Editor: **Daniel Zeng**, University of Arizona and Chinese Academy of Sciences, zengdaniel@gmail.com

Nonoccurring Behavior Analytics: A New Area

Longbing Cao, *University of Technology Sydney*
Philip S. Yu, *University of Illinois at Chicago*
Vipin Kumar, *University of Minnesota*

Behavior-related studies and applications, such as behavior analysis, data mining, machine learning, and behavioral science, have generally focused on behaviors that have occurred or will occur. Such behaviors are called *positive behaviors* (PBs) or *occurring behaviors* (OBs). Related work has focused on behavioral patterns, anomalies, impact, and dynamics. This constitutes the area of behavior analytics, which focuses on understanding, analyzing, learning, predicting, and managing past, present, and future behaviors. When behavior representation and modeling are also considered, we use the term *behavior informatics* or *behavior computing*[1] to describe the new perspective of modeling, reasoning about, verifying, analyzing, learning, and evaluating behaviors. This has emerged as an important and demanding area for comprehensively and deeply handling ubiquitous behaviors online, in business, government services, scientific activities, social activities, and economic and financial business.

Limited research has been conducted on analyzing, detecting, or predicting nonoccurring behaviors (NOBs), those which did not or will not occur. NOBs are also called *negative behaviors*, which are not straightforward, since they usually are hidden and difficult to understand, or one usually is not concerned with them. That NOBs are overlooked does not mean they are unimportant. For instance, if a patient misses an appointment with a specialist, and thus misses the opportunity to receive immediate and appropriate treatment for a health problem, the patient's health could worsen. Additionally, in many situations, failure to follow rules or policies could result in administrative or even legal obligations. Therefore, it is important to build a theoretical foundation for NOB study.

Unfortunately, few research outcomes of NOB study can be identified in the literature. Relevant work includes event analysis; negative association rule mining,[2] which identifies patterns comprising nonoccurring items; and negative sequential patterns,[3–8] which comprise sequential elements that do not appear in the business process. No systematic work has been conducted to understand, model, formalize, analyze, learn, detect, predict, intervene, and manage NOBs.

NOB is not a trivial problem. Some may argue that it is simple to treat an NOB as a special OB, and that all relevant techniques can then be used directly for NOB analytics. Unfortunately, this often does not work for reasons related to the different natures and complexities of occurring and nonoccurring behaviors. In this article, we outline the concept of NOBs and related complexities, draw a picture of NOB analytics, and present our view of NOB research directions and prospects.

What is NOB?

We briefly discuss the essence, intrinsic characteristics and complexities of NOBs, and the forms that NOBs can take, in order to understand the concept of NOB.

Intrinsic Characteristics

NOBs refer to those behaviors that should occur but do not for some reason. They are hidden, but are widely seen in behavioral applications in business, economics, health, cyberspace, social and mobile networks, and natural and human systems. Many businesses, services, applications, and systems involve NOBs, including healthcare and

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

medical services; government reporting systems, such as tax-related claims and social welfare; transport systems; safety and security management; and the changing physical world, including climate change.

NOBs are mixed with OBs and are widely seen in business, services, applications, and social activities, such as a missed medical appointment or the disappearance of significant climate phenomena in certain regions. Such NOBs often are more informative than OBs, because they are associated with important effects (such as the onset of heart problems if a patient misses an early appointment with a cardiologist, or threats to safety if proper airport security procedures are not followed).

NOBs are often important and can be related to significant living aspects and health, medical, economic, social, and cultural effects or impacts. NOBs can be more informative and irreplaceable than OBs. For instance, in transport, ignoring a stop sign and driving through an intersection could cause an accident; in social media, failing to protect private information could lead to identity theft. The effects or impacts could be significant and could even cause additional problems, change the way things develop, and affect future trends or status.

In summary, NOBs are complex because of their invisibility and their embedded intrinsic characteristics in many aspects,[4,9] including

- unclear working mechanisms about how a behavior or behavior sequence did not occur;
- hidden reasons for why a behavior or behavior sequence did not occur;
- many coupling relationships embedded between NOBs and OBs;[3]
- various occurrence possibilities and combinations of other behaviors that could replace those NOBs or occur subsequent to those missing behaviors; and
- the impact or effect that could result from an NOB or a NOB sequence.

