Predicting Behavior

Ahmed Abbasi, *University of Virginia*

Raymond Y.K. Lau, *City University of Hong Kong*

Donald E. Brown, *University of Virginia*

Behavior prediction is increasingly important, but traditional modeling approaches have shortcomings. A new framework aims to address them by advocating the integration of objective and perceptual information and decomposing behavior into a series of closely interrelated stages.

Predictive analytics is the use of statistical or machine learning methods to make predictions about future or unknown outcomes. While predictive modeling techniques have been researched by the data mining community for several decades, they've become increasingly pervasive in real-world settings in recent years, impacting every facet of our lives. One challenging yet highly important emerging research area is behavior prediction. In some respects, the ability to accurately forecast individuals' decisions and actions remains the "holy grail" of predictive analytics.[1] Use cases abound, ranging from marketing and e-commerce to health, security, economics, infrastructure, public policy, sports, and sociopolitical applications. However, modeling behavior presents some serious challenges:

- Most existing behavior prediction models rely on objective (observed) data. However, behavior is often based on an amalgamation of objective and perceptual considerations. As the saying goes, "perception is reality"—one that can't be ignored.
- Behavioral outcomes are often complex undertakings involving a series of "micro level" antecedent stages, decisions, or actions.

Here, we elaborate on both of these challenges, which we can summarize in data mining terms as representational richness of the feature space and predictive task complexity. Currently, many prediction tasks—such as customer churn, online conversions, user phishing susceptibility, criminal intent, traffic estimation, players' decisions during sporting events, and so on—rely on observed objective data such as transaction logs, sensor-based information, prior multimedia recordings, recent product purchases, historical patterns, past browsing clickstreams, demographics, weather data, socioeconomic indicators, and so forth. Such observed information is undoubtedly important, effective, and reliable. It's also generally easier to collect. However, such observed data can't explicitly surmise what people are thinking, feeling, or experiencing. Perceptual data encompasses an invaluable yet elusive set of independent variables, which if appropriately harnessed, has the potential to significantly bolster prediction capabilities.

For certain behaviors, the decision-making process is fairly habitual.[2] However, a large proportion of human behavior entails elaborate decision processes that can span several stages. These stages can include changing attitudes or intermediary actions, triggered by events, unfolding circumstances, time, or space.[2] Incorporating information about these stages can provide necessary context about links in the decision-making chain. A good example domain where we already

1541-1672/15/$31.00 © 2015 IEEE
Published by the IEEE Computer Society

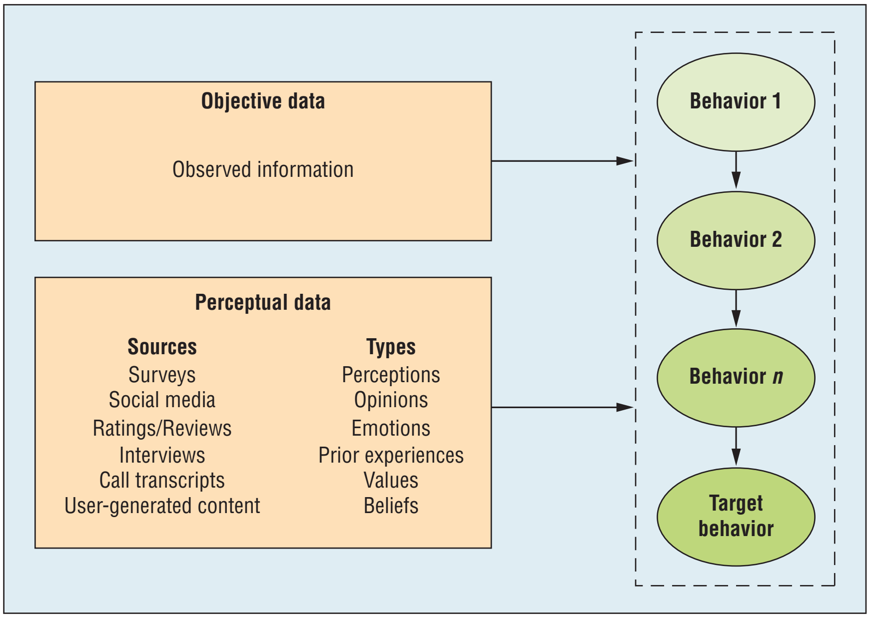

Figure 1. Behavior prediction framework. The framework decomposes a complex behavior prediction task into a sequence of interrelated prediction stages that take into account both perceptual and objective variables.

see widespread adoption of this idea is in e-commerce, where the purchasing process is decomposed into clickstream-based conversion funnels.[3] For example, a Web conversion funnel for an e-tailer might entail stages such as a visit to the homepage, visits to product pages, the addition of items to a shopping cart, account login, checkout, and order confirmation. These funnels aren't only used for descriptive analytics but have important implications for predicting visitors' likelihood of conversion—simply put, in such contexts decomposing the decision-making process into stages can enhance predictive power.[4]

Proposed Behavior Prediction Framework

To address these challenges, we propose the behavior prediction framework depicted in Figure 1. The framework advocates including perceptual data and the decomposing a target/dependent behavior variable prediction task into a sequence of more granular, interrelated prediction tasks. Answering questions such as "What are they thinking?" and "What are they feeling?" is a nontrivial problem. Traditionally, perceptual data collection has necessitated the use of expensive and obtrusive mailed or in-person surveys or interviews. In the era of Big Data, with the volume, velocity, and variety of behavior prediction tasks proliferating, these traditional methods often aren't computationally scalable or logistically feasible. However, with an ever-increasing variety of online crowdsourced survey data collection services, real-time popup surveys, social media, online ratings and reviews, call-center transcripts, and other forms of user-generated content, the inclusion of rich perceptual data such as individuals' intentions, opinions, emotions, prior experiences, values, and beliefs has never been more feasible or worthwhile.

