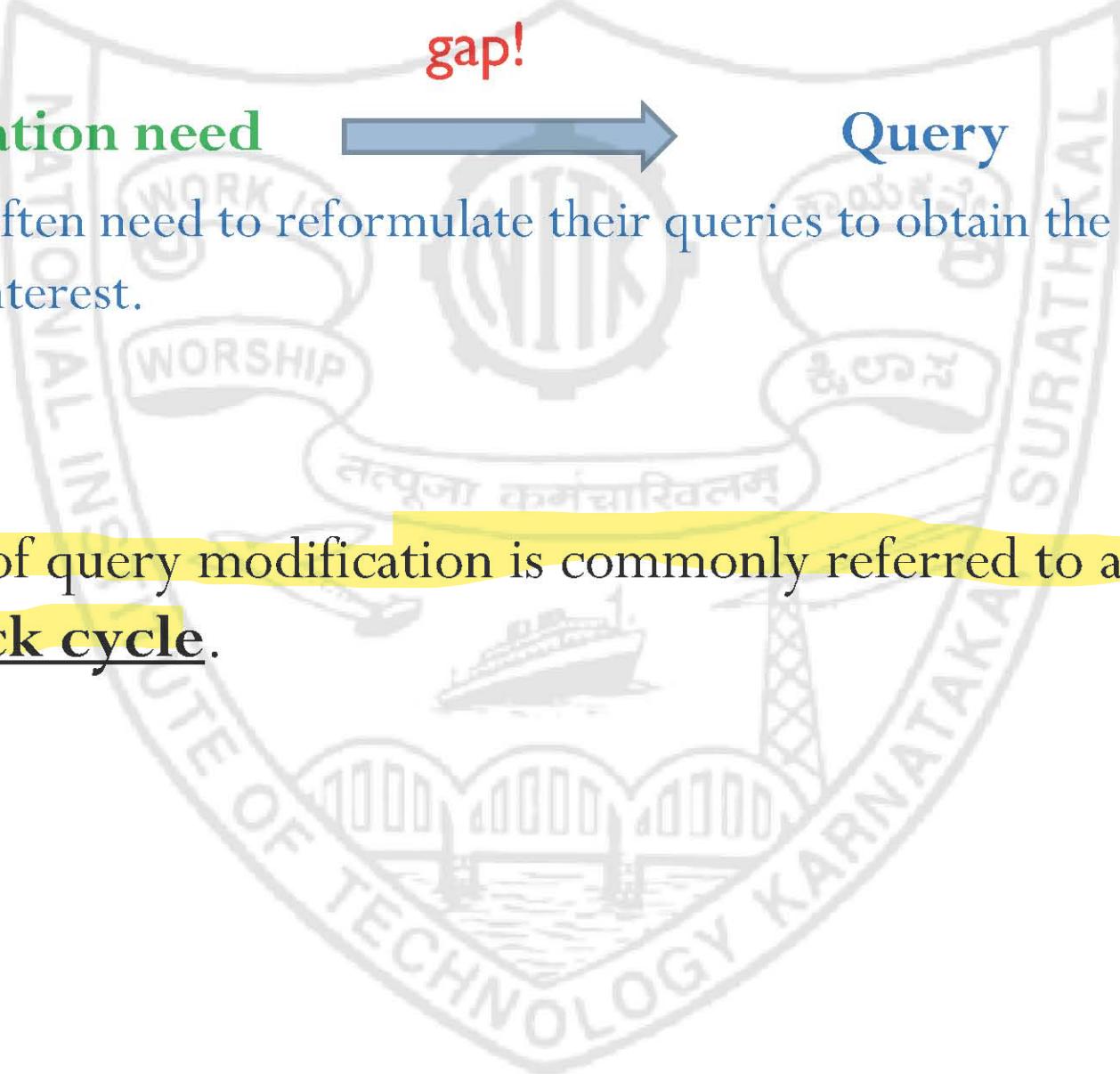




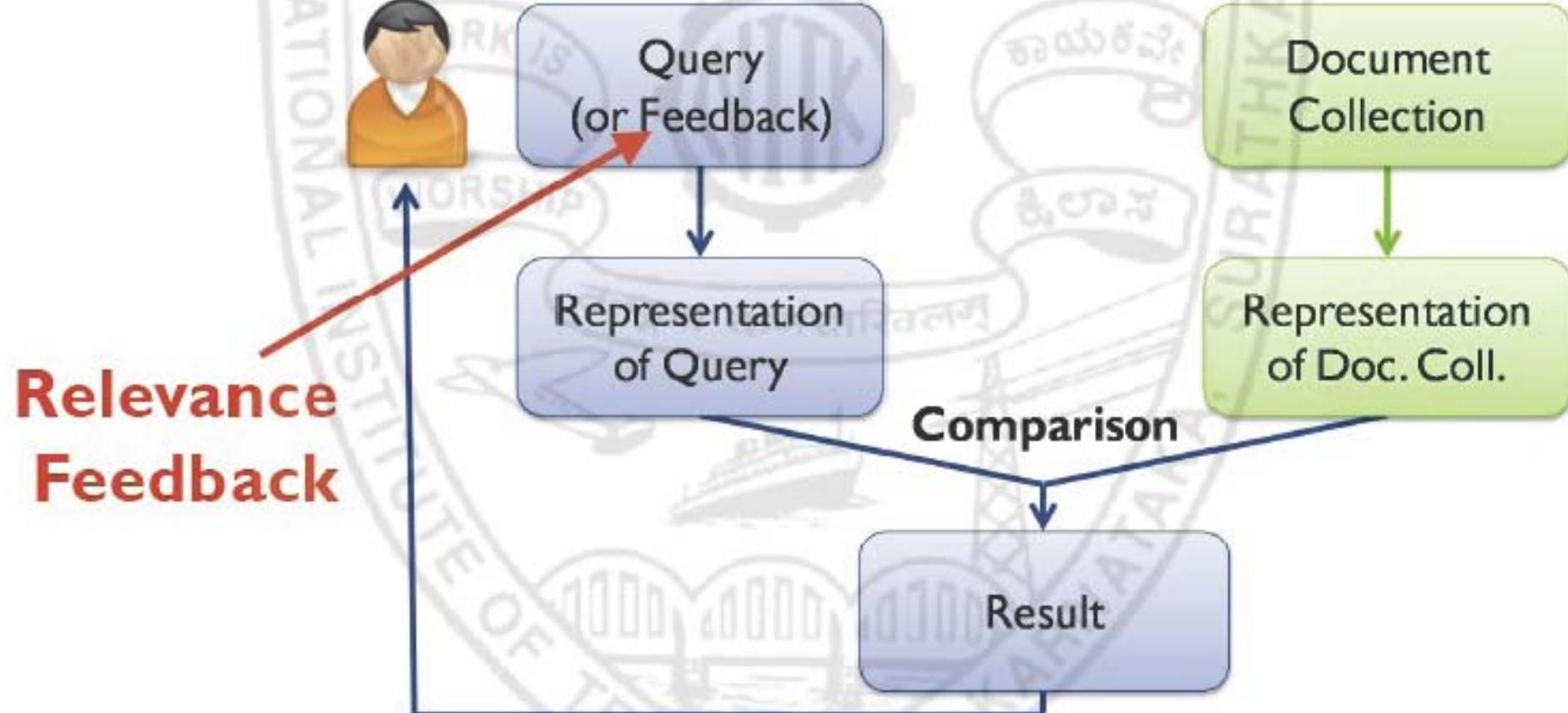
Capturing User Relevance

Relevance Feedback Approaches

Introduction

- 
- ▶ **Information need** → **Query**
- gap!
- ▶ users often need to reformulate their queries to obtain the results of their interest.
 - ▶ Process of query modification is commonly referred to as **feedback cycle**.

