Personalization Technologies: A Process-Oriented Perspective

By leveraging customer reactions to personalized products and services, companies continuously improve their personalization processes through an iterative feedback loop resulting in the 'virtuous cycle' of personalization.

By Gediminas Adomavicius and Alexander Tuzhilin

Much recent research and development in personalization has sought to create new technologies, understand personalization from a business perspective, and develop novel applications [11]. However, since personalization is a rapidly developing field, there still exist several points of view among researchers and practitioners as to what it actually is. Here, we synthesize these views, employing a process-oriented perspective to explore the field. The following are representative definitions:

- "Personalization is the ability to provide content and services tailored to individuals based on knowledge about their preferences and behavior" [Paul Hagen, Forrester Research, 1999];
- "Personalization is the use of technology and customer information to tailor electronic commerce interactions between a business and each individual customer. Using information either previously obtained or provided in real time about the customer, the exchange between the parties is altered to fit that customer's stated needs, as well as needs perceived by the business based on the available customer information" [Personalization Consortium, 2003];
- "Personalization is the capability to customize customer communication based on knowledge preferences and behaviors at the time of interaction [with the customer]" [Jill Dyche, Baseline Consulting, 2002]; and

- “Personalization is about building customer loyalty by building a meaningful one-to-one relationship; by understanding the needs of each individual and helping satisfy a goal that efficiently and knowledgeably addresses each individual’s need in a given context” [Doug Riecken, IBM, 2000].

Several important insights concerning personalization emerge from them. Collectively, they maintain that personalization tailors certain offerings (such as content, services, product recommendations, communications, and e-commerce interactions) by providers (such as e-commerce Web sites) to consumers (such as customers and visitors) based on knowledge about them, with certain goal(s) in mind.

Personalization takes place between one or several providers of personalized offerings and one or several consumers. Personalized offerings are delivered from providers to consumers through personalization engines by adapting one of three architectural approaches: provider-centric, consumer-centric, and market-centric (see Figure 1). The figure denotes providers and consumers of personalized offerings by white boxes, personalization engines by shaded boxes, and interaction between consumers and providers by lines.

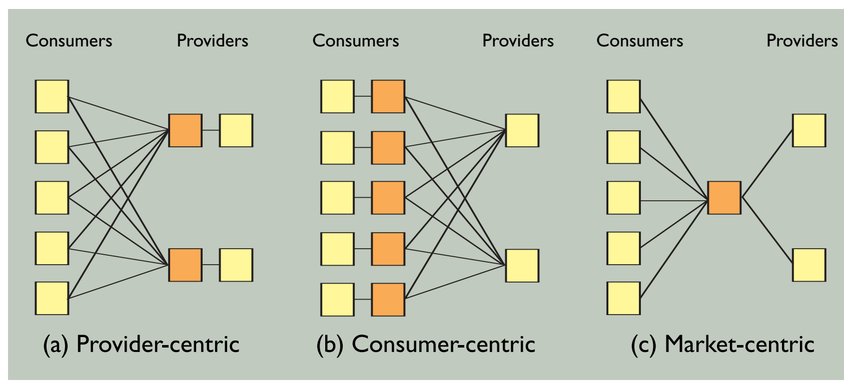

Figure 1. Classification of personalization approaches.

Figure 1(a) outlines a provider-centric personalization approach that assumes each provider has its own personalization engine to tailor its content to consumers. Provider-centric is the most common approach to personalization, as popularized by Amazon.com. The second, outlined in Figure 1(b), is consumer-centric, which assumes each consumer has his or her own personalization engine (or agent) that “understands” this particular consumer and provides personalization services across several providers based on this understanding. Such consumer-centric personalization delivered across a range of providers and offerings is called an e-Butler service [2]. The third approach, in Figure 1(c), is market-centric, providing personalization services for a specific marketplace in a particular industry or sector. The personalization engine is the infomediary, knowing the needs of the consumer and the provider’s offerings and trying to match them in the best ways possible according to their internal goals. While this approach is not used extensively, it has potential, especially in the context of various targeted e-marketplaces (such as Covisint, www.covisint.com, for the automotive and health care industries).

Although the three approaches are general enough to be deployed in online or offline personalization applications, they are in practice mainly applicable only to the online world. The reason for this is that personalization technologies in general are information-intensive; that is, they require the speedy collection and processing of large volumes of data about consumers, providers, and markets, as well as a quick response to the results of this analysis. Systems involving these requirements are implemented more effectively in the online world than the offline.