Essentially, our framework advocates decomposing the target behavior variable into a series of interrelated stage variables. We can leverage such granular information in several ways—for example, we can use prediction confidence scores from earlier stages as independent variables for prediction of outcomes in subsequent stages, including the target behavior variable. Alternatively, we could leverage statistical modeling methods such as conditional random fields, Bayesian networks, and hidden Markov models, which are commonly used for structured prediction tasks in domains such as natural language processing.[4] Our proposed framework is intentionally "classifier agnostic"; later, we use different techniques such as composite support vector machine kernels and Bayesian networks to demonstrate our framework's efficacy.

Table 1 presents some examples of applications using the proposed framework. For example, let's assume that existing customers have a negative experience, which might result in a call-center conversation. Inadequate perceived resolution of the issue might prompt negative sentiment expressed via social media or poor reviews, culminating in discontinuation. Drug patients experiencing adverse reactions often initially search for information online and inquire or comment about their experiences via health discussion forums or social networking sites. These discussions often lead to expressions of negative sentiment, increasing the likelihood of an official adverse event report to the regulatory agency such as the US Food and Drug Administration.[5] In the remainder of this article, we present more in-depth examples of how the framework can improve behavior prediction. Examples include predicting individual users' susceptibility to online auction fraud, consumer checkout at e-commerce sites, and customer churn for Internet video-on-demand (VOD) services.

Example 1: Predicting User Susceptibility to Auction Fraud

Online auction fraud remains one of the most commonly reported forms of cybercrime to complaint centers and law enforcement. Auction fraud is a type of semantic attack that exploits human vulnerabilities as opposed

Table 1. Example application areas for behavior prediction framework.

Domain	Prediction problem	Example perceptual data	Example behavior stages
Marketing	Customer churn	* Survey responses about experiences, satisfaction, loyalty * Online ratings/reviews * Social media-based sentiments and emotions about brand and products * Call-center conversation transcripts	Customer acquisition → Negative experience → Call-center conversation → Negative sentiment → Churn
E-commerce	Conversion	* Survey responses about site awareness, trust, perceived credibility, site design, and so on * Social media-based sentiments and emotions about brand and products * Call-center conversation transcripts	Visit → Browse products → Add to cart → Account login → Checkout
Security	Phishing susceptibility	* Survey responses about threat perceptions, prior online experiences, antiphishing security tool perceptions	Visit → Browse → Consider legit → Intend to transact → Provide information
Health	Drug adverse event report	* Social media-based questions, comments, and sentiments * Online ratings and reviews of company or products * Survey responses about drug experiences and side effects	Searching for drug information → Questions/comments about side effects → Negative drug sentiment → Adverse event report

to software vulnerabilities; common forms include failure-to-ship fraud, selling an inferior product, intentionally delaying shipment until the product's value has dropped, and so on.[6] Many fraud prevention mechanisms have emerged to help protect auction sellers and buyers, ranging from detailed buyer/seller ratings and reviews, third-party reputation monitoring sites, and predictive analytics for detecting fraudulent buyers/sellers.[6] Nevertheless, fraud susceptibility remains high, with certain Internet users more susceptible to auction fraud than others. For instance, we examined 10,000 fraud instances occurring on eBay over a five-year period from 2007 to 2012. Fewer than 20 percent of the incidents were associated with victims who were defrauded once—62 percent of incidents were associated with victims who were defrauded four or more times. These findings are consistent with related cybercrime research on antiphishing that has highlighted certain Internet users' persistent inability to avoid phishing website attacks.[7] Consequently, many stakeholder groups are interested in proactively predicting user susceptibility to various forms of cybercrime, including auction fraud, for a couple of reasons:

- improving effectiveness of education in cybercrime prevention programs by customizing curriculum based on users' predicted susceptibility levels; and
- personalizing protective policies and messages based on users' predicted susceptibility levels.

Consistent with the behavior prediction framework presented in Figure 1, Figure 2 shows our proposed auction fraud susceptibility prediction model. The left-hand side depicts the seller, product, and buyer factors that come into play, and the right-hand side depicts the auction fraud funnel—the series of potential decisions and actions taken by prospective buyers when encountering auction fraud. The seller-, product-, and buyer-related factors all encompass objective and perceptual data categories pertaining to specific user encounters with auction fraud, with the latter collected a priori using surveys. Seller information includes objective variables such as how long the seller has been active, number of transactions, number of unique buyers, sale price distribution and product category distribution for past transactions, transaction frequency distribution over time, overall rating, rating time series, and so on. The seller perceptions category includes the buyer's perceived credibility to sellers, perceived reputation of sellers with similar overall ratings, and perceived competence of auction sellers. Product characteristics are comprised of objective variables such as the product category, product type (hedonic, utilitarian), price, quantity available, shipping method, payment method, number of bids, and so forth. The product perception variables include the buyer's perceived interest in purchasing these types of products through online auctions, perceived likelihood of encountering bogus products, awareness of auction fraud, perceived susceptibly, and perceived severity of auction fraud. The demographics include attributes such as age, gender, and education level, whereas prior experience incorporates perceptual variables such as perceived reliance and trust in online auctions, familiarity with the product category and auction site, and past losses attributable to auction fraud.

