Predicting Subscriber Dissatisfaction and Improving Retention in the Wireless Telecommunications Industry

Michael C. Mozer\textsuperscript{+\*}
Richard Wolniewicz\textsuperscript{*}
Robert Dodier\textsuperscript{*}
Lian Yan\textsuperscript{*}
David B. Grimes\textsuperscript{+\*}
Eric Johnson\textsuperscript{*}
Howard Kaushansky\textsuperscript{*}

\textsuperscript{+}Department of Computer Science
University of Colorado, Boulder

\textsuperscript{*}Athene Software
Boulder, Colorado
The Wireless Industry

Extremely dynamic and competitive market

• Penetration rate 25% in 1998, 50% in 2004 vs. 71% in France (3Q 2003)
• Some local markets have as many as five carriers.
• Carriers announce new rates and promotions almost every month.
• New services and technologies are constantly introduced.

Competition has resulted in high rate of churn—customers switching from one carrier to another.

Monthly churn rates in US are ~2% of customer base.

• Feb 2004: 14% of AT&T customers thinking of churning in next 3 months

Churn cost industry nearly $10 billion in 2001.

• Signing new subscriber costs 5 times as much as retaining existing one.
• For carrier with 5M subscribers with an annual churn rate of 30%, that is a lost revenue of $870M. Cutting churn in half will save $435M.
Decision-Making Framework

- Subscriber data
- Reasons-for-dissatisfaction prediction
- Responsivity-to-intervention prediction
- Churn prediction
- Profitability estimation
- Credit-risk assessment
- Decision making
- Intervention strategy
## Factors Influencing Subscriber Satisfaction

<table>
<thead>
<tr>
<th>Factor</th>
<th>Importance</th>
<th>Nature of data required for prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>call quality</td>
<td>21%</td>
<td>network</td>
</tr>
<tr>
<td>pricing options</td>
<td>18%</td>
<td>market, billing</td>
</tr>
<tr>
<td>corporate capability</td>
<td>17%</td>
<td>market, customer service</td>
</tr>
<tr>
<td>customer service</td>
<td>17%</td>
<td>customer service</td>
</tr>
<tr>
<td>credibility / customer</td>
<td>10%</td>
<td>market, customer service</td>
</tr>
<tr>
<td>communications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>roaming / coverage</td>
<td>7%</td>
<td>network</td>
</tr>
<tr>
<td>handset</td>
<td>4%</td>
<td>application</td>
</tr>
<tr>
<td>billing</td>
<td>3%</td>
<td>billing</td>
</tr>
<tr>
<td>cost of roaming</td>
<td>3%</td>
<td>market, billing</td>
</tr>
</tbody>
</table>
Information Sources Related to Churn

Network
- usage patterns (peak/off peak, number and duration of calls, location of calls)
- dropped calls
- quality of service

Billing
- base fee
- charge for minutes beyond prepaid limit
- roaming charges

Customer Service
- nature of complaints and resolution

Application for Service
- rate plan
- handset type
- credit history
- active services (number, type, avenue of activation, cancellation dates)
- customer classification (corporate vs. retail)

Market
- competitor rate plans

Demographics
- population density
- average income
Data Set

Provided by national wireless carrier

Account profile

- 46,744 subscribers, primarily small businesses
- Average revenue per subscriber = $234
- No long-term contracts
- Four state region
- 20% in major metropolitan areas

Time period

- Training data from October to December, 1998
- Test involved predicting churn in January or February 1999
- 6.2% churn rate
Data Representation

Naive

134 variables → 148 element vector

discrete one-of-$n$ variables translated to an $n$-dimensional subvector
e.g., credit classification

Sophisticated

134 variables → 73 element vector

collapsed across some variables
e.g., different types of calls to customer service

expanded some variables
e.g., length of time with carrier

transformations
e.g., ratios, regression coefficients
Methodology

Ten-fold cross validation

Model classes
- logit regression
- neural network
- decision tree

Representations
- naive
- sophisticated

Model combination techniques
- single model
- majority vote
- Adaboost
Lift Curve

Interpret predictor output as a probability of churn.

For a given probability threshold, determine
• fraction of all subscribers above threshold, and
• fraction of all churners above threshold.