Such NOB forms bring about significant challenges to existing computing and behavioral studies.

Forms

The nonoccurrence of behaviors can take various forms when the nonoccurrence is considered together with OBs. We illustrate five abstract NOB forms, in which $\neg A$ means that A does not occur, and AB means two behavior elements A and B.

1. In $<\neg AB>$, one NOB or NOB sequence A is followed by an OB or OB sequence B.
2. In $<\neg(AB)C>$, two NOBs or NOB sequences AB form an element that is followed by an OB or OB sequence C.
3. In $(\neg AB)C>$, an NOB or NOB sequence A and an occurring item B form an element, which is followed by occurring C.
4. In $<\neg P \neg Q>$, an NOB or NOB pattern P (that is, a group of negative behaviors that may follow a pattern P such as one of the scenarios in 2 or 3) is followed by another NOB or NOB sequence Q.
5. In $<P \rightarrow \Psi>$, an NOB or NOB pattern P (one of the scenarios in 1 through 4) causes impact Ψ.

Such NOB forms bring about significant challenges to existing computing and behavioral studies.

Challenges

The intrinsic complexities and differences between NOB and OB bring about significant computing challenges that existing OB techniques cannot address. Here, we discuss whether we can treat an NOB as a special OB so that we can use existing OB techniques for NOB analytics. We also look at the fundamental computing challenges.

Can An NOB Be Treated As A Special OB?

In computing and behavioral science, researchers have made major efforts to analyze OBs. Very limited outcomes can be found in the literature for understanding NOBs and their patterns, sequential dynamics, impact, and utility. To the best of our knowledge, no work addresses such issues as the clustering and classification of NOB.

This indicates that significant gaps exist between the widespread NOB-related facts and needs and the respective theories and tools available in the knowledge base. Addressing these complexities significantly challenges the related work, including event analysis; both positive and negative sequence analysis;[10] behavior analysis;[11] data mining and machine learning, such as anomaly detection;[12] and behavioral finance.[13]

We must address why we cannot simply treat an NOB as a special OB and then directly apply OB analytics to NOB. Unfortunately, the issue is more complex than that, and there are many profound reasons, driven by the intrinsic differences and characteristics inbuilt in NOB, why this is so:

- An NOB element is fundamentally different from an OB in terms of characteristics.
- An NOB's distribution can differ from that of an OB.
- An NOB's impact can differ from that of an OB.
- An NOB's candidacy set has different characteristics and combinations than that of OB.
- NOB combinations often differ from OB combinations.

Therefore, it is necessary to develop effective theories and tools to analyze NOB accordingly.

NOB Computing Challenges

In addition to the NOB characteristics discussed and the different forms NOB can take, several fundamental computing challenges exist in modeling and analyzing NOB from the pattern mining perspective.

One such challenge is problem complexity. As we mentioned earlier, an NOB's hidden nature makes it complicated to define the problem of NOB analytics, particularly NOB formatting and negative containment.[14] This has been indicated in the limited research on negative sequence pattern (NSP) analysis,[2,3] where different and conflicting concepts have been defined. Because NOB is even more challenging, it is important to properly formalize the NOB analytics problem by considering intrinsic NOB characteristics and how they differ from OB.

Another challenge is the large NOB candidate search space. The large number and combinations of NOB candidates could result in a huge search space. The existing NSP approaches generate k-size negative sequential candidates (NSCs) by conducting a joining operation on $(k - 1)$-size NSP, which results in a huge number of NSCs.[4] For example, the NSC space of a k-size behavior sequence with 10 distinct items could be 20^k, whereas their positive sequential candidates are just 10^k. Furthermore, because NSC does not satisfy the Apriori principle, which holds for OB sequences, it is very difficult to prune the large proportion of meaningless NOB candidates.

Finally, high computational complexity is a challenge. Existing NSP mining methods calculate the support of NSC by additionally scanning the database after identifying positive sequential patterns, leading to low efficiency. There could be many possible combinations of replacement behaviors of an NOB. When the length of behavior sequences is very long and the number of behavioral items is big, NOB analysis could be very time consuming.

