proprietor.
A query from Users will go though the trusted Proxy, which encrypts the query and submits the query to Server. Server computes and returns the answer to Proxy. Proxy then decrypts the answer, and returns the answer to Users.

For example, Owner is a hospital, who outsources patient records to the cloud, and Users are various medical research labs, who post queries to retrieve patient records of interests.

4. SYSTEM ARCHITECTURE

This diagram tells us about the behaviour of the system the data accepted from the dataowner will be indexed and then will be converted into ciphered form and will be stored on the cloud if any of the customer want to access this data he has to first send his details to the datauser then the dataowner will verify the details and send the key to the customer ,and he will send the query request for the cloud server the cloud server will provide the data in the encrypted form .

This data the customer will receive and by using the key he will decrypt it.

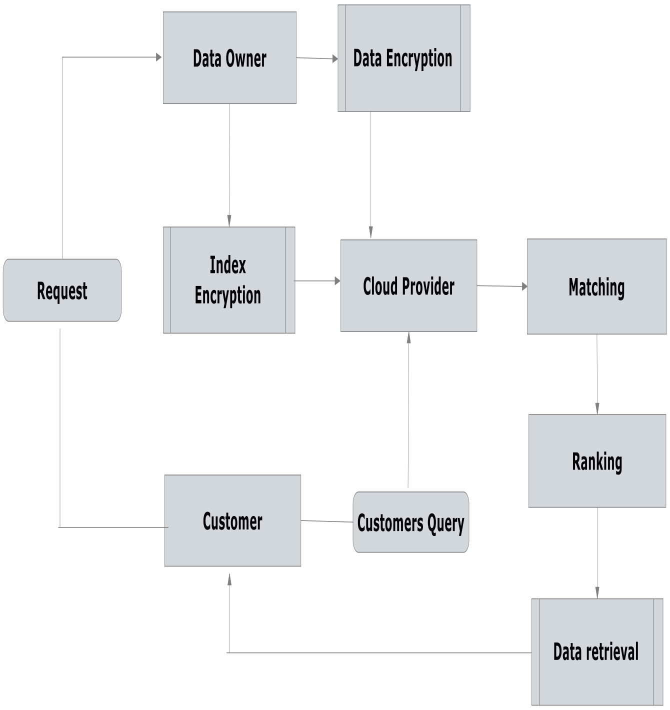


Fig-4.2:System Architecture

5.RESULT

The project result shows how the data owner uploads his data to the cloud and how the data will be accesse by the customer by maintaing security.

The result will be shown in the following screenshots.

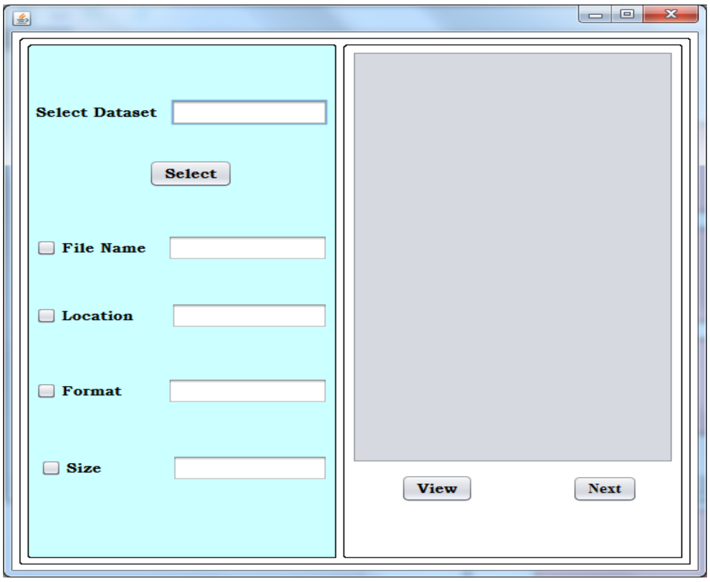
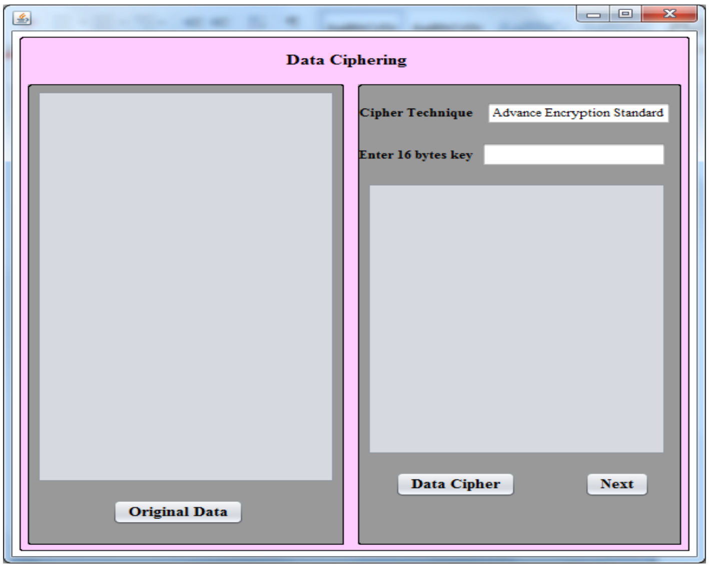


