Cutter Consortium
Access to the Experts

Cutter Business Technology JOURNAL

Vol. 30, No. 5, May 2017

Management, Innovation, Transformation

"The research articles in this issue advance the state of the art of fintech knowledge and provide detailed insights for financial services organizations that wish to gain an understanding of the ways technology can create value for them."

— Philip O'Reilly, Guest Editor

Beyond Fintech: New Frontiers

Opening Statement
by Philip O'Reilly 3

Wise Crowds, Safe Crowds: Balancing Diversity and Protection in Crowd Investing
by Jack Smith, Joseph Feller, Rob Gleasure, Philip O'Reilly, Jerry Cristoforo, and Shanping Li 6

Modernizing Legacy Systems with Machine Intelligence
by Albert Ma 13

A Robo-Advisor for China: Asset Allocation in Alpha UMa
by Jie Yang, Hanxi Ye, Yadan Wei, and Linqian Bao 21

Challenges and Success Factors in Global Test Outsourcing
by Xiaochun Zhu and Shanping Li 33

Refining Reconciliation: A Machine Learning Approach to the Financial Industry's Toughest Task
by Zhuo Li and Jianling Sun 39

Cutter Business Technology JOURNAL

As business models for creating value continue to shift, new business strategies are constantly emerging and digital innovation has become an ongoing imperative. The monthly *Cutter Business Technology Journal* delivers a comprehensive treatment of these strategies to help your organization address and capitalize on the opportunities of this digital age.

Founded in 1987 as *American Programmer* by Ed Yourdon, and until recently known as *Cutter IT Journal*, *Cutter Business Technology Journal* is unlike academic journals: Each monthly issue, led by an expert Guest Editor, includes five to seven substantial articles, case studies, research findings, and/or experience-based opinion pieces that provide innovative ideas and solutions to the challenges business technology professionals face right now — and prepares them for those they might face tomorrow. *Cutter Business Technology Journal* doesn't water down or delay its content with lengthy peer reviews. Written by internationally known thought leaders, academics, and practitioners — you can be certain you're getting the uncensored perspectives of global experts.

You'll benefit from strategic insight on how the latest movements in digital innovation and transformation, IoT, big data analytics and cloud, to name a few, are changing the business landscape for both new and established organizations and how cutting-edge approaches in technology leadership, enterprise agility, software engineering, and business architecture can help your organization optimize its performance and transition to these new business models.

As a subscriber, you'll also receive the *Cutter Business Technology Advisor* – a weekly bulletin featuring industry updates delivered straight to your inbox. Armed with expert insight, data, and advice, you'll be able to leverage the latest business management thinking to achieve your organization's goals.

No other journal brings together so many thought leaders or lets them speak so bluntly — bringing you frank, honest accounts of what works, what doesn't, and why. Subscribers have even referred to the *Journal* as a consultancy in print and likened each month's issue to the impassioned discussions they participate in at the end of a day at a conference!

Get the best in thought leadership and keep pace with the technologies and business models that will give you a competitive edge – subscribe to *Cutter Business Technology Journal* today!

Cutter Business Technology Journal®

Founding Editor: Ed Yourdon
Publisher: Karen Fine Coburn
Group Publisher: Christine Generali
Managing Editor: Karen Pasley
Production Editor: Linda Dias
Client Services: service@cutter.com

Cutter Business Technology Journal® is published 12 times a year by Cutter Information LLC, 37 Broadway, Suite 1, Arlington, MA 02474-5552, USA (Tel: +1 781 648 8700; Fax: +1 781 648 8707; Email: cbtjeditorial@cutter.com; Website: www.cutter.com; Twitter: @cuttertweets; Facebook: Cutter Consortium). ISSN: 2475-3718 (print); 2475-3742 (online).

Subscription rates are US $485 a year in North America, US $585 elsewhere, payable to Cutter Information LLC. Reprints, bulk purchases, past issues, and multiple subscription and site license rates are available on request.

Opening Statement

by Philip O'Reilly, Guest Editor

It gives me great pleasure to introduce the second of the fintech special issues of *Cutter Business Technology Journal (CBTJ)*. This special issue further showcases the R&D work undertaken in State Street Corporation's Advanced Technology Centres in University College Cork (UCC) and Zhejiang University (ZJU) and expands upon several of the concepts raised in last month's edition. Specifically, this issue focuses on key topics of interest for financial services organizations, namely equity crowdfunding, legacy systems migration, robo-advisors, test outsourcing, and refining the reconciliation process.

Financial services is a sector with significant information systems challenges. Even today, with the myriad of technology advancements that have taken place in the industry, legacy information systems remain very much *in situ*. Legacy systems are renowned for their inflexibility, which is hardly surprising considering most financial services organizations invested in information technology with a short-term view, not intending for it to last a significant period of time. Between poorly documented systems and the loss of original legacy information systems designers to retirement, organizations are finding both current maintenance and further development difficult. This has prompted many to look beyond their organizational boundaries for assistance with their information systems implementations.

Over the years, firms have pursued various strategies in designing, developing, testing, and implementing their information systems. Many have opted to outsource in various ways, from complete outsourcing of the IT function to engaging in strategic partnerships. Regardless of the model adopted, the upshot is that significant aspects of a financial services organization's IT are undertaken by third parties. While it is a common and mature practice, such outsourcing has not been without issues.[1] Much remains to be understood as to how financial services organizations can successfully outsource and partner with third parties.

