

---



# Enhancing the Security of Yioop Discussion Board

Masters Defense



Prajna Gururaj Puranik

# Yioop discussion board



## Yoop! - My Groups

Filter Go

**Demo** (1 posts, 1 threads)  
Last Post: [hola](#)

**FinalTest** (9 posts, 3 threads)  
Last Post: [--jsdha](#)

**Public** (500 posts, 500 threads)  
Last Post: [ad\\_program\\_terms Wiki Sayfasını Oluşturdu.](#)

User Home Page

## FinalTest:Talk

Start New Thread

2023-05-16

**user3** [jsdha](#) (2 posts, 33 views) .

**user3** [User3 thread](#) (4 posts, 14 views) .

2023-05-15

**user1** [New](#) (3 posts, 8 views) .

**user1** [user1 joined FinalTest!](#)

- [Blog](#) - [Privacy](#) - [Terms](#) - [ThisSiteBot](#) - [Developed at SeekQuarry](#) -

(c) This Site - [This Search Engine](#)

Group View

## FinalTest:Talk:New

2023-05-15

**user1** [New.](#)  
new user1

**user1** hello

**user2** hi

Comment

Thread View

---



# Importance of security



## **Protection of user data**

Safety of sensitive user  
information

## **Preserving user trust**

Instill confidence in users

## **Safeguarding against attacks**

Web portals are vulnerable to  
attacks

## **Compliance with regulations**

Comply with laws like CCPA and  
GDPR

## **Existing features**

Differential Privacy

Encrypted groups

External database



# Newly added Security Features

## Extending Differential Privacy

Hide number of users

## Flag and Moderation

Flag button and Moderation group


## Secret Sharing

Restrict access to encryption key

---



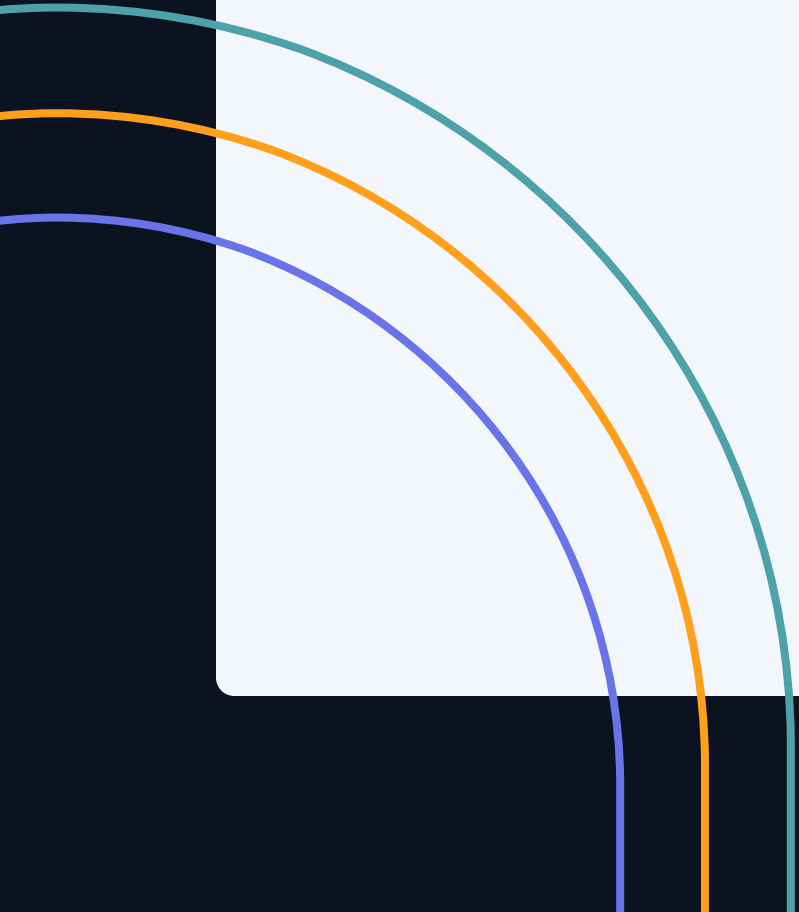
# Security considerations

- ❖ Security in social media sites similar to Yioop
  - ❖ Study by *Social Network Analysis and Mining* journal [1] → 60% of respondents support use of flags
  - ❖ Moderation success in Reddit → banned several subreddits in 2015
  - ❖ Banning posts saw an 80% decrease in hate speech usage [2]
- 

