On-line Sales and data analytics

STUDY GROUP MEMBERS:

R BEKKER, K BORN, I DENHAM-DYSON, P DUMANI, J JULYAN, T LEBESE, E NETSHIUNDA

SUPERVISED BY J SANDERS

January 13, 2018

Problem statement

Design the structure of an on-line store using Z notation, incorporating market analytics.

Overview

Client	Order	Item
Personal information	Status	Type of items
Log in information	Stock order history	Quantity
Payment information		Cost price
Wishlist		Selling price
Contact information		
Client order history		

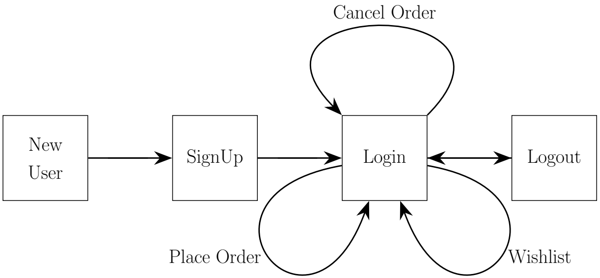

Figure: User interaction with the store

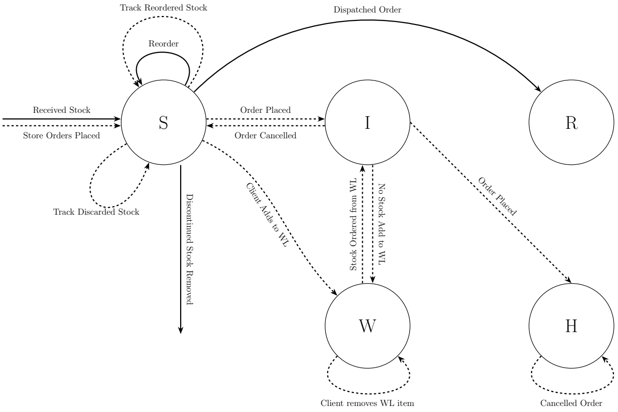

Figure: Movement of information and items around the store

Assumptions

- Financial security
- All orders are dispatched from a single location/warehouse
- Delivery to registered client's address
- After an item is dispatched, the delivery is handled by a courier

The Store

Store

$knownClient : \mathbb{P}\, Client$

$allOrders : \mathbb{P}\, Order$

$Stock : \mathbb{P}\, Item$

Client

$id : ID$

$name : Name$

$email : Email$

$password : Password$

$address : Address$

$cell : Cell$

$financials : Financial$

$history : \mathrm{seq}\ Order$

$wishlist : \mathbb{P}\ Item$

$online : \mathbb{B}$

$advertisement : \mathbb{P}\ Items$

$AddClient$

$\Delta knownClient$

$new? : Client$

$new?.id \notin \{c.id \mid c : knownClient\}$

$knownClient' = knownClient \cup \{new?[wishlist \backslash [\,], history \backslash [\,], online \backslash 0]\}$

$new?.id \in knownClient \Rightarrow Error$

$Error$

$\Xi Store$

$message! : seq\ Chars$

$message! = \text{User exists}$

$$
\begin{array}{l}
\textit{Login} \\
\hline
\Delta \textit{Client} \\
\textit{user} : \textit{Client} \\
\textit{username?}, \textit{pass?} : \text{seq } \textit{Chars} \\
\hline
\exists \ \textit{user} : \textit{Client} \\
\quad \textit{user.id} = \textit{username?} \\
\quad \textit{user.password} = \textit{pass?} \\
\quad \textit{user.online} = 0 \\
\quad \textit{user.online}' = 1 \\
\hline
\end{array}
$$

Logout

$\Delta Client$

$cl? : Client$

$cl?.online = 1$

$cl?.online' = 0$

Order

$orderID : OID$
$client : Client$
$time : Time$
$ord : Item \pfun \mathbb{N}^+$
$status : paid \mid dispatched \mid cancelled$
$successful : Item \pfun \mathbb{N}^+$
$orderCost : Currency$

$$
\begin{array}{|l}
\hline
\textit{PlaceOrder} \\
\Delta Store \\
order : Order \\
totalCost! : Currency \\
msg! : \text{seq } Chars \\
\hline
\vdots \\
\hline
\end{array}
$$