The result of personalization is the delivery of various offerings to consumers by the personalization engine(s) on behalf of the providers using any of the approaches in the figure. Personalized offerings include:

- Content (such as Web pages and links);
- Product and service recommendations (such as for books, CDs, and vacations);
- Email;
- Information searches;
- Dynamic prices; and
- Products for individual consumers (such as custom CDs).

Personalization objectives generally range from simply improving the consumer’s browsing and shopping experience (such as by presenting only content relevant to the consumer) to much more complex objectives (such as developing long-term relationships with consumers, improving consumer loyalty, and generating measurable value for the company).

Successful personalization applications depend on knowledge about consumers’ personal preferences and behavior typically distilled from volumes of granular information about them and stored in the form of consumer profiles [3, 9]. However, the idea that personalization constitutes an iterative process is missing from them and is not sufficiently addressed in the personalization literature. This is unfortunate, because personalization is indeed an iterative process. After personalized offerings are delivered, consumer satisfaction with them can be measured, and the consumer-response information used to adjust personalization strategies in an iterative feedback loop.

Personalization Process

Personalization constitutes an iterative process that can be defined by the three stages of the understand-deliver-measure cycle (see Figure 2):

Understand consumers by collecting comprehensive information about them and converting it into actionable knowledge stored in the form of consumer profiles;

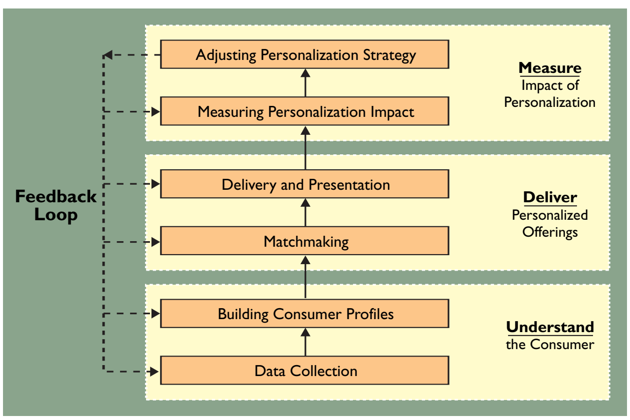

Figure 2. Personalization process.

Deliver personalized offerings based on the knowledge about each consumer, as stored in the consumer profiles; the personalization engine must be able to find the most relevant offerings and deliver them to the consumer; and

Measure personalization impact by determining how much the consumer is satisfied (and dissatisfied) with the delivered offerings. This measure provides information that enhances our understanding about consumers or identifies the deficiencies of the methods for personalized delivery. It also serves as feedback for possible improvements to each of the other components of the personalization process. When it completes one cycle of the process, it sets the stage for the next cycle in which improved personalization techniques are able to make better personalization decisions.

The technical implementation of the understand-deliver-measure cycle consists of the six stages outlined in Figure 2. The personalization process begins with the collection of data across different channels of interaction (such as the Web, phone, and direct mail) between consumers and providers, as well as from various other heterogeneous data sources, looking to obtain the most comprehensive picture of individual consumers.

Once the data is collected, a key issue in developing personalization applications is how to construct accurate, comprehensive consumer profiles based on it. Once the profiles are constructed, personalization systems must be able to use them to match appropriate content and services to individual consumers. Technologies for matchmaking include recommender systems, statistics-based predictive approaches, and rule-based systems where an expert specifies business rules governing delivery of content and services that depend on conditions stipulated by the rules.

Following matchmaking, personalized information is delivered to consumers. Delivery is accomplished in several ways using visualization, lists ordered by relevance, unordered lists of alternatives, and more. Next, various metrics (such as accuracy, consumer lifetime value, loyalty value, and purchasing experience) can be used to evaluate the effectiveness of personalized offerings. The quality of recommendations, as measured by these metrics, depends on the sophistication of technologies deployed in the previous four stages of the personalization process—from data collection to the delivery and presentation stages (see Figure 2).