With respect to behavior stages, the four key decisions or actions undertaken when encountering a fraudulent online auction include visiting the auction page, browsing the seller and product pages, considering the auction's legitimacy, and transacting with the seller. When first encountering the auction

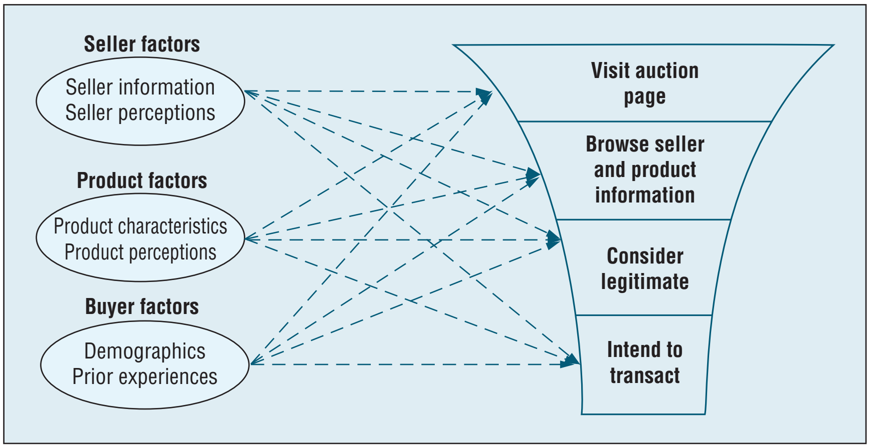

Figure 2. Approach for auction fraud susceptibility prediction. The left-hand side depicts the seller, product, and buyer factors that come into play, and the right-hand side depicts the auction fraud funnel—the series of potential decisions and actions taken by prospective buyers when encountering auction fraud.

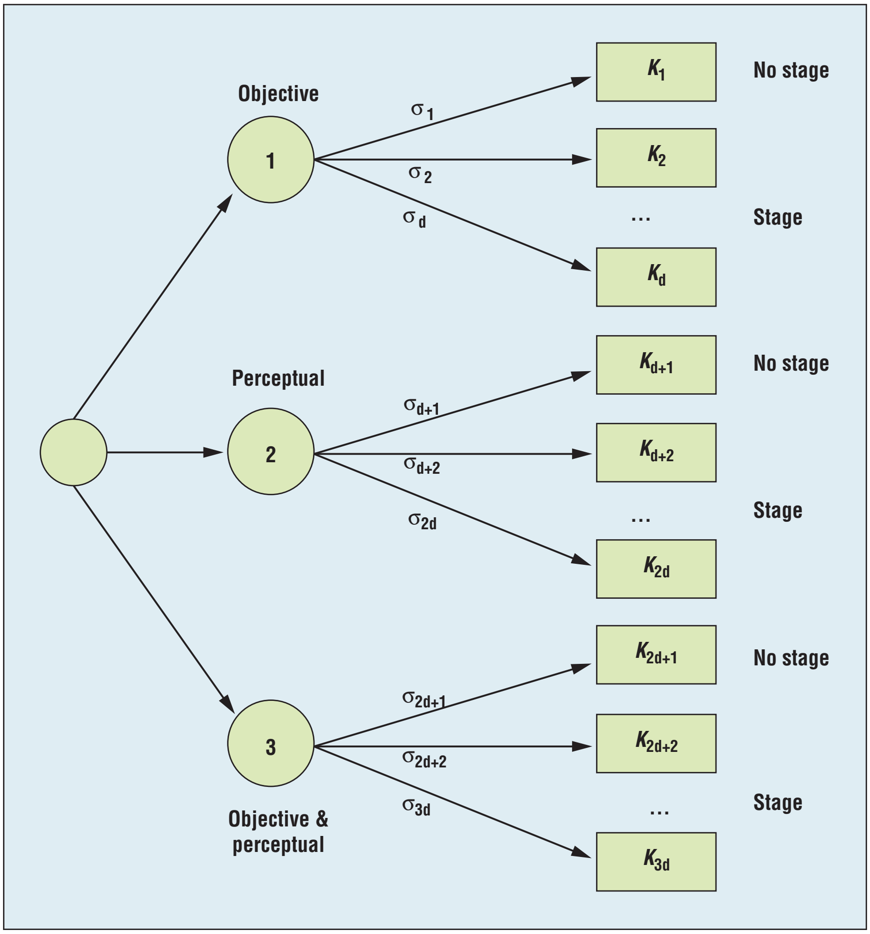

Figure 3. Composite kernel grouping for auction fraud susceptibility detection. The objective and perceptual feature categories and behavior stage variables are used to create classifier ensembles. The model illustrates one of several ways in which the proposed behavior prediction framework can be instantiated.

page, prospective buyers must decide whether to click on the link to visit the auction page. For those who choose to visit, they must then decide whether to browse the auction/product/seller pages or to exit. Similarly, for the third and fourth stages of the funnel, users must consider seller/auction legitimacy before possibly considering engaging in transactions (that is, the target variable).

