Forecasting Box Office Success of Movies: An Update and a DSS Perspective

Ramesh Sharda and Dursun Delen
Institute for Research in Information Systems
Department of Management Science and Information Systems
William S. Spears School of Business
Oklahoma State University

(Assistance from Michael Henry on recent data collection and trials; Ben Johnson and Xin Cao on MFG implementation)

Forecasting Box-Office Receipts: A Tough Problem!

"… No one can tell you how a movie is going to do in the marketplace… not until the film opens in darkened theatre and sparks fly up between the screen and the audience"

Mr. Jack Valenti
President and CEO
of the Motion Picture Association of America

Introduction

Pirates of the Caribbean

"When production for the film was first announced, movie fans and critics were skeptical of its chances of success"

– www.wikipedia.com

- 3rd highest grossing movie in 2003
- 22nd highest grossing movie of all time
- Sequel was the 6th highest grossing movie of all time

-www.the-movie-times.com

Our Approach – Movie Forecast Guru

- **DATA** –Movies released between 1998-2005
- **Movie Decision Parameters:**
  - Intensity of competition rating
  - MPAA Rating
  - Star power
  - Genre
  - Technical Effects
  - Sequel ?
  - Estimated screens at opening
  - …
- **Output:** Box office gross receipts (flop → blockbuster)

Class No.	1	2	3	4	5	6	7	8	9
Range (in Millions)	< 1 (Flop)	> 1 < 10	> 10 < 20	> 20 < 40	> 40 < 65	> 65 < 100	> 100 < 150	> 150 < 200	> 200 (Blockbuster)

Method: Neural Networks and others

- Output
  - Box office receipts: 9 categories
  - Flop (category 1)
  - Blockbuster (category 9)
- Prediction Results
  - Bingo
  - 1-Away

Updates of Previous Results

Original data from 1998 to 2002

834 Movies Tested

		Actual Categories									Avg.
		1	2	3	4	5	6	7	8	9	
Predicted Categories	1	37	35	5	4	0	0	0	1	2	
	2	33	37	13	14	0	1	0	1	1	
	3	5	13	28	21	1	4	8	7	4	
	4	15	3	16	38	0	2	3	4	9	
	5	0	0	6	13	55	30	7	3	2	
	6	0	1	2	3	31	26	19	13	4	
	7	0	0	8	5	5	12	24	21	10	
	8	0	0	5	2	3	7	24	20	16	
	9	0	0	9	1	2	7	8	22	43	
	BINGO	0.411	0.416	0.304	0.376	0.567	0.292	0.258	0.217	0.473	**0.369**
	1-Away	0.778	0.955	0.620	0.713	0.887	0.764	0.720	0.685	0.648	**0.752**

New Experiments

Method

- Collect Data from 2003 to 2005
- Run test on data from 2003 to 2005
- Compare with previous results from 1998 to 2002

Experiment One

Data

- Collect and test 475 movies
  - Independent variables: www.imdb.com
  - Dependent variables: www.the-movie-times.com

Experiment One

1998 to 2002	2003 to 2005
Bingo: 39.6%	Bingo: 54.1%
1-Away: 75.2%	1-Away: 74.6%

*1998 to 2002 results from Sharda and Delen

Experiment Two

Method

- Combine data from 1998 to 2005
- Test data from 1998 to 2005
- Compare with previous tests results

Experiment Two

Data

- Test included 1,323 movies
- 1998 to 2002 included 848 movies
- 2003 to 2005 included 475 movies

Experiment Two

1998 to 2002	**2003 to 2005**	**1998 to 2005**
Bingo: 54.5% 1-Away: 80.7%	Bingo: 54.1% 1-Away: 74.6%	Bingo: 49.12% 1-Away: 81.60%

*current 1998 to 2002 results

What about predictions in 2006?