Finally, the metrics can be used to improve each of the five other stages of the overall process. This feedback should be used to decide whether to collect additional data, build better profiles, develop better matchmaking algorithms, improve information delivery, or use additional personalization metrics. If it is properly integrated into the personalization process**,** the quality of interactions with individual consumers, as measured by the metrics, should improve over time, resulting in the virtuous cycle of personalization.[1]

If the virtuous cycle is achieved, personalization becomes a powerful process for delivering ever-increasing value to the stakeholders. Besides being essential for improving personalized service over time, it is also crucial to the personalization system's ability to keep up with the constantly changing environment (such as to adjust to changes in the tastes and preferences of individual customers, as well as in consumer fashion and product offerings).

The opposite of the virtuous cycle is the process of depersonalization, which can occur when the metrics of consumer satisfaction are low from the start, when they decrease over time, or when the system is unable to adjust in time to the changing environment (such as to

[1]The "virtuous cycle" is a situation in which improvement in one element of a chain of circumstances leads to improvement in another element, which leads to further improvement in the original element, and so on (www.wordspy.com/words/virtuous-cycle.asp).

changing consumer tastes). Consumers get so frustrated with the personalization system they stop using it. Depersonalization is largely responsible for the failures of some otherwise promising projects. One of the main challenges for personalization developers is how to achieve the virtuous cycle and not fall into the depersonalization trap.

Understanding the Consumer

Data about consumers is collected through multiple channels (such as the Web, phone, and direct mail) to generate the most comprehensive and useful picture of a consumer. It can be solicited explicitly (such as via surveys) or tracked implicitly and may include histories of consumer purchasing and online search, as well as demographic and psychographic information. After the data is collected, it is usually processed, cleaned, and stored in a consumer-oriented data warehouse for analysis.

The data is used to build consumer profiles. The profiles should include the information describing who the consumers are and how they behave in a way that is most relevant for the particular personalization application(s). Most personalization systems represent consumer profiles in terms of collections of facts about consumers, including demographics (such as name, gender, date of birth, and address). The facts can also be derived from past transactions (such as favorite product category or the value of the largest purchase made at a particular e-commerce Web site). This information is often defined as a record of values and stored in a relational database, one record per consumer.

However, due to the static nature of factual information, simple profiles specified through collections of facts may be insufficient in certain more complex personalization applications (such as high-precision personalized content delivery) and some advanced recommendation applications in certain industries (such as financial services, travel, and telecommunications). These applications may therefore require deployment of more advanced profiling techniques supporting the following methods for capturing more complex rules, sequences, signatures, and other advanced profiling methods for capturing more complex behavioral information about consumers:

Rules. Consider the rule "John Doe prefers to see action movies on weekends" (*Name = "John Doe" & MovieType = "action" → TimeOfWeek = "weekend"*), which can be a part of John Doe's profile, describing his movie-viewing habits [3]. Such rules can be learned from the transactional history of the consumer (John Doe in this case) using data mining techniques, including association and classification rule discovery methods [7].

Sequences. One may want to store in John Doe's profile his typical browsing sequence: "When John Doe visits the book Web site XYZ, he usually starts with the home page, then goes to the Home&Gardening section, then the Gardening section, and then leaves the site" (*XYZ: StartPage → Home&Gardening → Gardening → Exit*). This sequence might be learned from the transactional histories of consumers using frequent episodes [7] and other sequence-learning methods.

Signatures. Signatures are the statistics-based data structures used to capture the evolving behavior learned from large data streams of simple transactions [6]. For example, "top five most frequently browsed product categories over the past 30 days" is an example of a signature that could be stored in individual consumer profiles in an e-commerce application.

In some advanced personalization applications, the use of only simple factual consumer profiles is insufficient for accurately capturing consumer tastes and preferences; more advanced profiling techniques (such as rule-, sequence-, and signature-based methods) must be deployed to capture behavioral information. Although some research and development in these methods has been done [1, 6], more is required to better understand their applicability for consumer profiling and their integration into the personalization process.

Delivering Personalized Offerings

E-businesses usually deliver personalized information to consumers in several ways, including narratives, lists ordered by relevance, sets of alternatives, and various types of visualization. One classification of delivery methods is known as "pull, push, and passive" [12]. Pull methods notify consumers that personalized information is available but display it only when consumers explicitly request it. Push methods (such as sending email) are intended to reach consumers not already interacting with the personalization system (such as through a company's e-commerce Web site). And passive delivery displays personalized information as by-products of other consumer activities; for example, a consumer looking at a product on a Web site is also shown recommendations for related products.