---



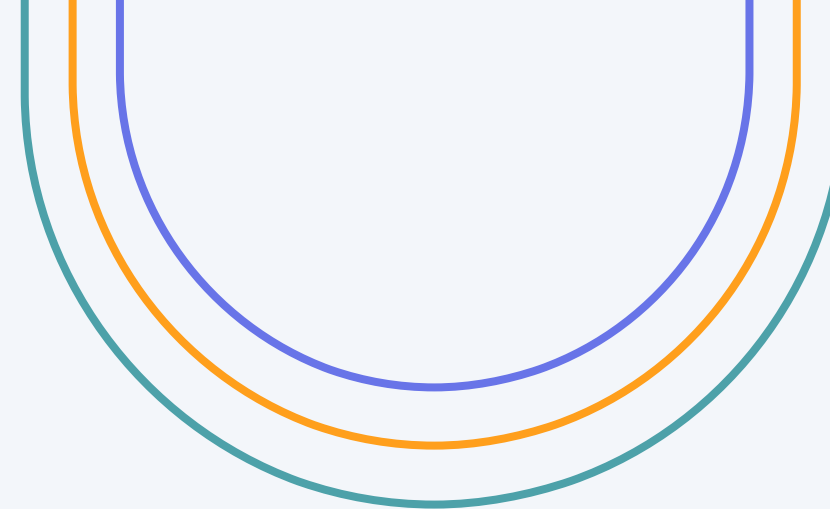
# Security considerations

- ❖ The California Consumer Privacy Act (CCPA) and the General Data Protection Regulation (GDPR) → keep sensitive user data safe and private [3]
  - ❖ Compelled businesses to be more open about how they gather user data and to give users greater control over it, which has resulted in the adoption of better privacy measures on the internet
- 

# Differential Privacy

- ❖ Statistical attacks → Extraction of private information by analyzing patterns/statistical properties of the data.
- ❖ Differential privacy → Mathematical framework for protecting users' privacy in datasets.
- ❖ Adds noise so that individual data points cannot be distinguished → causes statistical attacks to fail
- ❖ Yioop uses  $\epsilon$ -differential privacy →  $\epsilon$  is the privacy parameter
- ❖ Extended to mask the number of users in the group
- ❖ Benefits:
  - User anonymity
  - Avoiding Bias or Prejudice
  - Protection against Targeted Attacks

# Differential Privacy



❖ 3 UI instances where the group user count is displayed:

■ Group owner - Edit Group

Members: [3 users] With Selected ▾

	<a href="#">Name</a>	<a href="#">Join Date</a>	<a href="#">Status</a>	<a href="#">Action</a>	
<input type="checkbox"/>	root	10/19/2022	Active	Owner	Delete
<input type="checkbox"/>	user1	10/19/2022	Active	<a href="#">Ban</a>	<a href="#">Delete</a>
<input type="checkbox"/>	user2	10/19/2022	Active	<a href="#">Ban</a>	<a href="#">Delete</a>

[\[Invite More Users\]](#)

Members: [1 users] With Selected ▾


	<a href="#">Name</a>	<a href="#">Join Date</a>	<a href="#">Status</a>	<a href="#">Action</a>	
<input type="checkbox"/>	root	10/19/2022	Active	Owner	Delete
<input type="checkbox"/>	user1	10/19/2022	Active	<a href="#">Ban</a>	<a href="#">Delete</a>
<input type="checkbox"/>	user2	10/19/2022	Active	<a href="#">Ban</a>	<a href="#">Delete</a>

[\[Invite More Users\]](#)

■ Group owner - Edit Members

## Group Information

**Name:** test

**Owner:**  [root](#)

**Register:** Invite Only

**Access:** Members Can Read

**Voting:** No Voting

**Post Lifetime:** Never Expires

**Encryption:** Enabled

**Members:** 5

## Edit Group ?

**Name:**

**Owner:**

**Register:**

**Access:**

**Voting:**

**Post Lifetime:**

**Encryption:**

**Feed:**

**Members:**

■ Group user - Manage Group



# Encrypted Groups

- ❖ Yioop allows creation of encrypted groups
- ❖ Title and description of threads are encrypted and stored

TITLE	DESCRIPTION
%D8%A3%D8%B1%D8%A8%D8%B9%D9%85%D8%A7...	مناقشة صفحة في هذا الموضوع Mon.08 May 2023 17:03:50 -0700

- ❖ Key stored in a separate database
- ❖ GROUP\_ID attribute is used to access the key

KEY_ID	TYPE_ID	KEY_NAME
1	13	BE2jzxWra/v/LaBoN0YihuU7ytUKCdSWWErzSZm9TKc=

### Manage Groups

#### Create Group ?

Name:

Register:

Access:

Voting:

Post Lifetime:

Encryption:

# Secret Sharing

- ❖ GROUP\_ID is stored in public database - Not safe to access key
- ❖ So use secret sharing scheme to securely compute keys
- ❖ Sharing a secret among a group of participants in a way that no individual can deduce the secret by themselves
- ❖ Linear secret sharing is used here, where a line is used to generate shares that will be distributed to the users
- ❖ Note: Group owner adds users to the group