Introduction



Relevance Feedback

- ▶ User tells system whether returned documents are relevant to query (*information need*) or not.
- ▶ Helps augments keyword retrieval through *Query Reformulation*
 - ▶ give user opportunity to refine their query
 - ▶ tailored to individual
 - ▶ Iterative, subjective improvement

A Feedback Framework

- ▶ Feedback cycle –
 - ▶ composed of two basic steps:
 - ▶ Determine feedback information that is either *related* or *expected to be related* to the original query q (**Relevance Feedback**)
 - ▶ Determine how to transform query q to take this information effectively into account (**Query Reformulation / Expansion**)

A Feedback Framework

- ▶ Two basic approaches of feedback methods:
 - ▶ **Explicit feedback:**
 - ▶ the information for query reformulation/expansion is provided directly by the users.
 - ▶ **Implicit feedback:**
 - ▶ information for query reformulation/expansion is implicitly derived by the system.

Explicit Feedback

- ▶ feedback information is provided directly by the users
 - * collecting such feedback information is expensive and time consuming
- ▶ For e.g.
 - ▶ Direct questions asked to users by search engines or websites regarding retrieval performance.

Explicit Feedback

A Google search results page for the query "nitk". The top result is the official website of NITK Surathkal. Below the main search results, there is a "People also ask" section with several questions and their answers. A red oval highlights the "Feedback" link at the bottom right of this section.

Google nitk

All News Images Maps Videos More Tools

About 22,40,000 results (0.50 seconds)

<https://www.nitk.ac.in> :: **NITK Surathkal**

Academics · Research · Campus · Culture · Rankings · NIRF Engineering · NIRF Overall · QS India

Results from nitk

Fee Struct
FEE PAYMENT

Departmer
Water Resourc

Announcer
Announcement

Admission:
@nitk.ac.in; @i

People also

Is NITK a Tier 1

Which branch is best in NITK?

Is NIT Surathkal good for CSE?

Is NIT Surathkal top NIT?

Feedback

Explicit Feedback

9 Indian Institute of Technology Hyderabad >

10 National Institute of Technology Karnataka Surathkal. >

11 Jadavpur University >

 nirf.org

Source: NIRF via Careers360 · Learn more

MORE RANKINGS

People also search for


National Institute of Technol... Kozhikode


National Institute of Technol... Surathkal


National Institute of Technol... Warangal


National Institute of Technol... Rourkela

View 10+ more

Related searches

NIT college


National Institute of Technol... Warangal


National Institute of Technol... Calicut

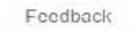

Motilal Nehru National In...

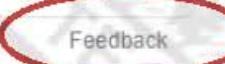

Maulana Azad National In...


Visvesvara... National Institute Of...


National Institute of Technol... Nagpur

See more →

 Feedback

 Feedback

Claim this knowledge panel

Explicit Feedback

Projectors (Showing 1 – 4 products of 4 products)

Sort By [Popularity](#) [Price – Low to High](#) [Price – High to Low](#) [Newest First](#)

ZEBRONICS Zeb-Pixa Play 12 with Dolby Audio Support & 7...

Black

4.1★ (1,281) Assured

₹8,999 ₹21,999 59% off

Free delivery

Saver Deal

ZEBRONICS ZEB-PIXAPLAY 10 LED with FHD 1080p support, 4...