A second aspect of the legacy systems problem that organizations must grapple with is systems migration. Indeed, systems migration issues pertain not only to legacy systems but also to the adoption of newer technologies such as blockchain. Regardless of the technology concerned, migration is a significant challenge for all financial services organizations, one that typically entails a time-consuming, costly, and difficult process. How can organizations handle systems migration effectively? Are there specific methodologies that can help? Can technology play a role?

Financial services is a sector with significant information systems challenges.

Another important area financial services organizations are focusing on at present is robo-advisors, and there is much discussion around their design, business models, and user adoption. One key question that remains to be answered is what is their operational value and the associated investment returns for users? Frankly, do these algorithms succeed in creating successful ROI margins in their selected portfolios? How do they perform in dynamic, volatile markets? Do they outperform existing (human) processes and methods?

Crowdfunding, the practice of funding a venture by raising small amounts of money from a large number of individuals, is typically performed via Internet-based intermediaries. Crowdfunding has received much attention in recent times, thanks to projects such as Oculus Rift, which received close to US $2.5 million in initial funding from investors on Kickstarter in 2012. Two years later, Oculus's owners sold the company to Facebook for approximately $2 billion. The Oculus Rift case became somewhat controversial because of its implications for investor protection. As *Guardian* technology reporter Alex Hern asks, "Were the backers, who paid almost $2.5m, engaging in a purchase (in which case the risk of failed projects seemed overly high), an act of philanthropy (which seems undercut by a billion-dollar sale), or an investment (but one in which they don't receive a share of the profits)?"[1] Such controversies have very much put regulation of crowdsourcing in the spotlight. Should crowdfunding intermediaries be regulated? What should be the nature of such regulation? This is a critical topic for discussion in today's multifaceted investor environment.

> **Alternatives to bank financing have drawn increased attention in recent years, as the financial crisis has restricted the amount of capital available through traditional means.**

In This Issue

Speaking of crowdfunding, our first article — by Jack Smith, Joseph Feller, Rob Gleasure, Philip O'Reilly, Jerry Cristoforo, and Shanping Li — focuses on equity-based crowdfunding, an alternative source of financing for organizations and a possible key to overcoming small and medium-sized enterprise liquidity issues. Alternatives to bank financing have drawn increased attention in recent years, as the financial crisis has restricted the amount of capital available through traditional means. One major advantage of crowdfunding is that it can be both a faster and cheaper source of financing. However, there are risks associated with the practice, including the potential for investors to be provided with inaccurate information. Educating and informing both investors and fund seekers is an important aspect of the crowdfunding process, and regulation will play an important part in this. Smith et al.'s research suggests that "equity crowdfunding regulations need to be specific and unique to this emerging investment mechanism; such platforms cannot be covered by existing investment regulations." Indeed they note that managing equity crowdfunding risk requires a specific set of regulations, making the important point that markets should not be overregulated from the start. Any regulations should ensure that the diversity of the crowd is maintained, a critical success factor in the context of equity crowdfunding. The authors note that "regulators and government departments seek to achieve a delicate balance between regulation for the safe participation of all involved and preservation of the unique investment environment that equity crowdfunding creates. Regulators and government departments are aware of how novel equity crowdfunding is and are cautious not to overregulate the market, which could kill it off completely."

As we observed last month, it seems that mainframe legacy systems will always be with us. This is a problem both in terms of increased operating expense and human capital — young people have little interest in learning procedural languages, and thus there are fewer and fewer people available to maintain these often mission-critical applications. While re-platforming COBOL in Linux or cloud containers might seem like the easiest fix, it doesn't address the problem of the dwindling talent supply. The only real solution is migrating mainframe systems to a modern technology stack, typically an expensive and time-consuming proposition. In our second article, Albert Ma tells us about BlueMorpho, a joint research project of InSigma Hengtian Software and Zhejiang University that uses machine intelligence and a new ontology-based methodology to "make the migration effort much more efficient and effective." While Ma acknowledges that BlueMorpho can't "automate the entire process flawlessly ... [w]hat it can do is to optimize cost savings and improve agility in migrating systems to a modern platform."

In last month's *CBTJ*, Jie Yang, Hanxi Ye, Yadan Wei, and Linqian Bao discussed robo-advisors, online platforms that use sophisticated algorithms to provide automated management of investment portfolios. In this issue, Yang et al. introduce Alpha UMa, the robo-advisor they created to help retail investors in China make sound investment decisions. They detail how Alpha UMa goes about selecting asset classes and making automated, threshold-driven trades, balancing the pursuit of high returns with the need to keep transaction costs low. While Nobel Prize winners and Harvard economists alike warn that the typical retail investor is unlikely to beat average market returns for very long, "Alpha UMa uses quantitative methods to generate views" that repeatedly yield above-market returns.

UPCOMING TOPICS

Roger Evernden
Leveraging Enterprise Architecture for Digital Disruption

Charalampos Patrikakis
Digital Transformation in the Industrial Sector

Don McIntyre
Agile Leadership

Indeed, the authors note, "Our simplified portfolio has an annualized return of more than 10%, which is a very good result in a turbulent market."

With so much money riding on the accuracy of algorithms, you can bet that the financial services industry is concerned about the quality of its software. High-quality systems require rigorous testing, which is the subject of our fourth contribution. In the article, Xiaochun Zhu and Shanping Li cite research that claims "product reliability will be better if independent test organizations conduct testing," which leads to their focus on test outsourcing. In looking at the subject, the authors found that "despite the growing interest in outsourcing in general and test outsourcing in particular, there has been no study that comprehensively investigates the types, processes, and challenges of test outsourcing." Fortunately, they've rectified this omission with their empirical study of test outsourcing at Insigma Technology, China's second-largest IT outsourcer. Through interviews and a quantitative survey, Zhu and Li identify the challenges and pinpoint the success factors in test outsourcing, making it easier for client companies to reap the benefits and avoid the pitfalls of this widespread practice.