**Members:** [\[1 users\]](#) With Selected ▾

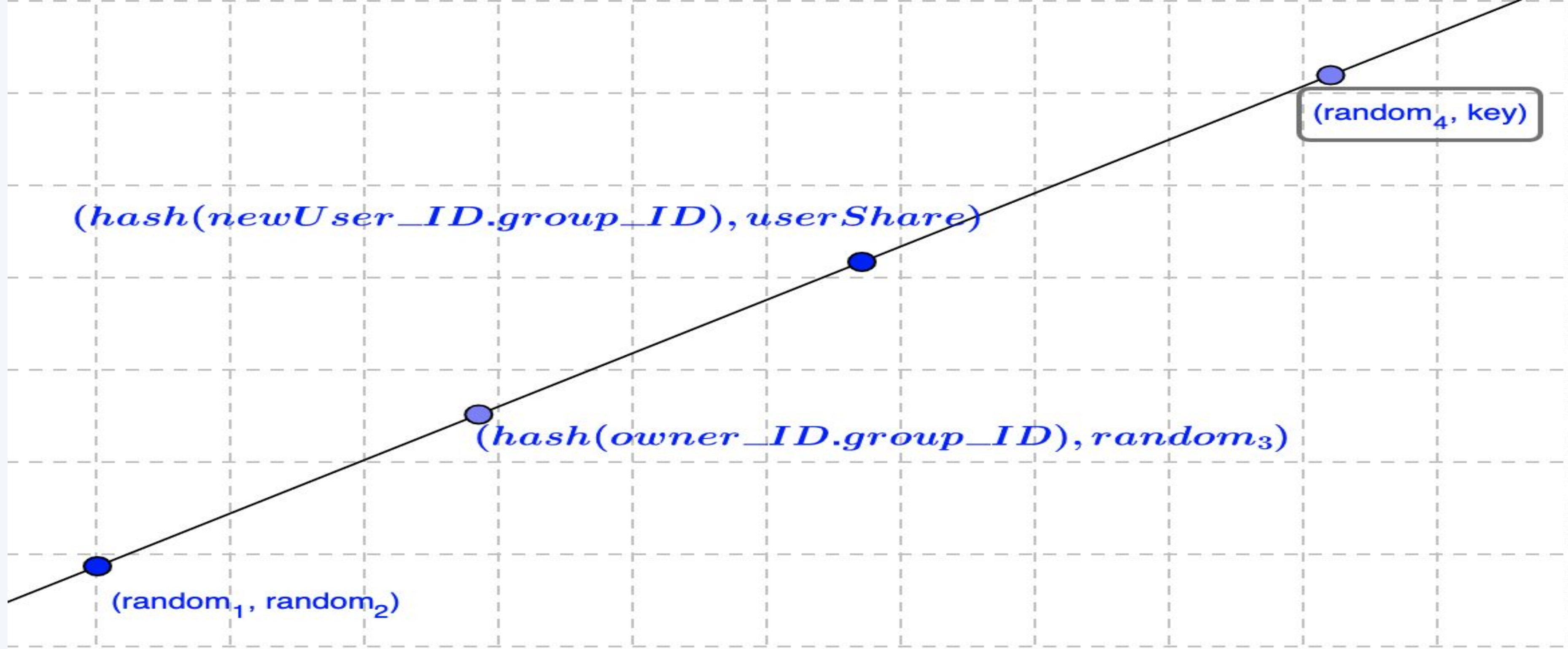
	<a href="#">Name</a>	<a href="#">Join Date</a>	<a href="#">Status</a>	<a href="#">Action</a>
<input type="checkbox"/>	user1	05/24/2023	Active	Owner Delete

[\[Invite More Users\]](#)

**Invite Users to Group** ✕

Name:

Usernames (space/comma delimited)



(random1, random2) from private DB

GROUP_KEYS		
GROUP_ID	RANDOM_1	RANDOM_2

OWNER\_ID, GROUP\_ID, random3 and random4 from public DB

public_db			
GROUP_ID	USER_ID	USER_HASH	SHARE

Use X = random4 to get Y value

public_db	
GROUP_ID	RANDOM_4

# Flagging

- ❖ Flagging → marking content that violates guidelines
- ❖ Flagged posts sent to moderators for review
- ❖ Benefits:
  - Early detection of harmful content
  - Transparency and user empowerment
- ❖ Considerations:
  - A user cant flag a post more than once
  - Appropriate dialog boxes to confirm choice to flag
  - Appropriate message if threshold reached
  - Care taken ensure encrypted groups have masked flag values

# Flag Feature



localhost:8080 says  
Are you sure you want to flag this post?