White

4★ (15) Assured

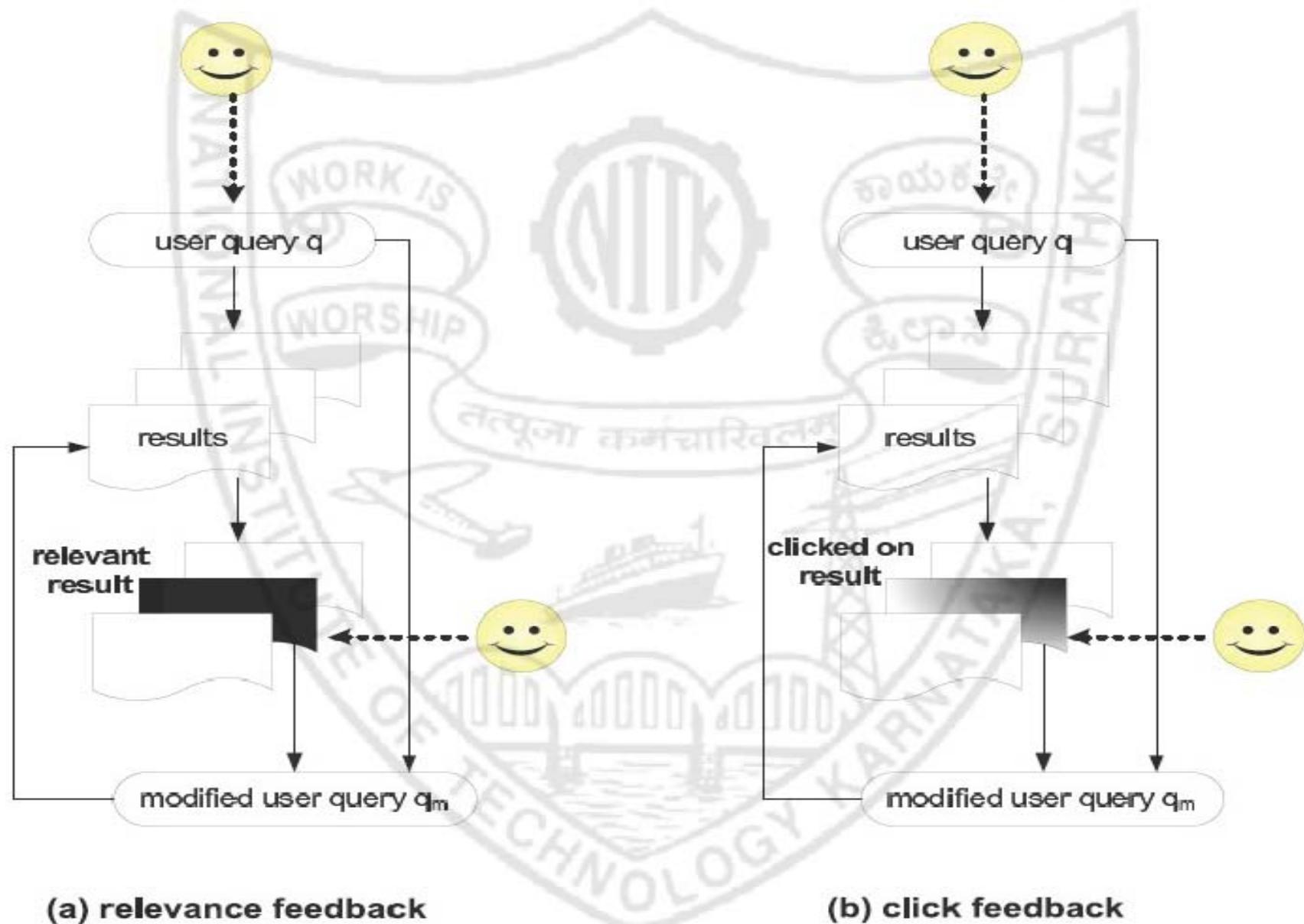
₹5,999 ₹10,999 45% off

Free delivery

Top Discount on Sale

Did you find what you were looking for?