In our final article, Zhou Li and Jianling Sun discuss the application of machine learning in the context of account reconciliation. The reconciliation process is critical to ensuring the completeness and accuracy of company accounts and likewise ensuring that organizations comply with various international accounting standards and principles (e.g., US GAAP). Li and Sun's research illustrates the efficiencies that can be realized through utilizing machine learning to reconcile accounting rules, resulting in reduced costs and less time spent on the onerous reconciliation process.

To sum up, in this issue we learn that:

- Regulation can contribute to equity crowdfunding success, but new regulations must be aligned with the principles of this novel form of financing.
- A combination of machine intelligence and an ontology-based methodology can greatly facilitate efforts to migrate legacy systems to modern platforms.
- Robo-advisors can play a vital role in asset selection and provide an above average ROI in a multi-asset portfolio invested in a rapidly evolving market.
- Test outsourcing can help deliver the high-quality software financial services firms depend on, and there are key success factors organizations need to consider to achieve this outcome.
- Machine learning can be applied to the accounting reconciliation process, thereby providing financial services companies with significant cost efficiencies and reducing the time they spend on this essential task.

The research articles in this special issue of *CBTJ* advance the state of the art of fintech knowledge and provide detailed insights for financial services organizations that wish to gain an understanding of the ways technology can create value for them. They also underscore the benefit of establishing partnerships between leading-edge financial services firms and universities to create new knowledge and lead the fintech revolution.

Endnotes

[1]Garland, Anna. "Five of the Biggest Outsourcing Failures." ITProPortal, 19 December 2015 (www.itproportal.com/2015/12/19/five-of-the-biggest-outsourcing-failures/).

[2]Hern, Alex. "Oculus Responds to Kickstarter Criticism with Free Headsets." *The Guardian*, 6 January 2016 (www.theguardian.com/technology/2016/jan/06/oculus-rift-kickstarter-free-headsets-virtual-reality).

Philip O'Reilly is a Senior Consultant with Cutter Consortium's Business Technology & Digital Transformation Strategies and Data Analytics & Digital Technologies practices. He is also Director of the State Street Advanced Technology Centre at University College Cork (UCC), the Academic Co-Director of Boole Business Labs, and a Senior Lecturer in Business Information Systems at UCC. Dr. O'Reilly received the Stafford Beer Medal in recognition of the most outstanding contribution to the philosophy, theory, and practice of information systems from the Operational Research (OR) Society at an awards ceremony that took place at the Royal Society in London. He has been invited to deliver keynotes and workshops by numerous multinational companies and at leading practitioner events, including the Banking & Payments Federation of Ireland National Conference*. Dr. O'Reilly has been a leading member of research teams that have been successful in securing in excess of €2 million in research funding in recent years. His research concentrates on digital business, with a specific focus on financial services. Dr. O'Reilly's work has been published in leading journals in the information systems field, including* European Journal of Information Systems, Journal of Strategic Information Systems, *and* Information Technology and People*. He can be reached at poreilly@cutter.com.*

Wise Crowds, Safe Crowds: Balancing Diversity and Protection in Crowd Investing

by Jack Smith, Joseph Feller, Rob Gleasure, Philip O'Reilly, Jerry Cristoforo, and Shanping Li

Small and medium-sized enterprises (SMEs) represent 99% of the businesses in Europe and are a major source of jobs and innovation.[1] SMEs, however, perpetually face a lack of sufficient funding.[2] Traditional financing mechanisms such as bank loans, venture capital, and angel investments are often not available to many SMEs.[3] Peer-to-peer financing in the form of crowdfunding is increasingly filling a funding gap for companies that are unable to obtain traditional financing or are too early in their lifecycle to attract angel investors and venture capitalists. Indeed, it was predicted the amount of funding received through crowdfunding would exceed venture capital in 2016.[4]

What Is Crowdfunding?

At its most basic level, crowdfunding is a form of capital financing that takes advantage of relatively small investments drawn from a large group of people, generally facilitated through online transactions on a crowdfunding platform. Crowdfunding is peer-to-peer funding behavior that bypasses conventional intermediaries by directly connecting funders and fund seekers.[5] In our work, we differentiate between four major forms of crowdfunding: rewards-based, donation-based, debt-based, and equity-based.[6]

This article focuses on equity-based crowdfunding, or *crowd investing*, which has become a promising instrument to help overcome SME liquidity issues, referred to as the early-stage equity gap.[7] The equity gap greatly reduces the success of smaller firms, and equity-based crowdfunding is a potential solution for reducing this gap because it removes barriers to equity.[8]

Crowdfunding Is Disruptive

Crowdfunding is potentially disruptive to traditional forms of financing. The current process for a company to go from startup to publicly traded company can be broken down into four stages:

1. **Founder investment.** All capital invested comes from the founders.
2. **Venture capitalist investment.** The company receives capital from angel investors and venture capitalists.
3. **Initial public offering (IPO).** The company is launched on a stock exchange.
4. **Public investment.** The company's shares are bought and sold on a daily basis on the stock exchange.