In addition to these pattern-mining-related NOB challenges, no efforts have been made to explore what challenges might be found in the clustering and classification of NOB-related behaviors or group and community NOB behaviors that are mixed with or hidden in extremely imbalanced OBs.

NOB Research Directions

We discuss possible research directions and opportunities to address the NOB-related characteristics and challenges, as well as the difference between NOBs and OBs from the perspective of modeling, pattern mining, clustering, classification, group, and community NOB analysis, impact, and evaluation. We also illustrate some of our related work in discussing these aspects.

Representation

NOBs are very different from widely seen behaviors because they are hidden and usually are complicated. The major challenge is how to explore their representation and verification

We must define an abstract NOB model to capture major characteristics and elements embedded with NOBs, such as actors, roles, forms, relationships and interactions, and impacts. We must consider the connection of these elements with OBs within certain contexts, in which we also consider environmental factors and impact. A key problem is how to capture the difference between NOBs and OBs.

We should develop both visual and formal modeling tools. A visual model symbolizes the main elements and properties of NOB elements, and a formal model formalizes an NOB to a property-based multidimensional vector by considering all of the NOB-related attributes. As a result, we can represent a behavioral application's individual and collective NOBs using these modeling building blocks.

We can build temporal logic-based ontological representation tools to collect the main properties of NOB, logic, statistical and probabilistic relationships between these properties, and various combinations of NOB and OB in the level of behavior items, elements (as basic behavior units, consisting of items), and sequences (formed by elements).

We must define negative containment and negation to specifically address NOBs' hidden nature, while taking into account the mixture of OB and contextual factors. The relationships and combinations within and between NOB and OB must be considered in defining the forms and formats of NOB negation and negative containment. This will create a generic understanding of fundamental problems: what forms NOB observations are in, and whether a NOB is contained in behavior data.

When the NOB representation and modeling tools are available, we must develop appropriate verification tools to verify and check the constraints, soundness, and consistency of the building blocks constructed. Constraints include ontology axioms and inferential couplings, and desired constraints will be formalized in linear temporal logic.

Pattern Mining

Pattern mining of NOB is a common need, especially in online, social, and organizational concerns. As we mentioned earlier, different forms of NOB patterns may exist. It is not easy to identify NOB patterns because we cannot simply treat an NOB item set or sequence as a special type of OB. Here, we discuss the main aspects that must be addressed in NOB pattern mining.

We use the e-NSP method to illustrate the main issues.[4] We recommend e-NSP because it is suitable for scalable behavior sequence analysis and it does not follow the approaches widely used in sequence analysis and pattern mining.

The e-NSP model is based on set theory, which identifies NOB through scanning only occurring behavioral pattern candidates, without additional database scanning, to substantially reduce the search space. A set-theory-based model will calculate the support of NOB patterns after obtaining the NOB positive partners (NOPs)[4] by analyzing the mathematical relationships between NOB candidates and their NOPs.

The problem statement is to explore how to systematically represent the NOB mining problem in terms of set theory and the process of set-based NOB mining. We must consider the intersection of two NOPs, {<*a*>} and {<*b*>}, from the set perspective. We can calculate the support of NOB sequence {<*abcdef*>} from the set perspective by considering possible NOP combinations {<*ace*>}, {<*abce*>},{<*acde*>}, and {<*acef*>}. Set-theory-based NOB mining first discovers NOP, then generates and converts NOB candidates to NOPs for calculating the support of NOB candidates.

Negative containment will be defined to determine whether a database sequence contains an NOB sequence. We can study different negative containment cases in terms of aspects that include NOB combination forms, NOB element versus sequence containment, left and right sides of sub- and super-NOB sequences, and conversion to maximal NOP sequences.

Negative conversion will develop rules, strategies, and theorems for converting whether a data sequence contains an NOB sequence to whether a data sequence does not contain other NOP sequences. We will do this from the perspective of set theory to look at the combination spaces and complement sets.

We will calculate the support of NOB partners with theorems and formulas defined for calculating the support of NOB items and sequences from known NOPs (that is, OB patterns) in terms of the set relations.

We will develop constraints and effective pruning strategies to filter and avoid irrelevant NOB candidates. Because the NOB combination space is usually huge, and not every NOB will appear in a specific application, we will develop strategies that involve domain knowledge and filter the least-frequent and low-impact (per contribution and utility) NOB, and we may need to explore the constraints on containment and conversion.