Explicit Feedback -- Capture methods

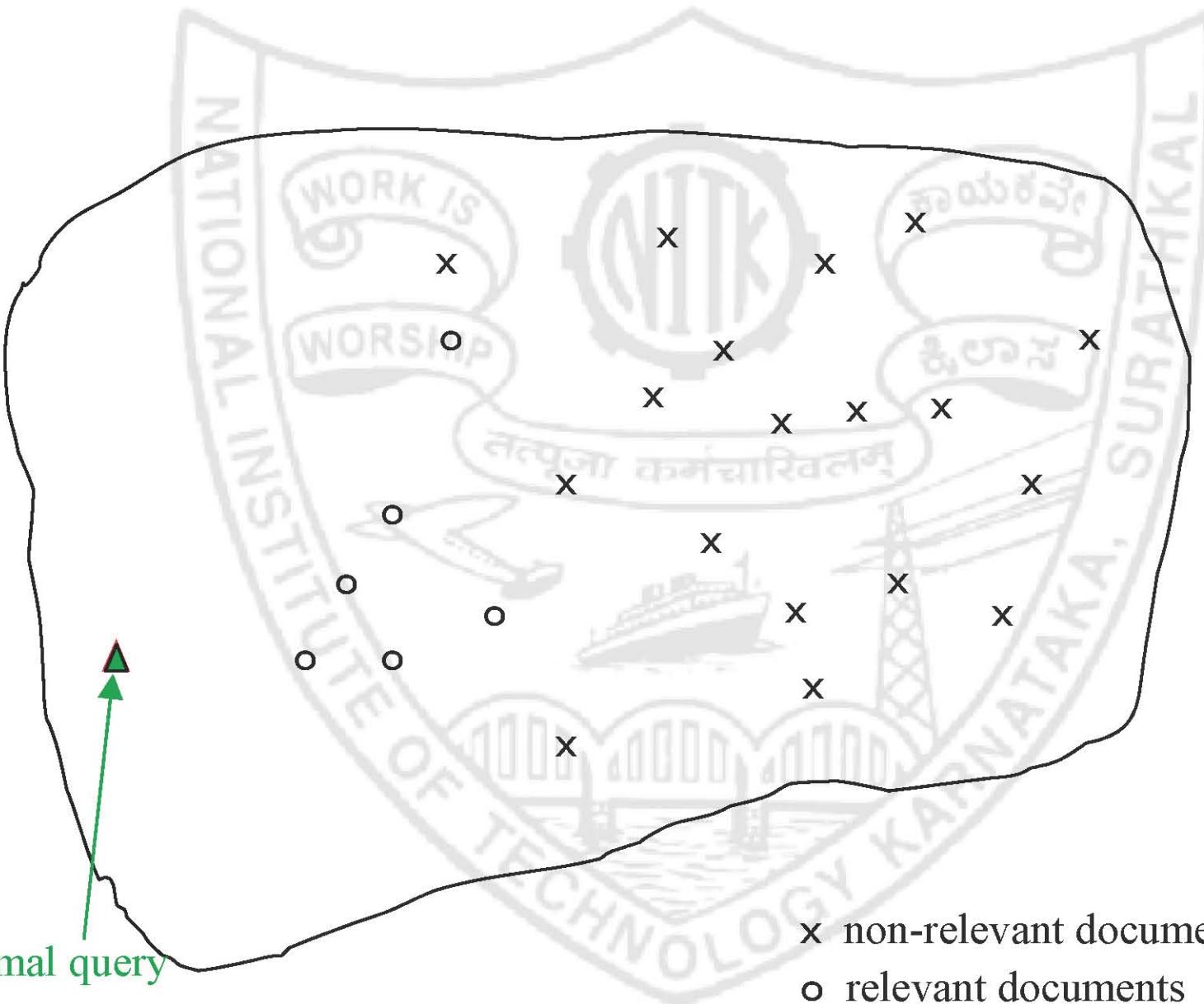


(a) relevance feedback

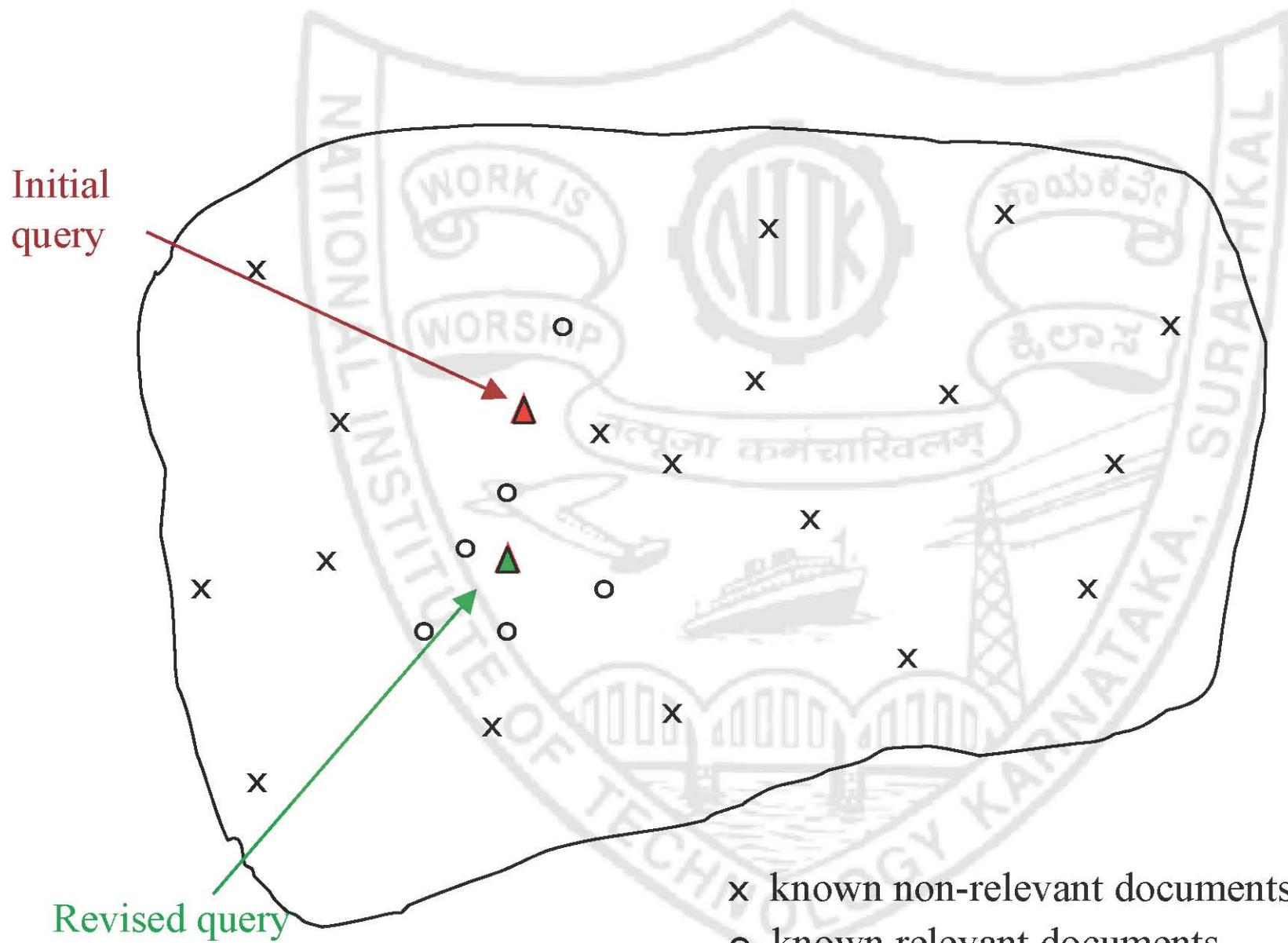
(b) click feedback

Relevance Feedback for Vector Space Model

The Theoretically Best Query



Relevance feedback on initial query



Rocchio's Method [1971]

- ▶ incorporates relevance feedback information into vector space model.
 - ▶ Introduced in the SMART system.
- ▶ Used with the Vector Space Model-
 - ▶ automatically re-weights terms
 - ▶ automatically adds new terms to modified query (found from relevant docs)
- ▶ Aim: to maximize $\{\text{sim}(\text{Q}, \text{Rel-docs}) - \text{sim}(\text{Q}, \text{Non-rel-docs})\}$

Rocchio's Method (1971)

$$Q_1 = Q_0 + \frac{\beta}{n_1} \sum_{i=1}^{n_1} R_i - \frac{\gamma}{n_2} \sum_{i=1}^{n_2} S_i$$

where

Q_0 = the vector for the initial query

R_i = the vector for the relevant document i

S_i = the vector for the non-relevant document i

n_1 = the number of relevant documents chosen

n_2 = the number of non-relevant documents chosen

β and γ tune the importance of relevant and nonrelevant terms
(in some studies best to set β to 0.75 and γ to 0.25)

Rocchio's Method (1971)

- ▶ *Positive feedback* - moves query closer to relevant documents

$$\frac{\beta}{n_1} \sum_{i=1}^{n_1} R_i$$

Positive Feedback

- ▶ *Negative feedback* - moves query away from non-relevant documents (but, not necessary closer to relevant ones)

$$\frac{\gamma}{n_2} \sum_{i=1}^{n_2} S_i$$

Negative Feedback

* *Most systems only use positive feedback*

Rocchio's Method: Example

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0
D3 (nr)	0	0	4	3	3

Term weights and relevance judgements for 3 documents returned after submitting the query Q_0