Importantly, each of these stages targets different levels of investor maturity, with IPOs in particular targeting quite sophisticated investors. Although accessible to very young companies, crowdfunding platforms behave in many ways like an IPO, as "crowdfunding sites are beginning to act more like stock exchanges in the services they offer their customers."[9] This is the heart of crowdfunding's disruptive nature — it skips the traditional steps for raising capital and provides financing directly through a crowd that is made up of *both* sophisticated and naive investors.

The Wisdom of the Crowd

This diversity is a major factor in making crowdfunding effective. Crowdfunding draws heavily on the "wisdom of crowds," a concept popularized by James Surowiecki. The wisdom of crowds concept argues that a crowd of individuals with diverse knowledge is likely to make better decisions or predictions than experts working independently, providing three conditions are met:[10]

1. Opinion diversity
2. Crowd decentralization
3. Crowd independence

Diversity of a crowd refers to the members' differences in terms of demographic characteristics, cultural identities, ethnicity, training, and expertise.[11] The value in crowdfunding thus lies in harnessing not just the capital but also the wisdom of the crowd, and the key to maximizing the wisdom of the crowd is diversification. The reason why diverse groups perform better is rooted in the fact that they are more able to take alternatives into account.[12] Experiments have confirmed that teams of randomly selected diverse agents can outperform teams made up of the most intelligent agents.[13] There is also evidence that the crowd outperforms professional analysts in financial predictions; investors can achieve a greater return based on recommendations of the crowd rather than those of the analysts.[14]

How Do You Regulate a Crowd?

Regulations are being introduced globally to aid in the safe participation for all parties in equity crowdfunding. The level of regulation differs greatly from country to country, with some countries having strict regulations, others adopting a laissez-faire attitude toward regulation, and still others leaving crowdfunding entirely unregulated. However, some commonalities do exist between regulations that have been introduced by various countries, such as requiring equity crowdfunding platforms to be licensed with the appropriate authority and imposing caps on how much a company can raise through equity crowdfunding within a certain period. In terms of the fund seekers, the disclosure required of companies before they can launch a crowdfunding campaign varies widely from country to country. As for investors, there are varying regulations restricting who can invest in any particular campaign and how much money one individual can invest. In the US, for example, there are tiered investment thresholds aligned to an investor's net worth or annual income.

In this article, we report on research carried out at the University College Cork/State Street Advanced Technology Centre in Cork, Ireland. The challenge we address in our work is how to leverage and nurture the diversity of the crowd while still ensuring the crowd behaves in a safe, responsible, and informed manner. We explore these four key questions:

1. What is the value of equity crowdfunding, and what enables this value creation?
2. What are the major risks associated with equity crowdfunding?
3. How should the identified risks be managed?
4. Will regulation impact the diversity of the crowd?

As part of this research, we held interviews with employees of a national European regulatory body (ERB), a national European government department (EGD), a US regulatory body (USRB), an active investor in crowdfunding (INV), and an individual from an equity crowdfunding platform (PFRM). Please note that the data gathered represents individual observation and opinion and not the official position of the employing organizations.

Figure 1 illustrates the overall model of regulation and diversity in equity crowdfunding that is emerging from our research. We discuss each of the components in the model in more detail in the following sections of the article.

> **Experiments have confirmed that teams of randomly selected diverse agents can outperform teams made up of the most intelligent agents.**

The Value of Equity Crowdfunding

Our findings reveal that the key perceived value of equity crowdfunding consists in creating an alternative form of financing for startups and SMEs, as opposed to such traditional forms as bank loans, angel investments, and venture capital investments. Crucial to the view of equity crowdfunding as an alternative form of financing is the idea that crowdfunding unlocks capital and takes advantage of previously unavailable capital; both EGD and USRB stated that crowdfunding is "an initiative to unlock capital markets." By uncovering previously unavailable capital, both the funder and the seeker benefit. The seeker is able to raise capital quickly and easily, and funders who may not ordinarily make equity investments are able to do so and potentially receive financial returns.

Faster and Cheaper

A key motivation for companies to use this alternative form of financing is that it can be faster than traditional methods, as companies do not have to go through multiple checks and red tape before they can start raising capital. ERB noted that crowdfunding "can aid companies raising capital with reduced bureaucracy." To raise capital through an equity crowdfunding campaign, a company simply launches a campaign on a platform and investors can invest through that platform based on the information the company has provided there.