Fig-5.1:Selecting Dataset

In the above screenshots the dataset will be selected and the filename,location,format and size can be seen. As soon as the application is started we need to select the dataset from where we need to access the data and this dataset will contain the required information that customer is needed

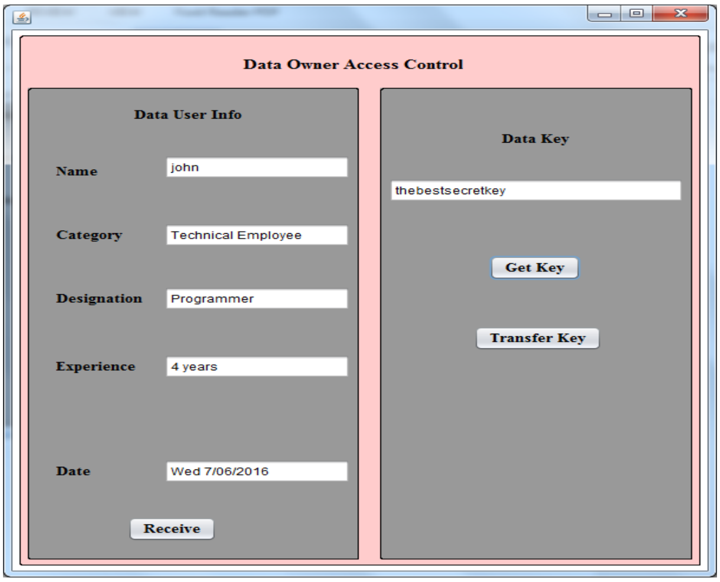


Cipher Technique: Advance Encryption Standard

Enter 16 bytes key

Fig -5.2:Encryption of Data

Here the data owners data will be converted into ciphered text to upload into the cloud.The data is converted into encryted data for maintaining the privacy.The data will be converted into bit vector and then converted to the binary format this will not to be in the readable format, it can be converted into readable form only by decrypting the data using the key provided by the data owner.