Theoretical analysis will provide the foundation for computational complexity analysis, considering the NOB-related factors, including the number of NOB candidates, the total number of comparison times required to calculate all union sets, and the time for processing data storage, and considering the factors related to data characteristics, including the average number of elements per NOB sequence, average number of items per element, average length of potentially maximal sequences, and number of sequences and items. We will define generic mathematical formulas to quantify and verify the complexity of our algorithms and baselines.

In addition, if we consider pattern relations, such as couplings,[9] it would be even more complicated and interesting to discover combined patterns[15] for NOB behaviors by considering ontological, temporal, inferential, party-based, and other couplings in behavioral and social systems.[14]

Clustering

Clustering NOB has not been attempted, to the best of our knowledge. It could be very complicated, and might require us to explore the following issues:

- *Problem statement*. We must define the problem of NOB clustering by considering NOBs' intrinsic nature and associated hidden attributes. We must define what an NOB is and how to judge whether two NOBs are similar, before we can apply the classic clustering process.
- *Feature analysis*. If we treat an NOB item as an object, we can explore two typical aspects—the hidden nature, and the potentially complicated coupling relationships.[9] The former concerns which properties are associated with this NOB item, and how to construct, select, or mine features that capture the hidden nature and that are discriminative to NOB learning. The latter involves the advancement opportunities of analyzing factors and feature selection by involving couplings and non-IIDness (that is, scenarios that are not independent and identically distributed)[14] across NOB features.
- *Similarity definition*. How can we define whether two NOB items are similar or dissimilar? Another issue is how to determine whether an NOB item is similar to an OB one. These issues are extended to higher levels of combinations of NOB and OB items, such as how to determine whether one NOB sequence is

similar to another, and whether an NOB sequence is similar to an OB sequence.

- *Coupling learning*. We must consider various couplings that exist in real-life scenarios within and between behavioral properties, behavioral items, sequences, and their business impact.[9] Taking such couplings into account in clustering objective functions and models is a new area related to non-IIDness learning.[14] The preliminary work on coupled clustering[16] and coupled ensemble clustering[17] gives some examples of how to incorporate coupling relationships into clustering.
- *Evaluation*. When we consider the complexities and characteristics of the nonoccurrence of behaviors in respect to clustering, the evaluation of models and outcomes essentially becomes very challenging. It could involve questioning what kind of data showing NOB and their characteristics is available, what kind of benchmark data is available for evaluation, and what kind of metrics are suitable for verifying the characteristics of NOB and the performance of models.

These issues indicate great innovation opportunities to develop new methods and algorithms to cluster NOBs.

Classification

Similar to NOB clustering, there is no work that reports on NOB classification. Accordingly, the issues surrounding NOB classification might well be similar to those surrounding NOB clustering.

Labeling NOB could be challenging because the NOB did not take place, so it would be difficult to evaluate its impact. It might not be as straightforward as simply reverting the NOB to OB from the business impact set. For instance, if a conference chair failed to appear at a conference session he or she was supposed to chair, the impact would be high if nobody else took on that role, but the impact could also be moderated if the chair informed the organizer and arranged for a replacement.

The widely explored issues in classic classification could also apply here; for instance, multilabels, imbalanced data, and limited label information, which are related to fully supervised, partially supervised, or fully unsupervised classification. Studies on issues including unsupervised classification, partial label-based classification, and imbalanced data classification could be interesting when applied to behavioral data.

A major challenge could be the definition of class labels for NOB, which may be different from OB with measurable outcomes. For this, experimental analysis, simulation, and domain-knowledge-based labeling could be useful. For some cases, one could create labeling for NOB by considering certain constraints.

When we consider the couplings between and within NOB properties, objects, and sequences, as well as between NOBs and labels, we will need to determine how such couplings can be modeled in the similarity metrics and classification objective functions. We might be able to extend some relevant work about coupled KNN.[18]

Compared with OBs, NOBs can be rarer, more sparse, and imbalanced. This means we must classify infrequent NOBs mixed with OBs for a specific purpose. When the combinations of NOBs are uncertain, we require more research to classify such NOBs.