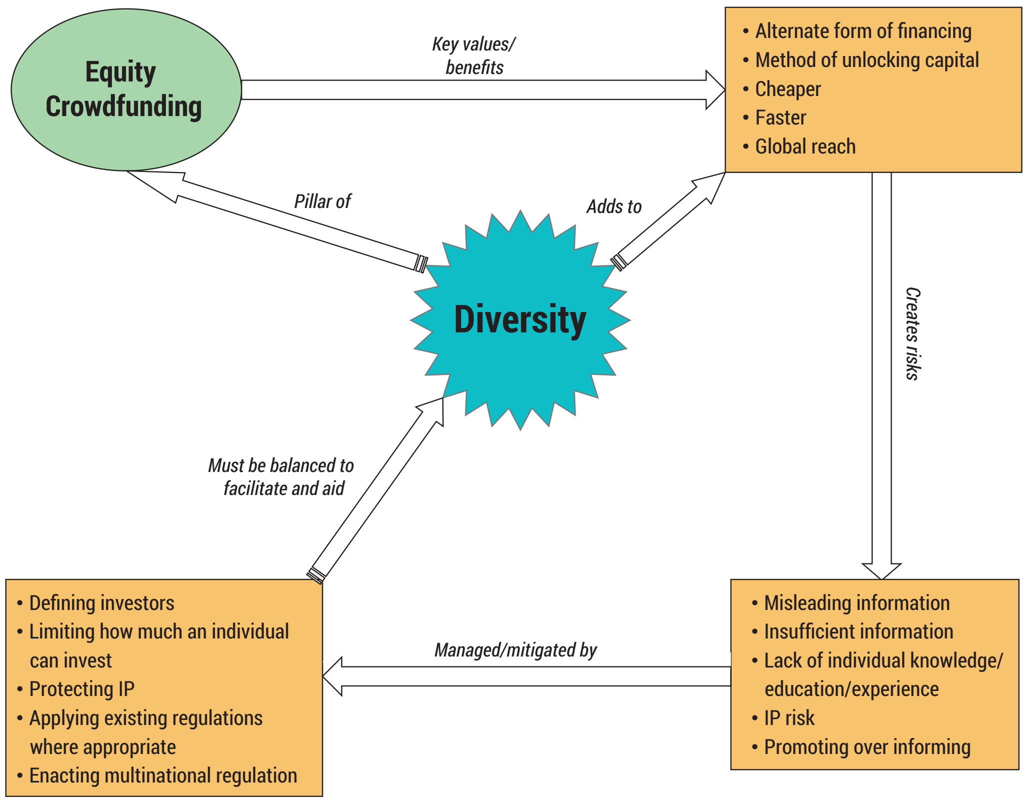


Figure 1 — Emerging model of regulation and diversity in equity crowdfunding.

In addition, equity crowdfunding is cheaper than traditional methods, as the company does not have to pay for all the process and documentation necessary for a bank loan or IPO. This idea was highlighted by EGD, who would like crowdfunding to "open up credit to people who wouldn't normally get it, and in a more administratively easy way than going into a bank." The cost of raising money is of crucial importance to startups and SMEs, as they are young companies with little cash reserves and cannot afford to pay for multiple audits and publication of a prospectus, among other costs.

More Than Money

Equity crowdfunding is much more inclusive, as it is open to a large crowd. More or less anyone can go onto a crowdfunding platform, view the campaigns that are being run, and decide if they want to invest or not. Furthermore, they can interact with companies, commenting on the campaigns and asking questions of them. This can provide the company with valuable feedback and help them gauge interest in their product or service. Utilizing an equity crowdfunding platform to raise money thus provides the company with another form of marketing through the platform. PFRM underscored this source of value, saying "a campaign serves as more than just a means of raising capital; the company can utilize it as an indicator for their product, almost a mini focus group."

Global Reach

As USRB put it, a "campaign can reach people all over the world." This global reach aids the unlocking of capital, as there is a wider audience of investors from which companies can get investment. This also amplifies the marketing benefits associated with equity crowdfunding. Once a campaign is live on a crowdfunding site, it can be viewed by people worldwide, thereby increasing the company's potential markets. The global reach of crowdfunding also means the crowd participating in a crowdfunding campaign is larger, which further adds to the diversity of the crowd.

The Risks of Equity Crowdfunding

The party most at risk in an equity crowdfunding campaign is the investor, and the main source of risk is inaccurate information provided through the crowdfunding platform. The information provided to the investor is integral in forming sound investment decisions, and misleading information increases the risk of both error and fraud. The threat of fraud was raised by INV, who observed "there is potential for fraud, as it is very easy to open a campaign or a platform and open it to the whole Web." The possibility of misleading information also arises, with USRB noting that "it's very easy to spoof a campaign."

Another major concern is insufficient information, as a fund seeker — while not intentionally misleading investors — might not provide enough information for an investor to make a sound decision as regards whether to invest. There may also be questions about the credibility of the source of this information. INV emphasized that one "cannot forecast without foundation"; this strengthens the observation that each piece of information provided must have a clear foundation and a credible source. On the other hand, it is important that the required information disclosure not be overly burdensome, as this could increase the costs of launching a crowdfunding campaign, thus discouraging a startup or SME from seeking funds through crowdfunding.

Educating and Informing or Advertising and Marketing?

There is often a tension between the hype a company uses to get noticed and the accuracy of the information provided. The problem with promotional videos and advertisements is that "promotional advertising material may not be independently verifiable," as EGD stated. This creates a similar challenge to the disclosure issue mentioned above. Investors require objective information, but should the standards and requirements for ensuring the integrity of promotional material prove too onerous, this may become an impediment for fund seekers.

Educating Investors

Because of the nature of equity crowdfunding and the diversity of the crowd, some participants will have considerably less investment experience than others. Therefore they are at a greater risk, as they may not completely understand the information that is being provided to them. As ERB commented, "consumers may not understand the risks associated with equity crowdfunding compared to other forms of financing." Investors need to be extremely aware of the high failure rate of startups and that they could be investing at a very early stage in a company's lifecycle. At such an early stage, the company is much more susceptible to bankruptcy, as it is still trying to establish itself in the market.

The aforementioned naive investor may run another risk in equity crowdfunding: lack of diversification in their investment portfolio, either in terms of equity and peer-to-peer lending forms of crowdfunding or other forms of investment. Any investor should have a range of investments in a diverse portfolio. INV noted that "lenders are at risk if they don't spread their loans across multiple borrowers in order to reduce the impact of a default."

The Regulation of Equity Crowdfunding

Our research suggests that equity crowdfunding regulations need to be specific and unique to this emerging investment mechanism; such platforms cannot be covered by existing investment regulations. Equity crowdfunding differs greatly from existing early stage financing due to the size of the crowd that can participate in a campaign, the diversity in investment sophistication and experience, the ease of participating in a crowdfunding campaign, and the fact that funds are invested through an online platform (meaning that trust plays a huge role in the decision to invest in a company through a crowdfunding campaign).