Assume $\beta = 0.5$ and $\gamma = 0.25$

$$Q_1 = Q_0 + \frac{\beta}{n_1} \sum_{i=1}^{n_1} R_i - \frac{\gamma}{n_2} \sum_{i=1}^{n_2} S_i$$

$$\begin{aligned}Q_1 &= (3, 0, 0, 2, 0) + 0.25 * (2+1, 4+3, 0, 0, 2) - 0.25 * (0, 0, 4, 3, 3) \\&= (3.75, 1.75, -1, 1.25, -0.25) \\&= (3.75, 1.75, 0, 1.25, 0)\end{aligned}$$

(Note: negative entries are changed to zero)

Rocchio's Method: Example

- After applying Rocchio's method, the new query -

$$Q_0 = (3, 0, 0, 2, 0) \rightarrow Q_1 = (3.75, 1.75, 0, 1.25, 0)$$

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
Q1	3.75	1.75	0.00	1.25	0.00
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0

Use the new query and compute similarities using matching functions
(e.g. Cosine Similarity)

Rocchio's Method: Example

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
Q1	3.75	1.75	0.00	1.25	0.00
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0

- After applying Rocchio's method, the new query -

$$Q_0 = (3, 0, 0, 2, 0) \rightarrow Q_1 = (3.75, 1.75, 0, 1.25, 0)$$

Use the new query and compute similarities using Cosine Similarity



	D1	D2	D3
Q0	?	?	?
Q1	?	?	?

Rocchio's Method: Some observations

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0
D3 (nr)	0	0	4	3	3

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
Q1	3.75	1.75	0.00	1.25	0.00
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0

- ▶ The new query Q1 added a weight for term 2.

Rocchio's Method: Some observations

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0
D3 (nr)	0	0	4	3	3

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
Q1	3.75	1.75	0.00	1.25	0.00
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0

- ▶ The new query Q1 added a weight for term 2.
- ▶ Terms initially not in user's vocabulary, may be added because it appeared as significant in enough relevant documents.

Rocchio's Method: Some observations

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0
D3 (nr)	0	0	4	3	3

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
Q1	3.75	1.75	0.00	1.25	0.00
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0

- Initial query resulted in high score for D3, though it was not relevant to the user (due to the weight of term 4). New query decreased score of D3 and increased those of D1 and D2.

Rocchio's Method: Some observations

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0
D3 (nr)	0	0	4	3	3

	T1	T2	T3	T4	T5
Q0	3	0	0	2	0
Q1	3.75	1.75	0.00	1.25	0.00
D1 (re)	2	4	0	0	2
D2 (re)	1	3	0	0	0

- Initial query resulted in high score for D3, though it was not relevant to the user (due to the weight of term 4). New query decreased score of D3 and increased those of D1 and D2.
- In general, fewer terms in the query, the more likely a particular term could result in non-relevant results**

Rocchio's Method – Some observations

- ▶ New query moves toward relevant documents and away from irrelevant documents
- ▶ *Tradeoff α vs. β/γ* : Positive feedback is more valuable than negative feedback
 - ▶ so, set $\gamma < \beta$; e.g. $\gamma = 0.25, \beta = 0.75$).
- ▶ Term weight can go negative.
 - ▶ Negative term weights are ignored (set to 0)

Relevance Feedback for Probabilistic Models

Relevance Feedback for Probabilistic Models

- ▶ Instead of reweighting in a vector space, if user provides some info on relevant and irrelevant documents, Probabilistic Relevance Feedback can be achieved.
- ▶ Incorporated in Binary Independence Model (BIM)
 - ▶ Relevance Feedback
 - ▶ Pseudo-relevance Feedback

More reading

- ▶ Salton, Gerard, and Chris Buckley. "Improving retrieval performance by relevance feedback." *Journal of the American society for information science* 41.4: 288-297.
- ▶ Rocchio, J. Relevance feedback in information retrieval. *The Smart retrieval system-experiments in automatic document processing*, 313-323.