Group Analysis

Research[19] shows that the respective behaviors conducted by an actor or a group are often coupled with each other in terms of various relationships[9,14] that are crucial for behavior formation, dynamics, and impact. Existing behavior analysis models and algorithms usually ignore such couplings and treat behaviors independently. We must explore the couplings embedded in NOB properties and behaviors, including the following aspects:

- *Definition of group NOB*. We identify both individual and cohort levels of NOBs with respective specified properties. At the group level, we specify common and exceptional NOBs shared by all or a proportion of group members.
- *Formalization of coupled group NOB*. Arriving at a formal understanding of coupled NOBs involves the exploration of forms and types of NOB couplings in terms of structural, dependent, semantic, inferential, and impact-related empirical aspects.[20] We can formalize intra- and inter-NOB couplings and their integration by treating an NOB space as a high-dimensional sequential matrix (or tensor)[19] and representing NOB properties via a vector.
- *Measuring relationships in coupled NOBs*. We must develop metrics to quantify the couplings in coupled NOB in terms of intra- and inter-NOB couplings, and integrative couplings to combine both intra- and inter-NOB couplings. Apart from the frequency and NOB-related factors, we can also consider behavior properties, actor properties, similarity between properties and their values, and impact on behavior outcomes. We must design similarity metrics to capture these couplings between and across different categorical and numeric aspects, for example, by extending the coupled similarity

metrics defined by Can Wang and his colleagues.[16]

- *Similarity-based coupled NOB analysis*. This can incorporate the coupled similarity into relation learning and clustering algorithms, such as the dynamic mixed-membership block model, to learn similar NOBs. This will address the intrinsic non-IIDness[14] in NOBs and OBs by involving both NOB and OB properties and couplings between and within NOBs and OBs.

These discussions show that NOB classification is not straightforward, which leads to new opportunities for innovative behavior analysis theories and tools.

Impact Modeling

Here, we examine the impact NOB will have on business or other aspects. This involves us quantifying NOB's consequences and analyzing what results might be delivered if certain NOBs exist under certain conditions. This requires the development of appropriate metrics for measuring NOB impact. In practice, we can quantify NOB impact as utility, profit, or cost benefit, or evaluate it in terms of other measurements acceptable to business.

Another important matter is to develop models and algorithms to identify impact-specific NOB and NOB patterns. Taking impact-oriented NOB pattern mining as an example, we discuss some of the relevant aspects, which include the following:

- Theoretical frameworks for measuring NOB—for instance, utility functions to be defined for NOB items, elements, sequences, and databases. Concepts such as element/sequence containment, length and size of utility sequences, and sequence matching will be formalized to form the statement and framework of impact-specific NOB analysis.
- Impact-specific NOB mining algorithms—for instance, for high-utility NOB mining, one option might be to extend the lexicographic quantitative sequence tree[21] to construct and organize utility-based NOB sequences, and to define strategies to append a new NOB item to form $(k + 1)$-NOB sequence concatenations.
- Pruning strategies are needed to efficiently select effective NOB patterns. As a utility-based NOB may be associated with multiple utilities in the NOB population, averaged utility and maximum utility-based NOB selection strategies may be developed. Pruning strategies need to be explored in terms of sequence-weighted utility, sequence-projected utility and sequence-reduced utility strategies.

When we consider impact in NOB analysis, we must discuss the difference between ways of identifying NOB and OB patterns. When impact is encoded into labels, NOB clustering and classification are then involved.

Measurement Metrics

Because NOBs occur in forms that are different from OBs, existing analytical methods and metrics are not directly usable. We must therefore develop metrics to evaluate NOB from both technical and business perspectives.

We must define mathematical and probabilistic metrics to measure the significance and probability of NOB items, elements, and sequences, and of NOB items, elements, and sequences mixed with OBs. We do not need to quantify the contribution of specific NOB elements causing particular NOB impact, such as fraud.

Similarity and dissimilarity metrics must involve the couplings[9] between and within NOB properties and their values, and within and between NOB items, and must integrate both intra- and inter-NOB similarity. Both categorical and numerical NOB properties will be considered in defining the coupled NOB similarity.