As mentioned in our discussion of the behavior prediction framework, the framework is intentionally classifier agnostic to ensure generalizability. In this context, a support vector machine (SVM) classifier with a composite kernel function is employed to facilitate incorporation of the key elements of the approach highlighted in Figure 2. Given an input space X—in this case, the set of all possible user-fraud encounter instances to be examined—we can formulate the learning problem as finding a classifier C: $X \to Y$, where Y is the class label to be assigned to the data points (in this case, susceptible, not susceptible). For kernel-based methods such as SVM, finding C relies on a kernel function K that defines a mapping K: $X \times X \to [0, \infty)$ from the input space X to a similarity score $K(x_i, x_j) = f(x_i) \times f(x_j)$, where x_i and x_j represent two data points—in this case, two user-fraud interaction instances—and $f(x_i)$ is a function that maps X to a higher-dimensional space without needing to know its explicit representation. The composite kernel $K\sigma$ is a convex combination of M kernels, $K_1, \ldots, K_M$, defined as follows[8]:

$$K_\sigma = \sum_{m=1}^{M} \sigma_m K_m ,$$

where

$$\sigma_m \geq 0, \ \sum_{m=1}^{M} \sigma_m = 1.$$

Figure 3 shows an illustration of the kernel. Using a two-level hierarchy, we group the kernels based on the input spaces (that is, feature sets) utilized: objective, perceptual, and all features. This creates something analogous to a weighted feature-based ensemble of SVM classifiers. Furthermore, given d behavior stages (including the target behavior), when predicting any subsequent stage $x \geq 2$, additional stage kernels K_2, …, K_x, K_{x+2}, …, K_{2x}, and K_{2x+2}, …, K_{3x} are incorporated, where the input space for K_x includes predictions for all prior stages 1, …, $x - 1$, and $\sigma_m = 1/M$.

Data were gathered for 3,000 incidents involving buyer auction fraud encounters pertaining to two diverse domains: consumer electronics and collectible dolls. For each encounter, the data included the objective and perceptual variables described in the prior paragraph as well as the user decisions and actions for the four funnel stages. The perceptual data was gathered in conjunction with our industry partner, a major North American online auction website, using surveys about seller and product perceptions, as well as prior experiences. These 3,000 incidents helped train the SVM classifier via a composite kernel. A separate set of 2,000 encounters were used as the test set.

To examine the utility of the behavior prediction framework, we used three different feature sets: objective (seller and product information plus demographics), perceptual (seller and product perceptions and prior experiences), and objective plus perceptual. Similarly, two different behavior stage approaches were employed: inclusion of prior stages' predictions as variables in subsequent stages or no inclusion. This resulted in six total experiment settings. The kernel hierarchy depicted in Figure 3 was adjusted based on the settings. The evaluation metrics included accuracy (AC) and class-level recall for users' "continue" or "exit" decisions (labeled CR and ER).

Table 2 shows the results. When using prior stages' predictions as input variables for subsequent stages, coupled with perceptual and objective features, overall AC and CR and ER recall rates were highest as evidenced by the values depicted in the first row of the table. These settings, which are in line with the proposed behavior prediction framework, boosted performance by nearly 20 percent over the "No" stage variable, under the objective data-only baseline (last row of Table 2). The performance improvement relative to the other rows underscores the utility of using perceptual and objective data in unison and in decomposing the target behavior task into a series of interrelated stages. The performance gains have important implications for auction fraud prevention, where the ability to better predict user susceptibility to fraud can translate into enhanced protective measures.

Examples 2 and 3: Predicting E-Commerce Conversions and Churn for VOD

Both the volume of transactions and the revenues of e-commerce companies have undergone rapid growth in recent years. For instance, the combined revenue for Taobao and Tmall, two Alibaba-run shopping sites, reached a whopping \$240 billion in 2013. There's considerable (and obvious) business value for e-commerce companies to predict their customers' purchasing behavior and then leverage such information as part of corresponding conversion rate enhancement, upsell, cross-sell, and lifetime value maximization strategies. On the other hand, predicting customer churn is a classic problem that's especially important in high-churn industries such as telecommunications, cable, and e-commerce-based VOD, where service providers constantly face the pressure of retaining existing customers.

Figure 4a highlights a model of predicting consumer checkout behavior at e-commerce sites. The four main behavioral stages undertaken by consumers include visiting an online shop, browsing products and product catalogs, adding selected products to shopping carts, and finally checking out.[4] Three categories of depicted factors impact customer conversions—namely, e-commerce site-specific factors, product factors, and consumer factors. Site factors include both perceptual variables such as consumers' perceived credibility of an online seller and objective variables such as an online shop's website. Product factors include perceptual variables such as sentiments in product-related reviews or comments, and objective variables such as price, category, type of product, and so on. Consumer factors include perceptual variables such as consumers' prior purchase experience for certain online shops and objective variables such as age and gender.