Name: john

Category: Technical Employee

Designation: Programmer

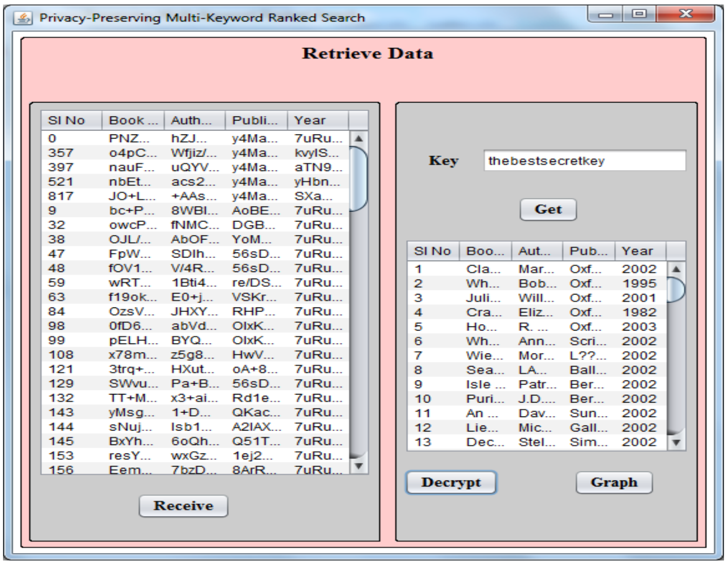
Experience: 4 years

Date: Wed 7/06/2016

Data Key: thebestsecretkey

Fig-5.3: Getting key from the data owner

If the customer wants to access the data present on the cloud, first they need to obtain the permission from the data owner by sending his details. After sending the details the owner will verify the details and will send that key to the customer who has requested for the key.



Sl No	Book ...	Auth...	Publi...	Year
0	PNZ...	hZJ...	y4Ma...	7uRu...
357	o4pC...	Wfjiz/...	y4Ma...	kvylS...
397	nauF...	uQYV...	y4Ma...	aTN9...
521	nbEt...	acs2...	y4Ma...	yHbn...
817	JO+L...	+AAs...	y4Ma...	SXa...
9	bc+P...	8WBl...	AoBE...	7uRu...
32	owcP...	fNMC...	DGB...	7uRu...
38	OJL/...	AbOF...	YoM...	7uRu...
47	FpW...	SDlh...	56sD...	7uRu...
48	fOV1...	V/4R...	56sD...	7uRu...
59	wRT...	1Bti4...	re/DS...	7uRu...
63	f19ok...	E0+j...	VSKr...	7uRu...
84	OzsV...	JHXY...	RHP...	7uRu...
98	0fD6...	abVd...	OlxK...	7uRu...
99	pELH...	BYQ...	OlxK...	7uRu...
108	x78m...	z5g8...	HwV...	7uRu...
121	3trq+...	HXut...	oA+8...	7uRu...
129	SWvu...	Pa+B...	56sD...	7uRu...
132	TT+M...	x3+ai...	Rd1e...	7uRu...
143	yMsg...	1+D...	QKac...	7uRu...
144	sNuj...	Isb1...	A2IAX...	7uRu...
145	BxYh...	6oQh...	Q51T...	7uRu...
153	resY...	wxGz...	1ej2...	7uRu...
156	Eem...	7bzD...	8ArR...	7uRu...

Key: thebestsecretkey

Sl No	Boo...	Aut...	Pub...	Year
1	Cla...	Mar...	Oxf...	2002
2	Wh...	Bob...	Oxf...	1995
3	Juli...	Will...	Oxf...	2001
4	Cra...	Eliz...	Oxf...	1982
5	Ho...	R. ...	Oxf...	2003
6	Wh...	Ann...	Scri...	2002
7	Wie...	Mor...	L??...	2002
8	Sea...	LA...	Ball...	2002
9	Isle ...	Patr...	Ber...	2002
10	Puri...	J.D....	Ber...	2002
11	An ...	Dav...	Sun...	2002
12	Lie...	Mic...	Gall...	2002
13	Dec...	Stel...	Sim...	2002

Fig-5.4: Retrieving the data

As soon as the customer or the data user gets the key from the data owner, the data user will send the request for the data present on the cloud to the cloud server by specifying the multiple keywords. Here in the screenshot we have taken book name, author, publisher, year as the multiple keywords.