Impact-oriented metrics must measure the business value and impact of NOB patterns. Unlike high-utility sequence mining,[21] in which utilities of unit items are obtainable from business, NOB item and sequence utilities might not be obvious. The utility of conducting specific NOB items, elements, and sequences will be measured in terms of their specific contribution from particular business concern perspectives, such as profitability, loss, cost benefit, and value at risk. This must be implemented in alignment with the forms of negative containment and the negation and conversion of NOBs to their positive partners.

We must also consider computational complexity analysis for each analytical method. For instance, we must look at the size k of NOPs, calculate the comparison times N of NOB candidates (NBC), the number m of NBC converted from k NOPs, and the combinations of taking m NOB from all candidates.

Application Areas

As long as human actions, operations, and events exist in the data, NOB will be involved in either an explicit or implicit way. NOB analysis will be useful and will present another direction for deeply understanding business problems and solutions. Many application areas can benefit from such analysis—for instance, retail and online customer behavior, web usage and interaction, trading behavior in capital markets, and activities captured by surveillance systems.

NOB has applications in several areas:

- in airway business, to detect travelers who maliciously avoid necessary processes and security checks;
- in banking, to conceal the finance borrowed from a bank when a customer applies for a home loan from another bank;
- in capital markets, to hold shares when the market presents favorable trading opportunities;
- in consumer behavior, when groups boycott certain brands of products;
- in education, when students intentionally miss classes;
- in government service, when a social security benefit recipient fails to declare an address change to the government;
- in health and medical business, when a patient misses an appointment with a specialist;
- in insurance, when someone hides the actual cause of a car accident;
- in marketing, when a marketing campaign strategy excludes a certain group of customers;
- in networks, when users avoid accessing websites during peak traffic;
- in online business, when a company fails to release a shopping outlet's address;
- in security, when security processes are not followed;
- in social media, when people do not respond to others' comments;
- in transport, when a traffic light is faulty; and
- in tourism, when road signs are missing.

These discussions show that NOB analytics is highly demanding in real life, and the findings from analyzing NOBs could play some unique roles in deeply understanding the situation of NOBs and their impact on business.

Examples

As we discussed earlier, little work has been reported on NOB-related topics and applications. Some of our preliminary work could be extended. For example, some work on negative association rule mining,[2] which identifies associations of unordered itemsets, may be fundamentally reinvented to capture both occurring and nonoccurring items or sets, . Negative sequence analysis discovers sequential patterns comprising nonoccurring sequential items and elements.[3–8] Some examples include e-NSP,[4] which was designed for large-scale NOB analysis; Neg-GSP,[5] which identifies NOBs using the GSP algorithm; and GA-NSP,[3] which identifies negative sequences based on genetic algorithms.

In addition, impact-based negative sequence analysis identifies negative sequence analysis associated with certain business impacts, such as a high risk of incurring government debt in social security.[6–8,22] Furthermore, utility-based analysis[21] could be implemented in NOB-related analysis, and group and community behavior analysis,[19] which is used to identify pool manipulation in stock markets, can be explored in group NOB analysis.

In the increasingly connected and networked world, one often involves multiple businesses and correspondingly different sources of information. This indicates that NOBs are affiliated with behaviors, the behavior actor's (or actors') demographics, and the context and impact. This requires analysis of the corresponding NOBs in a multimodal, cross-media, cross-domain, or context-aware way. Significant challenges and opportunities emerge in understanding complex NOBs in such complicated scenarios. This leads to the development of a new area—NOB analytics. ■