The first priority of the regulator is to protect the investor — in particular, the naive investor.

INV argued that, to begin with, the "equity crowdfunding regulations needed should be similar to existing regulations governing private equity and venture capital investing." However, because equity crowdfunding is still a relatively new form of raising capital, regulators have to be careful not to overregulate to the extent of squelching it. EGD supported this view, expressing the concern "that by regulating you might actually kill the industry before it gets started." It is necessary to regulate crowdfunding, as the risks are too great to leave the industry unregulated, but USRB recommended that regulators "start small." By starting small and building upon regulations slowly, this will allow regulators to see how the market reacts to regulations without strangling crowdfunding in its cradle.

Protecting Investors

The first priority of the regulator is to protect the investor — in particular, the naive investor. Due to the size of the crowd partaking in a crowdfunding campaign, there will be a wide range of people with varying degrees of knowledge of early stage capital investments (and all investments). INV proposed that investors be defined according to their experience and net worth: "people need to be defined as qualified or unqualified investors." This definition would determine how much each investor can invest in any given crowdfunding campaign over a certain period of time. Such a regulation would have at its core the goal of "restrict[ing] how much each can invest and how much they can buy, percentage-wise," thereby protecting any investor, particularly a naive investor, from investing too much in one campaign and losing a lot of money very quickly. USRB argued that "there is a fiduciary responsibility to investors to act responsibly for the customer" and suggested extending existing fiduciary responsibility to cover equity crowdfunding as well. Investment advisors need to be aware of and educated on the unique risks presented by equity crowdfunding before they can advise both qualified and unqualified investors on whether to participate in an equity crowdfunding campaign.

If a company chooses a platform and that platform crashes, fails, or disappears, there is a question as to what happens to the money already raised.

Protecting Fund Seekers

Regulation also needs to protect the fund seeker. The main risk for the seeker in a crowdfunding campaign is theft of intellectual property (IP). Because the seeker is posting their idea on a website that is available for all to see, it is crucial that the seeker "gets IP protected before putting the idea on a crowdfunding platform," EGD counseled.

A second risk to seekers is platform failure. If a company chooses a platform and that platform crashes, fails, or disappears, there is a question as to what happens to the money already raised. Both ERB and USRB highlighted the issue of platform regulation. Platforms need to be regulated and licensed before they can run crowdfunding campaigns; however, it is important that such regulations take the unique nature and requirements of crowdfunding platforms into account.

Facilitating Global Cooperation

Our findings indicate that crowdfunding regulations must be enacted at a multinational level because of the global reach of crowdfunding platforms. National regulations need to be compatible with other crowdfunding regulations across the world if cross-border investment is to take place effectively. This is, however, problematic for smaller countries where the crowdfunding market is small and there are no existing regulations, such as in Ireland. Indeed, ERB mentioned that while there were originally plans for a pan-European regulatory regime, that has changed as "there is no appetite for a pan-European regulatory regime anymore; it's now up to each individual country." In countries without any equity crowdfunding regulations at this time (Ireland included), "existing regulations are being implemented and adjusted without crowdfunding in mind." Central to this issue of multinational regulation is the differential in crowdfunding market size. For example, ERB continued, "In 2015 there [were] 3 million euros invested in crowdfunding in Ireland compared to 1.9 billion in the UK; less than 1% of [Irish] businesses have looked at crowdfunding as a form of raising capital." This point was further emphasized in discussions with PFRM and EGD, who commented: "Businesses look to P2P lending as one of the top two or three options when looking for funding in the UK, yet in Ireland it is down much further on the list."

The need for multinational regulation is further demonstrated by the present situation regarding platform regulation. Currently in Ireland, for example, crowdfunding platforms utilize self-regulation, doing their own credit checks and/or voluntarily aligning themselves with other countries' regulations. PFRM observed that "we have already adhered to FCA [UK Financial Conduct Authority] regulations and we hope that Irish regulations will be similar" and "some platforms are regulated at the moment in Ireland if they come under MiFID [Markets in Financial Instruments Directive] regulations." Our work shows that this issue needs to be addressed at a multinational level to aid and encourage cross-border investment without confusing investors by having various platforms operating under various regulations.

Evaluating Regulatory Impact on Crowd Diversity

As noted before, one of the major advantages of using equity crowdfunding as a method for raising capital is the global reach of campaigns. This global reach also adds to the diversity in the crowd that participates. By defining participants as qualified or unqualified investors and limiting how much each individual can invest depending on their experience, knowledge, and net worth, the size of the crowd is reduced, as is its diversity. By reducing diversity in the crowd, the wisdom of the crowd is also reduced. All interviewees were aware of this dilemma and conscious of the risk of overregulating.

Where Do We Go From Here?

As an alternative form of financing and a means of unlocking capital, equity crowdfunding has a clear benefit for younger and smaller companies. By enabling them to create jobs — and potentially grow into larger companies, with further job and wealth creation — such crowdfunding in turn benefits the economies of the countries in which those companies operate.

Regulators and government departments seek to achieve a delicate balance between regulation for the safe participation of all involved and preservation of the unique investment environment that equity crowdfunding creates. Regulators and government departments are aware of how novel equity crowdfunding is and are cautious not to overregulate the market, which could kill it off completely.

