Effective Recommendation in Knowledge Portals

The SKYbrary case study

Master thesis presentation
Roy Hoeymans

Content

- Introduction
- Problem description
- Research questions
- Experiment setup
- Case study: SKYbrary
- User navigation and content based algorithms
- Experiment 1 – Offline evaluation
- Hybrid algorithm
- Experiment 2 – Survey
- Recommendation engine demo application
- Conclusion
- Discussion

Introduction

- Focus of this research:
  building a recommender system for knowledge portals

Recommender systems

- Provide suggestions to the user
- Several types of data and techniques

Knowledge portals

- Web-based single point of access to information on a specific subject
- Can contain webpages, file systems, or applications
- Maintained by domain experts

Problem description

- Recommendations need to be relevant to current article
- Experts have to select recommendations manually
- Knowlegde portals: not a lot of data available
- Anonymous users
- Useful data sources:
  - Implicit user navigation data → user-navigation algorithm
  - Textual content → content based algorithm
- Combine algorithms into a hybrid algorithm
  - Compensate for separate disadvantages

Research questions

Main question

- Which type of recommendation algorithm is the most effective for a knowledge portal: a user-navigation based, a content based, or a hybrid algorithm that combines these two?

Subquestions

- Which type of user-navigation based algorithm works best for a recommender system in a knowlegde portal?
- Which type of content based algorithm works best for a recommender system in a knowlegde portal?
- Is a recommender system based on the combination of a content based and a user-navigation based algorithm more or less effective than a recommender system based on those separate algorithms?

Experiment setup

Case study

- SKYbrary: knowledge portal on aviation safety
- Implementation for knowledge portals using MediaWiki and Google Analytics

Experiment setup

- Develop several versions of user-navigation and content based algorithms
- Experiment 1 – offline evaluation
- Develop hybrid algorithm
- Experiment 2 – survey

Case study: SKYbrary

- Knowledge portal on aviation and air traffic management safety
- Developed by DNV-GL for EUROCONTROL
- Target audience: air traffic managers, pilots
- Set up as a Wiki site
- Content managed by team of experts to ensure quality
- ‘Related Articles’-section

Case study: SKYbrary

Hypoxia

Categories: Aeromedical | Loss of Control | *Operational Issues*

Contents [hide]

Article Information	
Category:	Aeromedical
Content source:	SKYbrary
Content control:	Air Pilots

Description

Hypoxia is defined as a lack of oxygen in the body tissues. This can be caused either by a shortage of oxygen in the air being breathed or by a number of physiological/pathological issues affecting blood circulation or the quantity of oxygen carried by haemoglobin in the blood.

The effects of hypoxia include fatigue, confusion, euphoria, inability to concentrate, impaired decision-making, impaired psychomotor performance, loss of consciousness and, eventually, death. Hypoxia does not cause discomfort or pain so its onset can be insidious and pass un-noticed by crews who are not fully aware of its dangers.

User-navigation based algorithm

Version 1

- Sessions that contain X and then Y
- Retrieved with Google Analytics API
- Accurate but obvious recommendations

Article Y	X → Y
Emergency or Abnormal Situation	175
Emergency Transponder Codes	155
Distress/Emergency Frequencies	154
Bird Strike	119
In-Flight Fire: Guidance for Flight Crews	115

Article X = ‘Emergency Communications’

User-navigation based algorithm

Version 2

- Number of sessions in which article X is not viewed but article Y is
- Corrects for the popularity of Y

$$\frac{\frac{X \rightarrow Y}{X}}{\frac{!X \rightarrow Y}{!X}}$$

Version 3

- Same as version 2
- Different formula

$$\frac{P(X \rightarrow Y)}{P(X) * P(Y)}$$

Content based algorithm

- Text retrieved from the MediaWiki database
- Similarity score for each pair of articles

Version 1

- Term frequency

Version 2

- Term frequency – inverse document frequency (tf-idf)
- Gives more importance to words that are rare

Experiment 1 – offline evaluation

Goal

- Select best version of each algorithm type
  - Hybrid algorithm
  - Experiment 2 – survey

Method

- Compare recommendations generated by algorithms to recommendations that are manually selected by the SKYbrary experts

Experiment 1 – offline evaluation

Recall for top S recommendations

User-navigation based

	S = 2	S = 5	S = 10	S = 30
Version 1	0.712	0.758	0.861	0.940
Version 2	0.398	0.500	0.677	0.917
Version 3	0.397	0.503	0.669	0.917

Content based

	S = 2	S = 5	S = 10	S = 30
Version 1	0.545	0.590	0.720	0.838
Version 2	0.605	0.721	0.880	1