Figure 4b shows an instantiation of the proposed behavior prediction framework for VOD service churn prediction. The three main stages are viewing streaming videos, posting feedback about the viewings, and switching service providers (that is, the target churn variable). In this case, the earlier stage variables provide indicators of customer engagement (viewing) and satisfaction (feedback) levels; lack of engagement and dissatisfaction are antecedents for churn. The three categories of independent variables employed are channel, viewing, and subscriber factors. Channel factors include both perceptual variables, such as subscribers' perceived quality of an online channel, and objective variables, such as the channel's inherent characteristics (for example, the frequency of adding

Table 2. Auction fraud susceptibility prediction results using different settings.*

Stage variables	Features	Visit			Browse			Consider legit			Intend to transact		
		AC	CR	ER	AC	CR	ER	AC	CR	ER	AC	CR	ER
Yes	All	**78.5**	**79.2**	**75.9**	**81.2**	**80.5**	**83.1**	**83.3**	**85.5**	**80.2**	**85.6**	**88.2**	**83.4**
	Perceptual	76.7	77.6	73.4	78.5	78.4	78.9	80.9	82.4	78.7	81.4	83.6	79.4
	Objective	69.2	70.1	66.0	66.9	67.8	64.7	68.0	70.2	64.7	72.7	69.5	75.4
No	All	**78.5**	**79.2**	**75.9**	79.0	78.8	79.6	79.5	80.6	77.8	80.6	82.3	79.1
	Perceptual	76.7	77.6	73.4	72.4	75.4	65.2	72.9	79.8	62.7	74.4	77.7	71.4
	Objective	69.2	70.1	66.0	64.8	65.5	63.2	64.0	64.1	63.8	66.3	68.2	64.6

* AC = accuracy, CR = continue decision, and ER = exit decision.

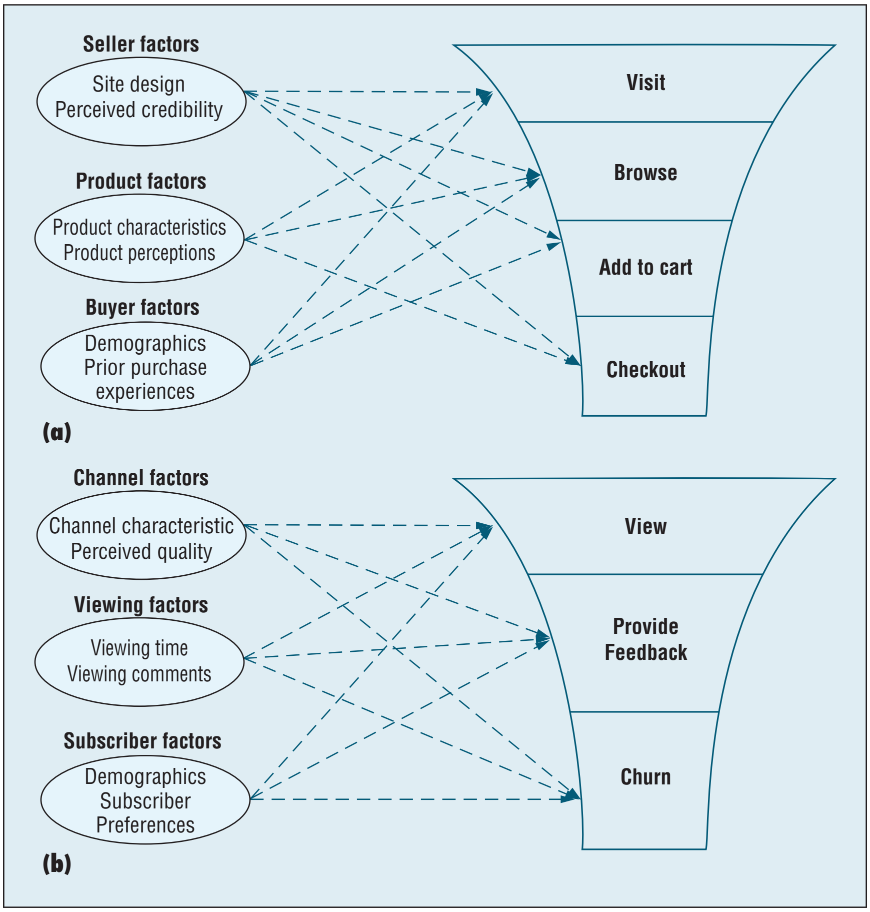

Figure 4. Models of customer behavior prediction: (a) customer conversion prediction in e-commerce and (b) customer churn prediction for Internet video-on-demand services. These models highlight the interrelated behavioral stages and the objective/perceptual variables influencing various stages.

new movies). Viewing factors refer to perceptual variables, such as the sentiments about channels and videos, and objective variables, such as the total time of viewed videos. Subscriber factors include perceptual variables, such as subscribers' preferences for certain kinds of videos, and objective variables, such as the demographics of subscribers.

Each variable can influence consumer behavior at multiple stages. Implementation-wise, there are arrays of statistical and machine learning methods to be chosen from. We employ Bayesian networks because they're adept at capturing the relationships among different behavioral stages in conjunction with the causal relationships between perceptual and objective variables and behavioral stages. They also offer an intuitive way to explain the underlying logic behind the predictions. Figure 5 shows a general Bayesian network model for predicting consumers' checkout behavior in e-commerce. Although the depicted network corresponds to the example in Figure 4a, a similar network model was also used for the churn prediction example in Figure 4b. A directed arc represents the causal relationship from a cause node to an outcome node in Figure 5. Each individual variable (causal factor) is represented as a separate node in our Bayesian network, but for brevity, we depict each category of variables as a single node in the figure. To achieve a good tradeoff between computational efficiency and prediction performance, some variables are assumed to influence a certain behavioral stage only, and the independence assumption holds among perceptual and objective variables. The actual Bayesian network structure was created via a manual knowledge-engineering process.