Plot one quantity against the other for various thresholds.
Results

- naive, logit regression
- naive, NN
- sophisticated, logit regression
- sophisticated, NN

The graph shows the comparison of different models in identifying subscribers and churners. The X-axis represents the percentage of subscribers identified, while the Y-axis represents the percentage of churners identified.
Results

% subscribers identified vs % churners identified for different models:
- NN, boosting
- NN
- decision tree, boosting
- decision tree

Graph shows the model performance across different thresholds.
Should a given subscriber be contacted and offered some incentive to remain with the carrier?

Offer incentive to all subscribers with churn probability $> \theta$

Select $\theta$ to maximize expected cost savings to carrier

Expected savings depends on

$C_I$ cost to carrier of providing incentive

$H$ time horizon over which incentive affects subscriber’s behavior (assume 6 months)

$P_I$ reduction in probability that subscriber will leave within time horizon as a result of incentive

$C_L$ lost revenue that results from churn (assume $500 acquisition cost + income over H months)$
Expected savings also depends on statistics from predictor

\[ N_{pc/ac} \]  \# subscribers predicted to churn who actually churn barring intervention

\[ N_{ps/ac} \]  \# subscribers predicted to stay who actually churn barring intervention

\[ N_{pc/as} \]  \# subscribers predicted to churn who actually stay

\[ N_{ps/as} \]  \# subscribers predicted to stay who actually stay

Net cost to carrier of performing no intervention

\[ \text{net}_{NI} = ( N_{pc/ac} + N_{ps/ac} ) C_L \]

Net cost to carrier of performing intervention

\[ \text{net}_I = ( N_{pc/ac} + N_{pc/as} ) C_I + ( P_I N_{pc/ac} + N_{ps/ac} ) C_L \]

Savings per churnable subscriber

\[ \text{savings} = ( \text{net}_{NI} + \text{net}_I ) / ( N_{pc/ac} + N_{ps/ac} ) \]
Expected Savings Per Churnable Subscriber Under Various Assumptions Concerning Intervention Cost and Resulting Retention Rate

- **25% retention rate**
- **35% retention rate**
- **50% retention rate**
- **75% retention rate**
Real World Testing

Six week experiment

Control and treatment groups

Treatment group contacted based on our recommendation

Churn rate 3.7% in control group, 2.2% treatment group → 40% retention

Intervention cost = $92 ($17 for incentive, $75 for call center)

By decision-theoretic framework, the savings per churnable customer is $417.

From subscriber in treatment condition with .81 churn score:

...I am writing this letter in regards to an employee there that I feel deserves special recognition. Your representative, Alicia Holmes, has single handedly encouraged me to stay on with X as my cellular service provider. She is professional, competent, polite, an expert with her skills and knowledge of your services... She turned a bad experience with X into a good one. If not for her I would have left X as soon as possible...
Observations

Positive correlation with churn
- average monthly bill
- total number of calls
- credit class of customer
- ratio of recent to earlier monthly bills

Negative correlation with churn
- number of active dispatch and messaging services
- time with carrier
- average number of billed minutes per month

Mutual information between each variable and churn is extremely low (<< .01 bit).

Domain intuitions almost always supported by mutual information scores and prediction accuracy.

In many data sets, few higher-order regularities (given appropriate representation).

Because of redundancy and low information content, input pruning is feasible with no loss in accuracy.
Further Experiments

Test window shifted in time
Replicability of results over time
Comparison of various ensemble techniques
Comparison of Wireless and ISP Data Sets
Test Window Experiment

1 2 3 4 5 6 7 8

- Two month test window
- One month test window
- Overlapping training windows
- One month test window
- Nonoverlapping training windows

Train input
Train prediction
Test input
Test prediction
Optimizing Performance on Lift Curve

% subscribers identified

% churners identified

churn score

0 1
Decision Network for Subscriber Profitability

One-time costs

- Fashion consciousness of user
  - Frequency of upgrade
- Product interoperability
- Product sophistication
- Fashionability of product
- User sophistication
- Channel cost
- Raw cost if offer accepted
- Fixed service cost

Acquisition cost

Recurring costs and revenues

- Cost of support by phone
  - Support contact cost
- Cost of support by email
  - Bad credit & collection cost
- Credit rating
- Seasonal factors
- Accessibility & availability of coverage
- Churn?
- No. services subscribed to
- Expected use of product
- Still using product?
- Long term fixed charge
- Known expiration

Use of service over month

Monthly rate plan

Future monthly service revenue

Long term service costs

Nominal net revenue