The key takeaway from our research to date is that diversity is key to forming a wise crowd, thus regulations need to make sure that the diversity of the crowd is not reduced too much. The challenge for regulators, platform builders, and other stakeholders is to ensure that risks throughout the system are managed without destroying the inclusivity that enables crowdfunding to operate effectively. In particular, regulation needs to allow for cross-border investment. By tailoring regulations to facilitate cross-border investment, the diversity among crowd participants can be safely sustained. Our research will continue to explore this complex balancing act, asking the question: "How can we create crowds that are both wise and safe?"

Acknowledgment

This article reports on research funded by State Street via the State Street/University College Cork Advanced Technology Centre in Cork, Ireland.

Endnotes

[1]De Buysere, Kristof, et al. "A Framework for European Crowdfunding," 2012 (http://eurocrowd.winball2.de/wp-content/blogs.dir/sites/85/2013/06/FRAMEWORK_EU_CROWDFUNDING.pdf).

[2]Bradford, C. Steven. "Crowdfunding and the Federal Securities Laws." *Columbia Business Law Review*, Vol. 2012, No. 1, 2012 (http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1118&context=lawfacpub).

[3]Hooghiemstra, Sebastiaan N., and Kristof de Buysere. "The Perfect Regulation of Crowdfunding: What Should the European Regulator Do?" Chap. 10 in *Crowdfunding in Europe: State of the Art in Theory and Practice*, edited by Dennis Brüntje and Oliver Gajda. Springer International, 2016 (https://link.springer.com/chapter/10.1007%2F978-3-319-18017-5_10).

[4]"2015CF Crowdfunding Industry Report." Massolution, 2015 (http://reports.crowdsourcing.org/index.php?route=product/product&path=20&product_id=54).

[5]Belleflamme, Paul, Thomas Lambert, and Armin Schwienbacher. "Crowdfunding: Tapping the Right Crowd." *Journal of Business Venturing*, Vol. 29, No. 5, September 2014 (www.sciencedirect.com/science/article/pii/S0883902613000694).

[6]Gleasure, Rob, and Joseph Feller. "Emerging Technologies and the Democratisation of Financial Services: A Metatriangulation of Crowdfunding Research." *Information and Organization*, Vol. 26, No. 4, December 2016 (www.sciencedirect.com/science/article/pii/S1471772716300148).

[7]Veugelers, Reinhilde. "Mind Europe's Early Stage-Stage Equity Gap." *Bruegel Policy Contribution*, No. 2011/18, December 2011 (http://bruegel.org/wp-content/uploads/imported/publications/111219_pc_rv_mind_europe_s_early-stage_equity_gap.pdf).

[8]Hagedorn, Anja, and Andreas Pinkwart. "The Financing Process of Equity-Based Crowdfunding: An Empirical Analysis." Chap. 5 in *Crowdfunding in Europe: State of the Art in Theory and Practice*, edited by Dennis Brüntje and Oliver Gajda. Springer International, 2016 (https://link.springer.com/chapter/10.1007%2F978-3-319-18017-5_5).

[9]Botero, Hector. "Why Equity Crowdfunding Platforms Are the Next Stock Exchanges." DealIndex, 3 September 2015 (https://dealindex.wordpress.com/category/equity-crowdfunding/).

[10]Surowiecki, James. *The Wisdom of Crowds*. Doubleday, 2004.

[11]Hong, Lu, and Scott E. Page. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences*, Vol. 101, No. 46, 16 November 2004 (www.pnas.org/content/101/46/16385.abstract).

[12]Nofer, Michael, and Oliver Hinz. "Are Crowds on the Internet Wiser Than Experts? The Case of a Stock Prediction Community." *Journal of Business Economics*, Vol. 84, No. 3, April 2014 (https://link.springer.com/article/10.1007/s11573-014-0720-x).

[13]Hong and Page (see 11).

[14]Nofer and Hinz (see 12).

Jack Smith is a researcher with the University College Cork (UCC)/State Street Advanced Technology Centre. Mr. Smith holds a BSc in business information systems from Cork University Business School, UCC, and is currently pursuing an MSc through research in fintech. Mr. Smith's research explores crowdfunding, its regulations, and how these regulations will impact the growth and value of this emerging market. He can be reached at jackdmsmith@gmail.com.

Joseph Feller is a Senior Consultant with Cutter Consortium's Data Analytics & Digital Technologies practice. He is Professor of Information Systems at the Cork University Business School at University College Cork (UCC), where he serves as Head of Business Information Systems. His research focuses on systems and organizational forms that leverage technological, legal, structural, and procedural openness to support mass collaboration and innovation. Such phenomena include open source software, content, hardware, and design; crowdsourcing and crowdfunding; open innovation and co-creation; open/citizen science; and so on. Dr. Feller has published 90+ peer-reviewed papers in journals such as Information Systems Research, Journal of the Association of Information Systems, Journal of Information Technology, European Journal of Information Systems, Journal of Strategic Information Systems, Information and Organization, *and* International Journal of Electronic Commerce, *as well as 130+ practitioner articles and reports. He has directed various international research projects funded by Irish, EU, and US bodies and by private industry and government partners. He is a frequent presenter and panelist at international conferences and has been invited to deliver keynotes, seminars, workshops, and policy briefings by numerous industry organizations, private firms, NGOs, and Irish and European government bodies. He can be reached at jfeller@cutter.com.*