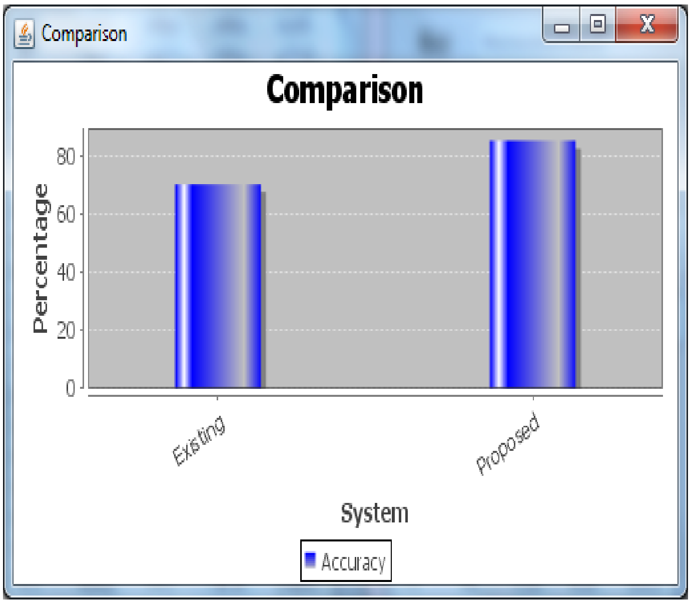


Fig-5.5: The graphical representation of the accuracy

The above graph represents the bar graph representation of the accuracy of the existing system and the proposed system. There are 0.5 increases in the accuracy when compared to existing system.

6. CONCLUSION

As we know maintain large amount of data in our own database is very difficult because it consumes lots of space and also maintenance is overhead. To overcome this, uploading of data to the cloud is must. After storing the data the privacy of the data has to maintained by cloud server because there could be many sensitive data that will be stored on the cloud. We consider the administration of noting the class of range inquiry look over numerical information and propose an information encryption plan to address these necessities. In this paper, surprisingly we characterize and tackle the issue of multi-watchword positioned seek over encoded cloud information, and build up an assortment of protection necessities. Among different multi-catchphrase semantics, we pick the proficient likeness measure of "coordinate matching", i.e., whatever number matches as could be expected under the circumstances, to viably catch the pertinence of outsourced reports to the inquiry watchwords, and utilize "internal item similitude" to quantitatively assess such comparability measure. For addressing the difficulty of supporting multi-watchword semantic without protection breaks, we propose an essential thought of MRSE utilizing secure internal item calculation. At that point, we give two enhanced MRSE plans to accomplish different stringent security necessities in two distinctive danger models. We likewise examine some further improvements of our positioned look instrument, including supporting more pursuit semantics, i.e., TF IDF, and dynamic information operations. Intensive examination researching security and effectiveness assurances of proposed plans is given, and investigations on this present reality information set demonstrate our proposed plans present low overhead on both calculation and correspondence.

7. REFERENCE

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, “Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data,” Proc. IEEE INFOCOM, pp. 829-837, Apr, 2011.
[2] D. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Proc. IEEE Int. Symp. Security Privacy, Nagoya, Japan, Jan. 2000, pp. 44–55.
[3] E. Goh. (2003). Secure indexes [Online]. Available: http://eprint.iacr.org/
[4] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in Advances in Cryptology-Eurocrypt 2004, Springer, 2004, pp. 506–522.
[5] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, “Secure ranked keyword search over encrypted cloud data,” in Proc. IEEE Distrib.Comput. Syst., Genoa, Italy, Jun. 2010, pp. 253–262
[6] P. Xu, H. Jin, Q. Wu, and W. Wang, “Public-key encryption withfuzzy keyword search: A provably secure scheme under keyword guessing attack,” IEEE Trans. Comput., vol. 62, no. 11, pp. 2266–2277, Nov. 2013.
[7] Z. Xu, W. Kang, R. Li, K. Yow, and C. Xu, “Efficient multikeyword ranked query on encrypted data in the cloud,” in Proc. IEEE 19th Int. Conf. Parallel Distrib. Syst.Singapore, Dec.2012, pp. 244–251