Movie	Actual	Prediction
1. Pirates of Caribbean 2	1. Class 9	1. Class 9
2. The Break Up	2. Class 7	2. Class 7
3. Inside Man	3. Class 6	3. Class 6
4. V for Vendetta	4. Class 6	4. Class 6
5. Underworld 2	5. Class 5	5. Class 4
6. BloodRayne	6. Class 2	6. Class 4
	7. Class 1	7. Class 2

Results So far…

- The more data available to train and test model, the higher the prediction rate.
- Re-evaluating the data to ensure consistency and accuracy improved the prediction rate.
- Neural networks can handle complex problems in forecasting in difficult business situations.

Web-Based DSS

- Information fusion (multiple method forecasting)
- Use of models not owned by the developer
- Sensitivity Analysis

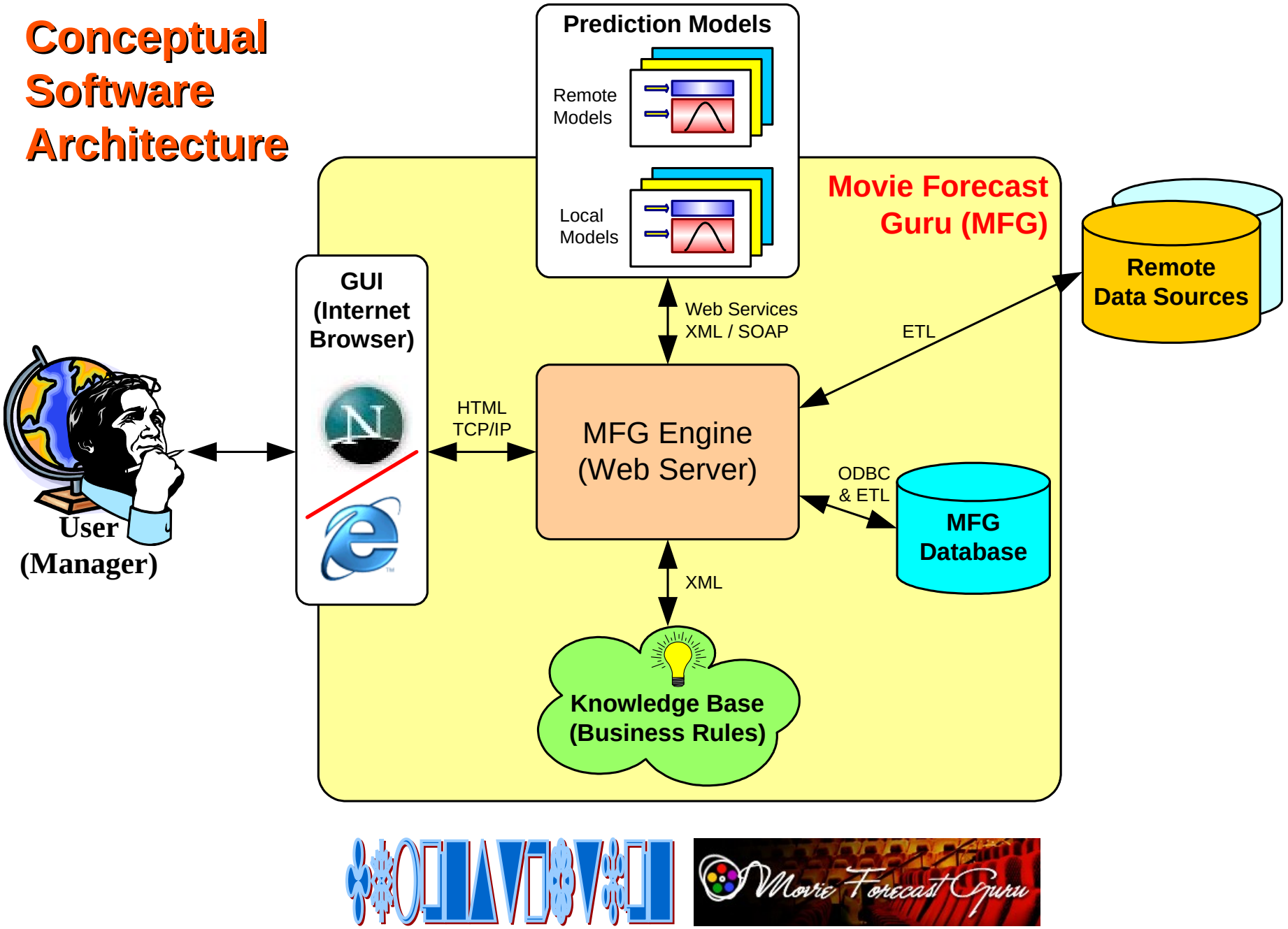
Web-Based DSS

- Collaboration among stakeholders
- Platform independence
- Forecasting models change frequently
  - Versioning
- Web services a good method for updates
- Web-based DSS!

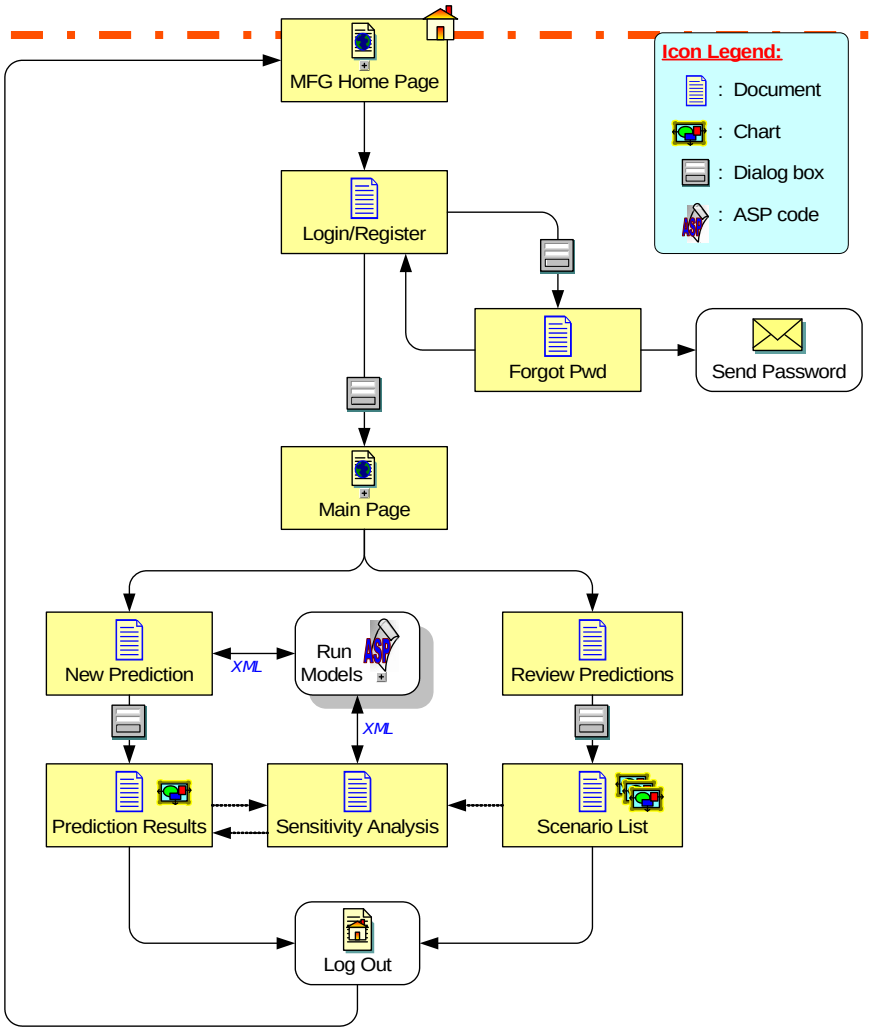
DSS: Movie Forecast Guru

- Forecast Methods:
  - Neural Networks
  - Decision Tree (CART & C5)
  - Logistic Regression
  - Discriminant Analysis
  - Information Fusion
- .Net server