By taking the first-order Markov chain assumption, the joint probability distribution of the set of variables (nodes) $V = \{x_1, x_2, \ldots, x_{|V|}\}$ captured in a Bayesian network G and the inference rule of G are depicted as follows:

$$p(x_1, x_2, \cdots, x_{|V|} \mid G) = \prod_{i=1}^{|V|} p(x_i \mid \mathbf{x}_{pa(i)})$$

$$p(\mathbf{x}_t \mid \mathbf{x}_k) = \frac{p(\mathbf{x}_t, \mathbf{x}_k)}{p(\mathbf{x}_k)} = \frac{p(\mathbf{x}_t, \mathbf{x}_k)}{\sum_{\mathbf{x}'_t} p(\mathbf{x}'_t, \mathbf{x}_k)}.$$

The notation $\mathbf{x}_{pa(i)}$ represents the set of parent nodes of the node $\mathbf{x}_i$. Moreover, $\mathbf{x}_t$ and $\mathbf{x}_k$ are the set of targeting prediction variables and the set of known variables, respectively. Basically, inference in a Bayesian network is operationalized by conditioning on the data via clamping the known variables $\mathbf{x}_k$ with their observed values, and then normalizing to derive $p(\mathbf{x}_t \mid \mathbf{x}_k)$ from $p(\mathbf{x}_t, \mathbf{x}_k)$. If each node x_i has at most F parents and each parent has at most K states, the general computational complexity of model learning and inference of G is characterized by $O(VK^F)$. However, if a Bayesian network has a tree-like structure as the one shown in Figure 5, the computational complexity is reduced to $O(VK^2)$.[9] Accordingly, our proposed behavior prediction model is efficient, and the prediction process consumes time linear with respect to the number of nodes in the network. For this example, a Bayesian classifier that operates according to the aforementioned formulas is applied to predict customers' checkout decisions in e-commerce.

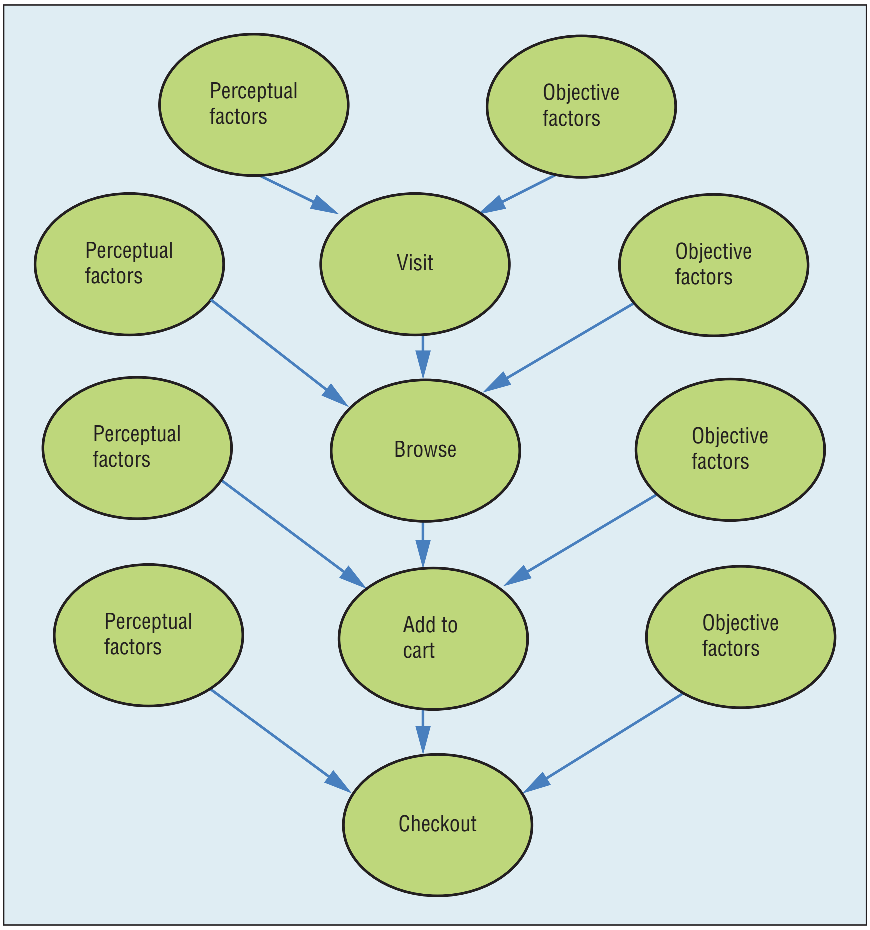

Figure 5. A general Bayesian network model for consumer behavior prediction in e-commerce. A directed arc represents the causal relationship from a cause node to an outcome node. Although each individual variable (causal factor) is represented as a separate node in our Bayesian network, for brevity, we depict each category of variables as a single node. The model represents another illustration of how the proposed behavior prediction framework can be instantiated.

To evaluate the proposed behavior prediction model in the context of e-commerce conversions and VOD churn, we constructed testbeds for each task. For e-commerce conversions, we retrieved the perceptual and objective data of consumers through our research partner, a large e-commerce company in China. Some perceptual data were collected through a survey performed by our research partner. In addition, product sentiments were extracted from the product reviews posted to various online shops hosted on their e-commerce platforms. More specifically, a domain-specific sentiment analysis method was applied to extract these product sentiments.[10] Overall, there were a total of 11,650 observations in our e-commerce conversion dataset, with the earliest 60 percent used for training and the subsequent 40 percent reserved for testing.