References

1. L. Cao and P. Yu, eds., *Behavior Computing*, Springer, 2012.
2. Y. Li, A. Algarni, and N. Zhong, "Mining Positive and Negative Patterns for Relevance Feature Discovery," *Proc. 16th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining*, 2010, pp.753–762.
3. Z. Zheng et al., "An Efficient GA-Based Algorithm for Mining Negative Sequential Patterns," *Advances in Knowledge Discovery and Data Mining*, LNCS 6118, 2010, pp. 262–273.
4. X. Dong et al., "e-NSP: efficient negative sequential pattern mining based on identified positive patterns without database rescanning," *Proc. 20th ACM Int'l Conf. Information and Knowledge Management*, 2011, pp. 825–830.
5. Z. Zheng et al., "Negative-GSP: An Efficient Method for Mining Negative Sequential Patterns," *Proc. 8th Australasian Data Mining Conf.*, 2009, pp. 63–67.
6. Y. Zhao et al., "Debt Detection in Social Security by Sequence Classification Using Both Positive and Negative Patterns," *Machine Learning and Knowledge Discovery in Databases*, LNCS 5782, 2009, pp. 648–663.
7. Y. Zhao et al., "Mining Both Positive and Negative Impact-Oriented Sequential Rules From Transactional Data," *Advances in Knowledge Discovery and Data Mining*, LNCS 5476, 2009, pp. 656–663.
8. Y. Zhao et al., "Efficient Mining of Event-Oriented Negative Sequential Rules," *Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence and Intelligent Agent Technology*, 2008, pp. 336–342.
9. L. Cao, "Coupling Learning of Complex Interactions," *Information Processing and Management*, vol. 51, no. 2, 2015, pp. 167–186.
10. M.J. Zaki, "Spade: An Efficient Algorithm for Mining Frequent Sequences," *Machine Learning*, vol. 42, no. 1, 2001, pp. 31–60.
11. L. Cao, "In-Depth Behavior Understanding and Use," *Information Science*, vol. 180, no. 17, 2010, pp. 3067–3085.

12. V. Chandola, A. Banerjee, and V. Kumar, “Anomaly Detection for Discrete Sequences: A Survey,” *IEEE Trans. Knowledge and Data Eng.*, vol. 24, no. 5, 2012, pp. 823–839.
13. E. Cartwright, *Behavioral Economics*, Routledge, 2011.
14. L. Cao, “Non-IIDness Learning in Behavioral and Social Data,” *Computer J.*, vol. 57, no. 9, 2014, pp. 1358–1370.
15. L. Cao, “Combined Mining: Analyzing Object and Pattern Relations for Discovering and Constructing Complex Yet Actionable Patterns,” *Wiley Interdisciplinary Rev.: Data Mining and Knowledge Discovery*, vol. 3, no. 2, 2013, pp. 140–155.
16. C. Wang et al., “Coupled Nominal Similarity in Unsupervised Learning,” *Proc. 20th ACM Int’l Conf. Information and Knowledge Management*, 2011, pp. 973–978.
17. C. Wang, Z. She, and L. Cao, “Coupled Clustering Ensemble: Incorporating Coupling Relationships Both Between Base Clusterings and Objects,” *Proc. IEEE 29th Int’l Conf. Data Eng.*, 2013, pp. 374–385.
18. C. Liu, L. Cao, and P.S. Yu, “A Hybrid Coupled k-Nearest Neighbor Algorithm on Imbalance Data,” *Proc. Int’l Jt. Conf. Neural Networks*, 2014, pp. 2011–2018.
19. L. Cao, Y. Ou, and P.S. Yu, “Coupled Behavior Analysis with Applications,” *IEEE Trans. Knowledge and Data Eng.*, vol. 24, no. 8, 2012, pp. 1378–1392.
20. C. Wang and L. Cao, “Modeling and Analysis of Social Activity Process,” *Behavior Computing*, L. Cao and P.S. Yu, eds., Springer, 2012, pp. 21–35.
21. J. Yin, Z. Zheng, and L. Cao, “USpan: An Efficient Algorithm for Mining High Utility Sequential Patterns,” *Proc. 18th ACM SIGKDD Int’l Conf. Knowledge Discovery and Data Mining*, 2012, pp. 660–668.
22. L. Cao, Y. Zhao, and C. Zhang, “Mining Impact-Targeted Activity Patterns in Imbalanced Data,” *IEEE Trans. Knowledge and Data Eng.*, vol. 20, no. 8, 2008, pp. 1053–1066.

Longbing Cao is a professor in the Faculty of Engineering and Information Technology at the University of Technology Sydney. Contact him at longbing.cao@uts.edu.au.

Philip S. Yu is a UIC Distinguished Professor and Wexler Chair in Information Technology in the Department of Computer Science at the University of Illinois at Chicago. Contact him at psyu@uic.edu.

Vipin Kumar is a professor in the Computer Science and Engineering Department at the University of Minnesota. Contact him at kumar@cs.umn.edu.