$$
\begin{aligned}
&\ldots \\
&order.time' = CurrentTime \\
&order.orderID' = order.orderID \\
&\qquad\qquad\qquad = F(order.client.id, order.time) \\
&\forall\, i\ dom\ order.ord \bullet order.ord(i) \leq stock(i) \\
&\quad \Rightarrow stock(i) := stock(i) - order.ord(i) \\
&i\ dom\ order.ord \bullet order.ord(i) > stock(i) \\
&\quad \Rightarrow order.client.wishlist' = order.client.wishlist \cup \{i\} \wedge msg! = i \text{ out of stock} \\
&totalCost! = \sum_{j \in \{i:Item \mid order.ord(i) \leq stock(i)\}} i.sellingPrice(j) * order.ord(j) \\
&\qquad\qquad = order.orderCost' \\
&order.client.financials' = order.client.financials - totalCost! \\
&order.status' = paid \\
&order.successful' = \{i : Item \mid order.ord(i) \leq stock(i)\} \lhd order.ord \\
&order.client.history' = order.client.history +\!\!+ [order]
\end{aligned}
$$

CancelOrder

$$\Delta Store$$
$$o? : Order$$
$$stock : Items \nrightarrow \mathbb{N}^{+}$$

$$o?.status = paid$$
$$\forall\, i : dom(o?.successful) \bullet stock(i) = stock(i) + o?.successful$$
$$o?.client.financials' = o?.client.financials + o?.orderCost$$
$$o?.client.history' = o?.client.history \setminus [o?]$$
$$o?.status' = cancelled$$

DispatchOrder

$\Delta Order$
$do? : Order$
$m! : \mathrm{seq}\, Chars$

$do?.status = paid$
$do?.status' = dispatched$
$m! = \text{Order dispatched}$

$$
\begin{array}{l}
\textit{EditWishlist} \\
\hline
\Delta Client \\
wl? : Client \\
newIt? : \mathbb{P}\, Item \\
trash? : \mathbb{P}\, Item \\
\hline
wl?.wishlist' = wl?.wishlist \cup newIt? \setminus trash?
\end{array}
$$

$$
\begin{array}{|l}
\hline
\textit{Item} \\
\hline
itemDetails : Details \\
quantity : Item \rightarrow \mathbb{N} \\
costPrice, sellingPrice : Item \rightarrow Currency \\
\hline
\end{array}
$$

Reorder

$$\Delta Stock$$
$$newR? : Item \pfun \mathbb{N}^+$$

$$Stock' = Stock \cup newR?$$

EditStock

$$\Delta Stock$$
$$newS? : Item \pfun \mathbb{N}^+$$
$$garbage? : \mathbb{P}\, Item$$

$$(\mathrm{dom}\ Stock) \cap (\mathrm{dom}\ newS?) = \emptyset$$
$$Stock' = garbage? \ndres Stock$$

You might also like... (Collaborative filtering)

CF algorithms provide a way of quantifying preferences and predicting them in order to make recommendations to active users.

- Assuming if users X & Y both rate n items similarly they will have similar behaviours.
- Typically internet sites use a rating system to gauge users preferences. Such as the 5-star rating system of Amazon and Youtube.
- In the future a rating system will be preferable but a binary system is simple and easy to illustrate the task.
  - If the item is in the user history the rating the user has 'given' the product is 1.
  - If it is not, the user has rated it 0.

User-Item matrix

Client/ID	LOTR	Cinderella	50 Shades of Gray
241	Bought	Ignored	Bought
111	Unknown	Bought	Ignored
81	Bought	Ignored	Bought
Target User	Unknown	Unknown	Bought

Recommendation system

$$
\begin{array}{l}
\textit{Predict} \\
\hline
\Delta Client \\
activeClient? : Client \\
Similar : Client \times kc \rightarrow \mathbb{P}\, Items \\
\hline
activeClient?.advertisement' = Similar(activeClient?, kc)
\end{array}
$$

But what does the function *Similar* look like?

Collaboration vs Content

- Collaboration
- Content
- Hybrids

Measures of sameness

A typical method is to use correlation although dot products can be used too.

$$w_{u,v} = \frac{\sum_{i \in I}(r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I}(r_{u,i} - \bar{r}_u)^2}\sqrt{\sum_{i \in I}(r_{v,i} - \bar{r}_v)^2}} \tag{1}$$

These weights can be used to predict a user's rating of an item.

Challenges in data-analytics

- Data Sparsity: Spread and ‘cold start’ problems
  - Coverage
  - Neighbourhood transitivity
- Scalability
- Synonymy
- Gray and black sheep
- Shilling attacks

Future work

- Ranking architecture that is not boolean.
- Response systems to adverts (Not interested).
- Search history.
- Seasonal aspects.
- Current Affairs.
- Leverage off client details.
- Feature extraction.

References

- Su X, Khoshgoftaar T M. A Survey of Collaborative Filtering Techniques. Advances in Artificial, Vol 2009 (2009).