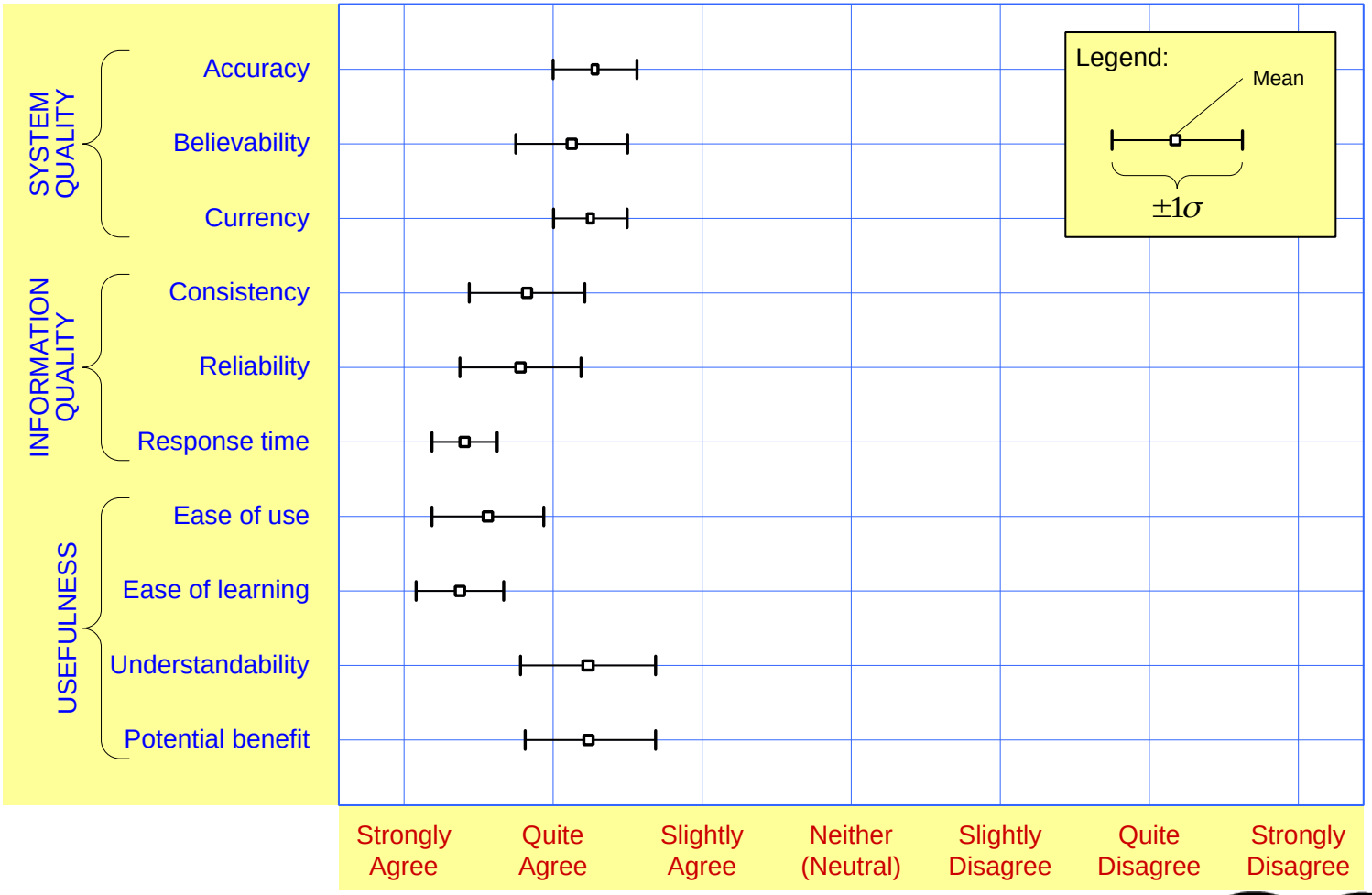
Conceptual Software Architecture



User Interaction with MFG



Preliminary Assessment



Conclusions

- Interesting problem for DSS Implementation
- Marketing challenge remains!
- Many other similar problems in forecasting
- Web-DSS framework