To construct the VOD churn dataset, we retrieved the behavior and objective data of subscribers from our research partner, a large Internet VOD service provider. Some behavioral data (such as perceived quality about an online channel) were collected by our research partner through its regular online service survey. In addition, the sentiments about channels or videos were extracted from the online comments posted to the VOD service website using a domain-sensitive sentiment extraction method.[10] There were a total of 5,410 observations in the VOD churn dataset, again with the earliest 60 percent used for training and the subsequent 40 percent for testing.

We used a similar setup to the prior experiment, with six total settings for the Bayesian inference network. We adopted standard performance evaluation metrics such as precision (P), recall (R), and F-score (F). Because each class label of our dataset only consists of two possible values (such as checkout versus not checkout), the performance scores

Table 3. Experimental results of various methods for e-commerce conversion and VOD churn prediction tasks.*

		E-commerce conversion prediction											
		Visit			Browse			Add to cart			Checkout		
Stage variables	Features	R	P	F	R	P	F	R	P	F	R	P	F
Yes	All	**75.7**	**75.3**	**75.5**	**77.7**	**76.6**	**77.1**	**78.5**	**79.1**	**78.8**	**83.5**	**84.6**	**84.1**
	Perceptual	73.4	75.3	74.3	75.0	75.3	75.2	76.6	75.9	76.3	80.1	82.7	81.8
	Objective	68.5	68.9	68.7	70.4	70.1	70.2	71.8	73.6	72.7	78.2	79.5	78.8
No	All	**75.7**	**75.3**	**75.5**	75.7	75.8	75.8	76.6	76.1	76.3	75.2	73.6	74.4
	Perceptual	73.4	75.3	74.3	73.2	72.7	72.9	74.0	73.6	73.8	72.8	72.4	72.6
	Objective	68.5	68.9	68.7	68.7	68.4	68.5	70.9	70.2	70.6	69.2	68.2	68.7
		VOD churn prediction											
		View			Feedback			Churn					
Stage Variables	Features	R	P	F	R	P	F	R	P	F			
Yes	All	**76.8**	**75.8**	**76.3**	**79.4**	**76.9**	**78.2**	**83.2**	**82.4**	**82.8**			
	Perceptual	75.2	74.8	75.0	77.4	75.5	76.4	82.4	80.4	81.4			
	Objective	73.1	71.5	72.3	73.9	72.5	73.3	80.5	78.3	79.4			
No	All	**76.8**	**75.8**	**76.3**	76.9	76.3	76.6	75.4	75.5	76.8			
	Perceptual	75.2	74.8	75.0	75.2	74.9	75.1	74.2	73.1	73.6			
	Objective	73.1	71.5	72.3	71.8	72.2	72.0	71.5	73.1	72.3			

* F = F-score, P = precision, and R = recall.

refer to predicting the positive outcome for each target class (for example, visit = true or checkout = true). Table 3 summarizes our experimental results.

Based on the values depicted in the first row for each testbed, the proposed behavior prediction method, which combines objective and perceptual features and includes decision stage variables, outperforms other baseline methods across stages in both cases: e-commerce conversions and VOD churn. In particular, the proposed method can boost performance of checkout prediction and VOD churn detection by 9.7 percent and 6.0 percent, respectively, in terms of F-score over its stage-less counterpart (the fourth rows of Table 3). Our experiment results once again reveal that perceptual features markedly outperform the use of objective attributes alone. In the context of VOD churn prediction, based on our postexperiment analysis, four variables with high discriminatory potential were dwindling subscriber viewing activity, lack of posted comments for an extended period of time, strong negative sentiments in viewing comments, and perceived channel quality. Overall, the results underscore the utility of using prior engagement-oriented behavior stages' predictions, coupled with objective and perceptual factors as input variables to predict subscriber's churn behavior. Once again, our experimental results confirm that perceptual features are more useful than objective evidence for predicting subscriber behavior. This is especially interesting considering that such service providers routinely ascertain perceptual data via regular online surveys.

Results Discussion

In all three of the examples we presented, incorporating perceptual and objective data in unison, coupled with the inclusion of decision stage information, resulted in markedly improved predictive power. This enhanced predictive capability has important implications for practitioners. In the context of online conversion rate analysis, models that are 10 percent more accurate can allow e-commerce companies to more accurately deploy exit-avoidance strategies such as real-time promotions and webpage personalization. For a mid-sized e-commerce company, this can translate into increased annual sales revenue of more than $20 million. Similarly, the enhanced churn prediction performance depicted for the VOD service provider (see the bottom half of Table 3) can enable the deployment of more effective customer retention campaigns. Such analytically-driven initiatives have been shown to reduce churn rates by more than half.[11]

The main contributions of this work are twofold. First, we advocate a novel behavior prediction framework that captures the interrelated microbehavioral stages and leverages a fusion of perceptual and objective features. Second, we validate the effectiveness of our proposed framework through instantiations of it in three different application areas. We show that it's relatively straightforward to model the interrelated microbehavioral stages of prediction tasks arising in different application areas. Moreover, rich perceptual data (such as users' subjective comments or survey responses) are often readily available to enhance the overall prediction performance.