The performance of these delivery and presentation methods depends on the quality of underlying matchmaking technologies. Much recent research has focused on this subject, including rule-based matchmaking, statistics-based predictive approaches, and recommender systems. Several industrial personalization solutions, initially developed in the mid-1990s by BroadVision (www.broadvision.com) and subsequently integrated into various personalization servers, support rule-based matchmaking where the rules are defined by a domain expert. For example, a marketing manager may specify that if a consumer of a certain type visits an online grocery store on a Sunday night, this person should also be offered discount coupons for baby diapers. Here, we focus on recommender systems because they represent the most developed matchmaking technologies.

Much research over the past decade has focused on developing recommendation-based matchmaking technologies [1, 5, 8, 9, 12] featuring methods from such disciplines as human-computer interaction, statistics, data mining, machine learning, and information retrieval. These methods are often classified into broad categories according to their recommendation approach, as well as to their algorithmic technique; [4] classifies them based on recommendation approach as:

Content-based recommendations. The system recommends items to consumers (such as content, services, and products) like the ones they preferred in the past. Content-based methods analyze the common features among the items a consumer has already rated highly. Only the items similar to consumer's past preferences are then recommended;

Collaborative recommendations (or collaborative filtering). The system recommends items to the consumer that people with similar tastes and preferences have liked in the past. Collaborative methods find the closest peers for each consumer, that is, the people with the most similar tastes and preferences. Only the items most liked by the peers are then recommended; and

Hybrid approaches. Can be realized in several ways by combining collaborative and content-based methods. One is to implement separate content-based and collaborative systems, then combine their results to produce the final recommendations. Another is to use content-based and collaborative techniques in a single recommendation model, rather than implementing them separately.

Recommender systems methods can be classified [5] based on the algorithmic technique as:

Heuristic-based. These techniques constitute heuristics that calculate recommendations based on consumers' previous transactions. One such heuristic—for a movie recommender system—is to find consumer X whose taste in movies is closest to the tastes of consumer Y, then recommend to consumer Y everything X liked but that Y has not seen; and

Model-based. These techniques use previous transactions to learn a model (usually using some

ONE OF THE MAIN CHALLENGES FOR PERSONALIZATION DEVELOPERS IS HOW TO ACHIEVE THE VIRTUOUS CYCLE AND NOT FALL INTO THE DEPERSONALIZATION TRAP.

machine-learning or statistical technique) that is then used for making recommendations. For example, based on the movies consumer X has seen, a probabilistic model is built to estimate the probability of how consumer X will react to each of the as-yet unseen movies.

These two classifications are orthogonal and give rise to six classes of matchmaking methods corresponding to six possible combinations of the classifications [1]. Although much research has sought to produce matchmaking methods, most current methods do not address all issues crucial to deploying personalization technologies in real-life applications. A detailed list of limitations of the current generation of recommender systems is included in [1]. Here, we focus on context awareness and multi-criteria recommendations.

Context awareness. Many recommendation technologies provide their offerings based solely on the consumer information (and information about the offering) while ignoring additional contextual information. However, in many applications the utility of an offering (such as a product) to a consumer may depend on the context in which it was offered, as established through market research. It may also depend on the people with whom the product will be consumed or shared, along with the circumstances under which such sharing will take place. In such situations, simply recommending items to consumers may be insufficient; the recommender system would have to weigh such additional contextual information as time, place, and the companions of the consumer when recommending products. For example, when recommending a vacation package, the system should consider the time of year, the companion(s) with whom the consumer plans to travel (such as children), the prevailing traveling conditions and restrictions, and other contextual information. Including context awareness improves the system's matchmaking abilities.

Multi-criteria recommendations. Most recommender systems deal with single-criterion ratings (such as those by consumers of movies and books), but in some applications (such as restaurant recommenders), multi-criteria ratings must be incorporated into the methods. For example, many restaurant guides (such as *Zagat's Guide)* provide three criteria—food, decor, service—because consumers may have separate sets of preferences for each. Therefore, in order to be effective, the matchmaking engine must provide the personalized offerings that match well across all criteria. Although multi-criteria ratings have not been examined in the recommender systems research literature, multi-criteria optimization problems have been studied in the operations research literature for decades, and some of it is potentially useful for designing and developing matchmaking technologies. Adding context awareness and multi-criteria matchmaking abilities to the current generation of personalization systems would improve their matchmaking and content-delivery methods.