Cancel OK

11:33 am

**user1** [New.](#)  
new thread

**user1** flag this post

**user3** flagged your pos, user1

**user2** hello

Comment

[TestModerat:Talk:New](#) **Post flagged!**

11:33 am

**user1** [New.](#)  
new thread

[TestModerat:Talk:New](#) **You already flagged this post!**

11:33 am

**user1** [New.](#)  
new thread

# Moderation

- ❖ A group of moderators to review flagged posts
- ❖ Root user can add other moderators to the group
- ❖ Each flagged post appears as a separate thread
- ❖ Each thread allows moderators to:
  - Comment
  - Approve
  - Delete
  - Check original posts
- ❖ Benefits:
  - Ensures a safe atmosphere for users
  - Risk mitigation: Detect malicious content, phishing attempts, spams, harassment etc

# Moderation



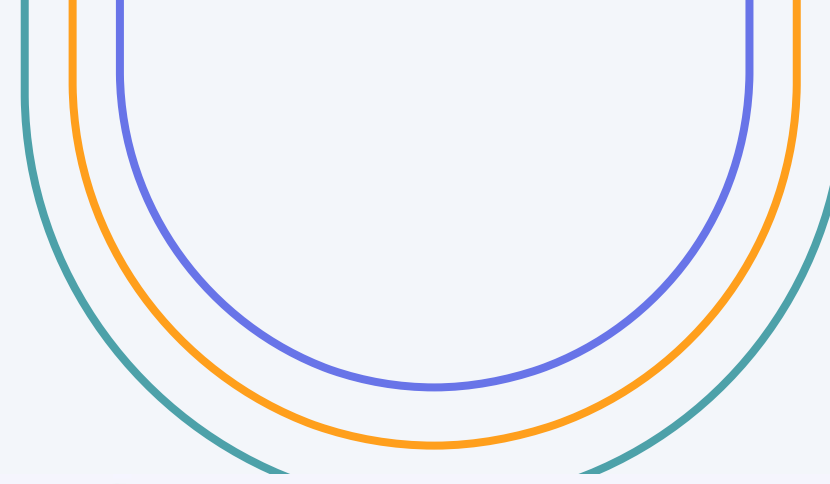
Start New Thread

1 m 51 s ago

[Bad Post group 10](#)

15 hours ago

[root joined Moderation!](#)



12:03 pm

[Bad Post group 10](#)  
user1 flag this post

Comment Approve Delete

Original Post

Flagged post approved



Filter Go

Public (500 posts, 500 threads)

Last Post: [%E9%9A%90%E7%A7%81Wiki网页创建了!](#)

11:33 am

[New.](#)  
user1 new thread

This post has been flagged  
user1

# Response Time Testing

Differential Privacy		
Test Type	Baseline	Post Implementation
System Load Time	0.091s	0.091s
Page load time - Edit Group Page	0.19s	0.19s
Page load time - View Group Page	0.15s	0.16s
Page load time - Manage Group Page	0.1s	0.1s
Flagging		
Test Type	Baseline	Post Implementation
System Load Time	0.091s	0.092s
Page load time - Group Thread page	0.29s	0.32s
Response Time - Flag post	1.6s	0.92s
Moderation		
Test Type	Baseline	Post Implementation
System Load Time	0.091s	0.095s
Page load time - Root login and load	0.513s	0.515s
Time taken to approve/delete	-	0.3s
Time taken to view original thread	0.14s	0.12s
Time taken to comment	0.13s	0.13s
Time to add new users	0.16s	0.18s
Secret Sharing		
Test Type	Baseline	Post Implementation
System Load Time	0.091s	0.096s
Page load time - Group Creation	0.24s	0.25s
Time taken to generate a key	-	0.72s





---

# Conclusion

- ❖ Implemented new mechanisms to enhance security
- ❖ Incorporated content moderation features to elevate user experience and increase security of the platform
- ❖ Enhanced security of encryption keys

# Future Work

- ❖ Content moderation can be extended to include features like banning users, locking threads etc
- ❖ New avenues to extend the implementation of differential privacy
- ❖ Encryption techniques like homomorphic encryption can be explored to protect the existing upvote/downvote feature



**Thank you!**



---



# References

[1] C. Lanius, R. Weber, and W. I. MacKenzie Jr., "Use of bot and content flags to limit the spread of misinformation among social networks: a behavior and attitude survey," *Social Network Analysis and Mining*, vol. 11, no. 32, Mar 2021.

[2] E. Chandrasekharan, U. Pavalanathan, A. Srinivasan, A. Glynn, J. Eisenstein, and E. Gilbert. "You can't stay here: The efficacy of reddit's 2015 ban examined through hate speech." 2017. [Online]. Available: <https://doi.org/10.1145/3134666>