It's important to reiterate that the practical implications for enhanced behavior prediction extend beyond e-commerce and marketing. Better predicting user behavior in health and security contexts could translate into far-reaching societal benefits.[5] Beyond the practical implications, for academic research, predicting behavior can help inform behavioral theory.[2] Why a certain predictive model works better (specific perceptual predictors and their interplay with decision stages) could enhance our understanding of behavioral phenomena in various application domains.[1]

Behavior prediction has received increasing attention in a wide variety of application areas, including marketing, e-commerce, e-health, cybersecurity, economics, public policy, and sports. Existing behavior prediction models often lack the representational richness of both perceptual and objective features, as well as the interrelated microbehavioral stages of a prediction task.

Future work will improve the external validity of our research by applying the proposed framework to more application domains such as cybersecurity and health. For instance, in the case of many cybersecurity threats such as malware and phishing, users are often considered the proverbial "weak links" in the security chain. Recent evidence suggests that factors such as demographics, personality traits, awareness, prior history, and perceived vulnerability to attacks are strong predictors of future susceptibility.[12] Similarly, in the health arena, users' overall sentiments and emotions, accounts of pharmaceutical drug-related experiences, and online search behavior are strong predictors of adverse drug reactions.[5] As understanding and inferring behavior becomes increasingly important and pervasive, the framework proposed here signifies an important first-step towards enhanced predictive analytics. Ultimately, our hope is that future work will build on this foundation. ■

THE AUTHORS

Ahmed Abbasi is an associate professor of information technology, director of the Center for Business Analytics, and a member of the Predictive Analytics Lab at the University of Virginia. His research interests include data mining, machine learning, text mining, social media analytics, health, and Big Data. Abbasi is a senior member of IEEE and an associate editor for *ACM Transactions on MIS, IEEE Intelligent Systems, Information Systems Research,* and *Decision Sciences Journal*. Contact him at abbasi@comm.virginia.edu.

Raymond Y.K. Lau is an associate professor in the Department of Information Systems and associate director of the Center for Intelligent Information Systems at City University of Hong Kong. His research interests include information retrieval, text mining, and social media analytics. Lau is a senior member of IEEE and a life member of the Hong Kong Computer Society. Contact him at raylau@cityu.edu.hk.

Donald E. Brown is the director of the Data Science Institute and W.S. Calcott Professor in the Department of Systems and Information Engineering at the University of Virginia. His research focuses on data fusion, statistical learning, and predictive modeling with applications to security and safety. Brown is a fellow of IEEE, past president of the IEEE Systems, Man, and Cybernetics Society, and former editor in chief of *IEEE Transaction on Systems, Man, and Cybernetics, Part A: Systems and Humans*. He's also the recipient of the Norbert Wiener Award and IEEE Millennium Medal. Contact him at deb@virginia.edu.

Acknowledgments

This work was funded in part by the following grants from the US National Science Foundation: IIS-1236970, CNS-1049497, and ACI-1443019. Raymond Lau's work was partially supported by a grant from the Research Grants Council of the Hong Kong SAR (project CityU 145712) and the Shenzhen Municipal Science and Technology R&D Funding—Basic Research Program (project JCYJ20130401145617281 and JCYJ20140419115614350).

References

1. G. Shmueli, "To Explain or to Predict?" *Statistical Science*, vol. 25, no. 3, 2010, pp. 289–310.
2. H. Aarts, B. Verplanken, and A. Knippenberg, "Predicting Behavior from Actions in the Past: Repeated Decision Making or a Matter of Habit?" *J. Applied Social Psychology*, vol. 28, no. 15, 1998, pp. 1355–1374.
3. A. Kaushik, *Web Analytics 2.0: The Art of Online Accountability and Science of Customer Centricity*, Wiley Publishing, 2011.
4. A.L. Montgomery et al., "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, vol. 23, no. 4, 2004, pp. 579–595.
5. D. Adjeroh et al., "Signal Fusion for Social Media Analysis of Adverse Drug Events," *IEEE Intelligent Systems*, vol. 29, no. 2, 2014, pp. 74–80.
6. B. Gavish and C.L. Tucci, "Reducing Internet Auction Fraud," *Comm. ACM*, vol. 51, no. 5, 2008, pp. 89–97.
7. A. Abbasi et al., "Enhancing Predictive Analytics for Anti-Phishing by Exploiting Website Genre Information," to be published in *J. MIS*, vol. 31, no. 4, 2015.
8. M. Szafranski, Y. Grandvalet, and A. Rakotomamonjy, "Composite Kernel Learning," *Machine Learning*, vol. 79, nos. 1 and 2, 2010, pp. 73–103.
9. K. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.
10. R.Y.K. Lau et al., "Web 2.0 Environmental Scanning and Adaptive Decision Support for Business Mergers and Acquisitions," *MIS Quarterly*, vol. 36, no. 4, 2012, pp. 1239–1268.
11. M.S. Hopkins and L. Brokaw, "Matchmaking with Math: How Analytics Beats Intuition to Win Customers," *Sloan Management Rev.*, vol. 52, no. 2, 2011, pp. 35–41.
12. F.M. Zahedi, A. Abbasi, and Y. Chen, "Fake-Website Detection Tools: Identifying Elements that Promote Individuals' Use and Enhance Their Performance," to be published in *J. AIS*, 2015.