Measuring Personalization

This part of the personalization process addresses personalization impact and how to adjust the various stages of personalization strategy. The most common are accuracy-based metrics for measuring personalization impact, reflecting how much the consumer liked or disliked a specific personalized offering in terms of, say, the accuracy and relevance of the recommendation [5, 9]. Because accuracy-based metrics are unable to capture the more complex and subtle aspects of personalization, attempts have been made to develop the more general aspects of personalization effectiveness by advocating comprehensive personalization metrics (such as consumer lifetime value, loyalty value, purchasing and consumption experience, and return on customer) [10, chapter 11]. However, they constitute only initial steps; much more work must be done to develop ways to measure personalization impact.

If we are dissatisfied with the results of the personalization process, we must identify the causes and adjust some of the methods associated with the previous five stages—from data collection to measuring personalization impact—of the process, based on the feedback loops in Figure 2. If the performance metrics indicate the personalization strategy is not performing well, we need to understand if the cause is poor data collection, inaccurate consumer profiles, poorly chosen techniques for matchmaking, or ineffective content delivery. Such problems must be fixed through a feedback mechanism. Alternatively, we may determine that the selected performance metrics measure factors that are irrelevant for the personalization application and need to be replaced with more relevant metric(s). We call this a feedback-integration problem, since it determines how to adjust different stages of the personalization process based on feedback from the performance measures.

The feedback-integration problem is recursive, that is, even if we are able to identify the underperforming stage(s) of the personalization process, we may still face similar challenges when deciding which specific adjustments are needed within each stage. For example, if we need to improve the data-collection phase, we would have to decide whether to collect more data, different data, or just use better data-preprocessing techniques.

Because feedback integration in the personalization process has not been studied extensively in the personalization literature, more research is needed to understand how to transform e-business measurements into specific adjustments to the various stages of the process.

INTEGRATED PERSONALIZATION

Here, we discuss how these stages might be integrated into a single coherent process. No existing system implements all six stages and feedback loops or integrates them into an overall personalization process. Most state-of-the-art personalization systems implement only the first four stages (up to delivery and presentation) with rudimentary measurements and without feedback loops. This is unfortunate, because developing good evaluation measures, sound methods for adjusting personalization strategies, and proper feedback loops constitute an urgent task of personalization. The ability to achieve the virtuous cycle of personalization (and avoid the trap of depersonalization) depends critically on these steps.

The virtuous cycle requires deployment of:

- The best technologies for each of the six stages of the personalization process; and
- Sound design principles for integrating the various stages of the personalization process.

The data-driven (or "forward") method is the most popular for designing such a process. In it, data is usually collected first (or is already collected). Consumer profiles built from the data are then used in the matchmaking algorithms, and personalization impacts are measured.

However, we advocate designing the process backward in accordance with the well-known business-management dictum: You cannot manage what you cannot measure. Designing the personalization process should start by specifying the measures used to determine the impact of the personalization process. The selected measure(s) should then determine what types of personalized offerings should be delivered to consumers. The profiling and matchmaking technologies needed to deliver personalized offerings are then determined, as are the types of information that should be stored in the profiles. Finally, the types of relevant data that need to be collected for building comprehensive consumer profiles are also determined. We call this approach "goal-driven" (as opposed to "data-driven"), because it starts with a predefined set of goals (or measures).

Because it starts with goals, the goal-oriented approach is more likely to realize the virtuous cycle of personalization than the data-driven approach and is therefore more likely to deliver greater value to the stakeholders. However, the goal-oriented approach has never been studied in the personalization literature, and our conjecture must still be validated as part of any

THE GOAL-ORIENTED APPROACH IS MORE LIKELY TO REALIZE THE VIRTUOUS CYCLE OF PERSONALIZATION BETTER THAN THE DATA-DRIVEN APPROACH AND THEREFORE IS MORE LIKELY TO DELIVER MORE VALUE TO THE STAKEHOLDERS.

future research. Moreover, other issues affecting the design of the personalization process, including trustworthiness and the potential intrusiveness of personalization technologies, must also be studied and incorporated into the process.

Trustworthiness. How much can we trust the results produced by personalization technologies? Untrustworthy results may follow personalized offerings biased by the ulterior motives of recommending parties trying to further their own or some third-party interests. In this regard, consumer- and market-centric approaches to personalization (see Figure 1) promise to provide more objective offerings to consumers than the provider-centric approach that might potentially focus more on providers than on consumers.