Experiment 1 – offline evaluation

Recall for top S recommendations

User-navigation based

	S = 2	S = 5	S = 10	S = 30
~~Version 1~~	~~0.712~~	~~0.758~~	~~0.861~~	~~0.940~~
Version 2	0.398	0.500	0.677	0.917
Version 3	0.397	0.503	0.669	0.917

Content based

	S = 2	S = 5	S = 10	S = 30
Version 1	0.545	0.590	0.720	0.838
Version 2	0.605	0.721	0.880	1

Hybrid algorithm

Compensate for separate disadvantages

- User-navigation based
  - Cold start problem
- Content based
  - Does nothing with user preferences
  - Over-specialization

Intuition

- Adjust articles recommended based on content similarity with its relative popularity
- High similarity but not popular -> decrease score
- Low similarity but very popular -> increase score

Hybrid algorithm

- Article X = ‘Hypoxia’

Content based

Article Y	Score
Emergency Depressurisation	0.48
Pressurisation Problems	0.38
Aircraft Pressurisation Systems	0.34
Cabin Altitude	0.34
Explosive Depressurisation	0.26
Rapid Depressurisation	0.24
Time of Useful Consciousness	0.19

User-navigation based

Article Y	Score
Time of Useful Consciousness	1
Explosive Depressurisation	0.68
Emergency Depressurisation	0.46
Rapid Depressurisation	0.44
Aircraft Pressurisation Systems	0.38
Cabin Altitude	0.23
Pressurisation Problems	0.13

Hybrid algorithm

- Article X = ‘Hypoxia’

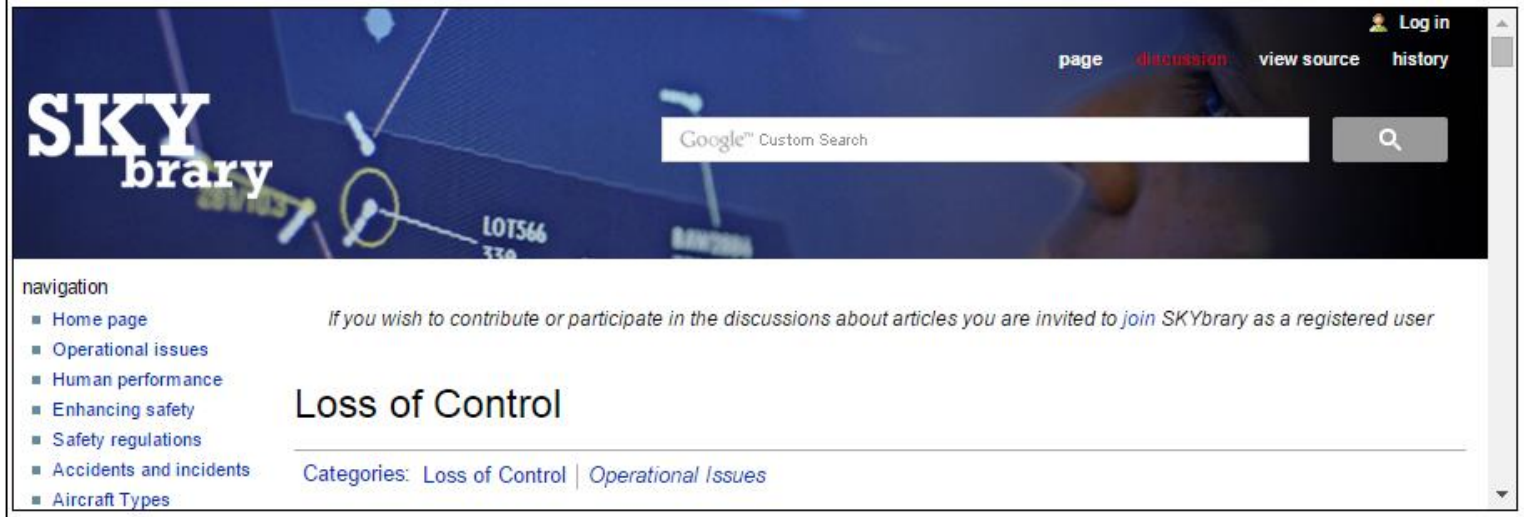
Hybrid

Article Y	Score
Time of Useful Consciousness	0.6
Emergency Depressurisation	0.47
Explosive Depressurisation	0.47
Aircraft Pressurisation Systems	0.36
Rapid Depressurisation	0.34
Cabin Altitude	0.27
Pressurisation Problems	0.26

Experiment 2 – survey

- Goal: determine which algorithm performs best
- Respondents: SKYbrary experts
- 15 questions
- Each question shows recommendations:
  - 2 articles recommended by user-navigation based algorithm
  - 2 articles recommended by content based algorithm
  - 2 articles recommended by hybrid algorithm
  - 1 randomly selected article (baseline)
- Respondents rate recommendations on a scale of 1 to 5

Example question - Loss of Control (link)



Question

Imagine that a visitor of SKYbrary is interested in the article displayed above, and the recommender system produces the recommendations for the articles below. Please indicate for each recommended article how relevant you think it is for the visitor.