Providing an intuitive explanation to the consumer regarding why a particular personalized offering was made plays an important role in getting consumers to trust personalization technologies. For example, Amazon.com provides some explanation as to why certain recommendations are made, helping consumers decide which ones are truly trustworthy and which are not. However, some other personalization solutions constitute black boxes that provide no transparency of the personalization process to consumers.

Intrusiveness. A personalization system sending too many feedback requests to the consumer is intrusive and often irritating. Therefore, minimizing intrusiveness is usually desirable; methods toward this end are classified as privacy-related and feedback-related. Although consumers want personalized offerings, they may be unwilling to provide some of the personal information needed to improve the system's ability to make recommendations. Striking an optimal balance between the usefulness of recommendations and the amount of privacy the consumer needs to give up to receive quality personalized offerings is the main research issue in privacy-related intrusiveness.

Feedback-related intrusiveness involves consumers providing explicit feedback to personalization systems that may require significant consumer involvement in reacting to personalized offerings. Several nonintrusive feedback-determination methods have been proposed (such as in [8]), though they are often inaccurate and cannot fully replace explicit feedback from consumers. Determining how to minimize intrusiveness (at a data-collection stage) while maintaining personalization quality (at a measurement stage) must be addressed; any solution must integrate information across several stages of the personalization process, as outlined in Figure 2.

Conclusion

We've described the process-oriented approach to personalization, particularly the importance of vertical personalization research and the need for solid design principles for integrating all stages of the personalization process leading to the virtuous cycle of personalization. The three most important issues in developing these principles and the related personalization applications are:

- Develop good metrics to determine personalization impact;
- Study the feedback-integration problem and develop novel methods to address it; and
- Investigate the goal-driven design process in order to achieve better personalization solutions.

We hope the personalization community explores these issues further, thereby developing better personalization solutions. C

References

1. Adomavicius, G. and Tuzhilin, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering 17,* 6 (June 2005), 734–749.
2. Adomavicius, G. and Tuzhilin, A. e-Butler: An architecture of a customer-centric personalization system. *International Journal of Computational Intelligence and Applications 2,* 3 (Oct. 2002), 313–327.
3. Adomavicius, G. and Tuzhilin, A. Using data mining methods to build customer profiles. *IEEE Computer 34,* 2 (Feb. 2001), 74–82.
4. Balabanovic, M. and Shoham, Y. Fab: Content-based, collaborative recommendation. *Commun. ACM 40,* 3 (Mar. 1997), 66–72.
5. Breese, J., Heckerman, D., and Kadie, C. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence* (Madison, WI, July 24–27). Morgan Kaufman, San Francisco, 1998, 43–52.
6. Cortes, C., Fisher, K., Pregibon, D., Rogers, A., and Smith, F. Hancock: A language for extracting signatures from data streams. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Boston, Aug. 20–23). ACM Press, New York, 2000, 9–17.
7. Hand, D., Mannila, H., and Smyth, P. *Principles of Data Mining.* MIT Press, Cambridge, MA, 2001.
8. Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., and Riedl, J. Applying collaborative filtering to Usenet news. *Commun. ACM 40,* 3 (Mar. 1997), 77–87.
9. Pazzani, M. A framework for collaborative, content-based, and demographic filtering. *Artificial Intelligence Review 13,* 5-6 (Dec. 1999), 393–408.
10. Peppers, D. and Rogers, M. *Managing Customer Relationships: A Strategic Framework.* John Wiley & Sons, Inc., Hoboken, NJ, 2004.
11. Riecken, D., guest ed. Personalized views of personalization (special section). *Commun. ACM 43,* 8 (Aug. 2000).
12. Schafer, J., Konstan, J., and Riedl, J. E-commerce recommendation applications. *Data Mining and Knowledge Discovery 5,* 1–2 (Jan.–Apr. 2001), 115–153.

Gediminas Adomavicius (gedas@umn.edu) is an assistant professor in the Department of Information and Decision Sciences of the Carlson School of Management at the University of Minnesota, Minneapolis.

Alexander Tuzhilin (atuzhili@stern.nyu.edu) is an associate professor in the Information, Operation and Management Sciences Department of the Stern School of Business at New York University, New York.