Score meanings: 1 = completely irrelevant | 2 = irrelevant | 3 = neutral | 4 = relevant | 5 = very relevant

Recommended article	Link	Score	Remark (optional)
Stall	link	1 ○ 2 ○ 3 ○ 4 ○ 5 ○	
Recovery from Unusual Aircraft Attitudes	link	1 ○ 2 ○ 3 ○ 4 ○ 5 ○	
Aerodynamic Stall Awareness and Avoidance	link	1 ○ 2 ○ 3 ○ 4 ○ 5 ○	
In-Flight Icing	link	1 ○ 2 ○ 3 ○ 4 ○ 5 ○	
C404, Kulusuk Greenland, 2002 (WX LOC HF)	link	1 ○ 2 ○ 3 ○ 4 ○ 5 ○	
Emergency Evacuation on Land	link	1 ○ 2 ○ 3 ○ 4 ○ 5 ○	

Experiment 2 – survey results

- 6 respondents
- Algorithms perform much better than random baseline
- Hybrid performs slightly better than content based

Descriptive statistics

Algorithm	Mode	Mean
Baseline	1	1.27
User-navigation based	4	3.42
Content based	4	3.65
Hybrid	5	3.69

Experiment 2 – survey results

- Test for statistical significance
  - Mann Whitney-U test
  - 95% confidence interval ($P < 0.05$)
  - Null hypothesis: performance of algorithms is not different

Algorithm pair	P-value	Significant
HY – CB	0.52	No
HY – UN	**0.02**	**Yes**
CB – UN	0.09	No
HY – BL	**0.00**	**Yes**
CB – BL	**0.00**	**Yes**
UN – BL	**0.00**	**Yes**

- Hybrid algorithm performs significantly better than a user-navigation based algorithm but not significantly better than a content based algorithm

Recommendation engine demo application

- Developed during research
- Tool that shows recommendations for a selected article and algorithm
- Allows content managers to decide on recommendation algorithm
- Supports process of manually selecting recommendations

SKYbrary The single point reference for aviation safety knowledge

Welcome, editor | Log out

Recommendation engine demo

Welcome to the recommendation engine demo.
Choose a SKYbrary article that you want to show recommendations for.

Article: Hypoxia
Algorithm: Content based

Apply filters:

- [x] Hide Accidents & Incidents articles
- [ ] Hide recommended articles that are linked in the text
- [ ] Hide recommended articles that are in the Related Links section

Recommended article	URL	Place in article	Score
Emergency Depressurisation	link	Related Links	0.48
Pressurisation Problems: Guidance for Controllers	link	None	0.38
Aircraft Pressurisation Systems	link	Related Links	0.34
Cabin Altitude	link	Link in article	0.32
Explosive Depressurisation	link	Related Links	0.26
Differential Pressure	link	None	0.26
Rapid Depressurisation	link	None	0.24
Pilot Incapacitation	link	None	0.23
High Altitude Flight Operations	link	None	0.19
Time of Useful Consciousness	link	Related Links	0.19

Showing 1 to 10 of 25 entries

Previous 1 2 3 Next

Conclusion

Research question:

- Which type of recommendation algorithm is the most effective for a knowledge portal: a user-navigation based, a content based, or a hybrid algorithm that combines these two?

Answer:

- Hybrid is significantly more effective than a user-navigation based algorithm
- Not enough evidence for claims:
  - Hybrid is more effective than a content based algorithm
  - Content based algorithm is more effective than a user-navigation based algorithm

Discussion

- Subtle difference between content based and hybrid algorithm
- Difficult to find significant difference

Limitations

- Amount of respondents of survey

Future work

- Different hybrid methods
- More extensive or different type of evaluation

Summary

- Research on recommender system in knowledge portals
- Several versions of user-navigation and content based algorithms
- Experiment 1 – offline evaluation
  - Selected best algorithm of each version for hybrid algorithm and survey
- Experiment 2 – survey
  - Hybrid works better than user-navigation based
  - Not enough evidence that hybrid works better than content based
- Recommendation engine demo application: support manual process
- Future work: more extensive evaluation

The end

Questions? Comments?
Suggestions?