Rob Gleasure is a Lecturer at University College Cork and Associate Lecturer at the Irish Management Institute. Dr. Gleasure's teaching includes a range of undergraduate, postgraduate, and executive education courses relating to digital business and online crowd behavior. His research focuses on design science, the psychology and neuropsychology of digitally mediated interaction, and crowdsourcing/crowdfunding. Dr. Gleasure has published work on these topics in many leading international journals, including Journal of Information Technology, Journal of the Association for Information Systems, Journal for Strategic Information Systems, Information Systems Journal, Information and Organization, *and* International Journal of Electronic Commerce. *Outside of research, Dr. Gleasure performs consultancy work in the areas of crowd engagement and interaction design. He can be reached at R.Gleasure@ucc.ie.*

Philip O'Reilly is a Senior Consultant with Cutter Consortium's Business Technology & Digital Transformation Strategies and Data Analytics & Digital Technologies practices. He is also Director of the State Street Advanced Technology Centre at University College Cork (UCC), the Academic Co-Director of Boole Business Labs, and a Senior Lecturer in Business Information Systems at UCC. Dr. O'Reilly received the Stafford Beer Medal in recognition of the most outstanding contribution to the philosophy, theory, and practice of information systems from the Operational Research (OR) Society at an awards ceremony that took place at the Royal Society in London. He has been invited to deliver keynotes and workshops by numerous multinational companies and at leading practitioner events, including the Banking & Payments Federation of Ireland National Conference. *Dr. O'Reilly has been a leading member of research teams that have been successful in securing in excess of €2 million in research funding in recent years. His research concentrates on digital business, with a specific focus on financial services. Dr. O'Reilly's work has been published in leading journals in the information systems field, including* European Journal of Information Systems, Journal of Strategic Information Systems, *and* Information Technology and People. *He can be reached at poreilly@cutter.com.*

Jerry Cristoforo is Executive VP and CTO for State Street Corp. In addition, Mr. Cristoforo is Chairman and General Manager of State Street Zhejiang Technology Company (SSTZ), where he pioneered State Street's partnership with Zhejiang University in Hangzhou, China, in 2000. He is a senior computer/securities industry executive with in-depth technology and management experience in advanced distributed systems architecture, product development, workflow reengineering, large-scale systems integration, and project management. He can be reached at jacristoforo@statestreet.com.

Shanping Li has been with Insigma Hengtian since 2004, where he serves as CIO. Dr. Li has been a professor of Zhejiang University since 1998. He works at the College of Computer Science and the College of Software Engineering. Dr. Li's main research areas are software engineering, distributed computing, and the Linux kernel. He received his master's degree in 1987 and his PhD in 1993, both in computer science, from Zhejiang University. He can be reached at shan@zju.edu.cn.

Modernizing Legacy Systems with Machine Intelligence

by Albert Ma

Overview

For the majority of organizations using mainframe systems, the most important fact to note is that they are spending over 70% of their development resources simply to maintain existing applications.[1] With the broad adoption of open systems, cloud technologies, and open source development, companies prefer to develop new applications on modern platforms. Despite the fact that mainframes are still highly reliable in transaction processing and have been widely adopted by companies that require intensive data crunching, it is difficult to attract young people to learn mainframe languages like COBOL, PL/I, RPG, and Progress. One of the greatest growing risks for these legacy systems is not the applications themselves, but rather finding people with the skills necessary to continue to develop, maintain, and operate them. After operating cost, human capital becomes the second most important reason to migrate mainframe applications to modern technology stacks.

Modernization efforts like re-platforming COBOL in Linux or cloud containers do not solve the problem of inadequate mainframe talents. Some companies have tried converting mainframe business analysts to take over the programming job. Such efforts are costly and can only be small in scale, as young people prefer to work with modern technologies and will plan their careers accordingly. Indeed, a Micro Focus survey of 119 universities in North America and around the world found that 73% of the universities polled do not even have COBOL programming as part of their curriculum.[2] Given this reality, applications must be translated to a modern language in order to be sustainable in this rapidly changing business world.

Yet making the migration to a modern platform is no easy task. The project itself must be cost-effective and not disrupt the current business. New systems must provide comparable business functions with enhanced value, and maintenance cost must also be lower than for the legacy system.

Before an application can be migrated, its business rules must first be extracted, documented, and then reengineered to cope with the new architectural design. This invariably requires significant up-front investment and is a lengthy process. Portions of this process can be automated, such as using static code analyzers to scan source code and plot the program structure, and then following through the program logic to determine what needs to be rewritten. One of the pitfalls in this approach, however, is that it always generates static information in batches. If a programmer modifies the source code, the whole system may need to be scanned, as it is difficult to know which parts of the program have been changed.

We believe that about 50%-60% of overall migration costs can be saved with the use of machine intelligence and a new methodology.

BlueMorpho is a joint research project between InSigma Hengtian Software Ltd. and Zhejiang University in Hangzhou, China, the goal of which is to empower the legacy system modernization effort and cloud migration. My BlueMorpho colleagues and I believe that about 50%-60% of overall migration costs can be saved with the use of machine intelligence and a new methodology. We have developed a special parser that can parse COBOL source code and generate a system ontology. With a dynamic tagging mechanism using semantic triples,[3] the system ontology can be expanded to cover different degrees of granularity. Machine intelligence is then applied to mine the source code with repetitive patterns associated with designated business rules. Combining prediction analytics with data flow analysis, a "code signature" can be discovered to determine what business functions a code snippet is about. The end results will be stored in the same ontology, thereby forming a dynamic knowledge repository. To enable machine learning, the code signature is persisted in a graphic database, which is accustomed to semantic search. All these tools make the migration effort much more efficient and effective.

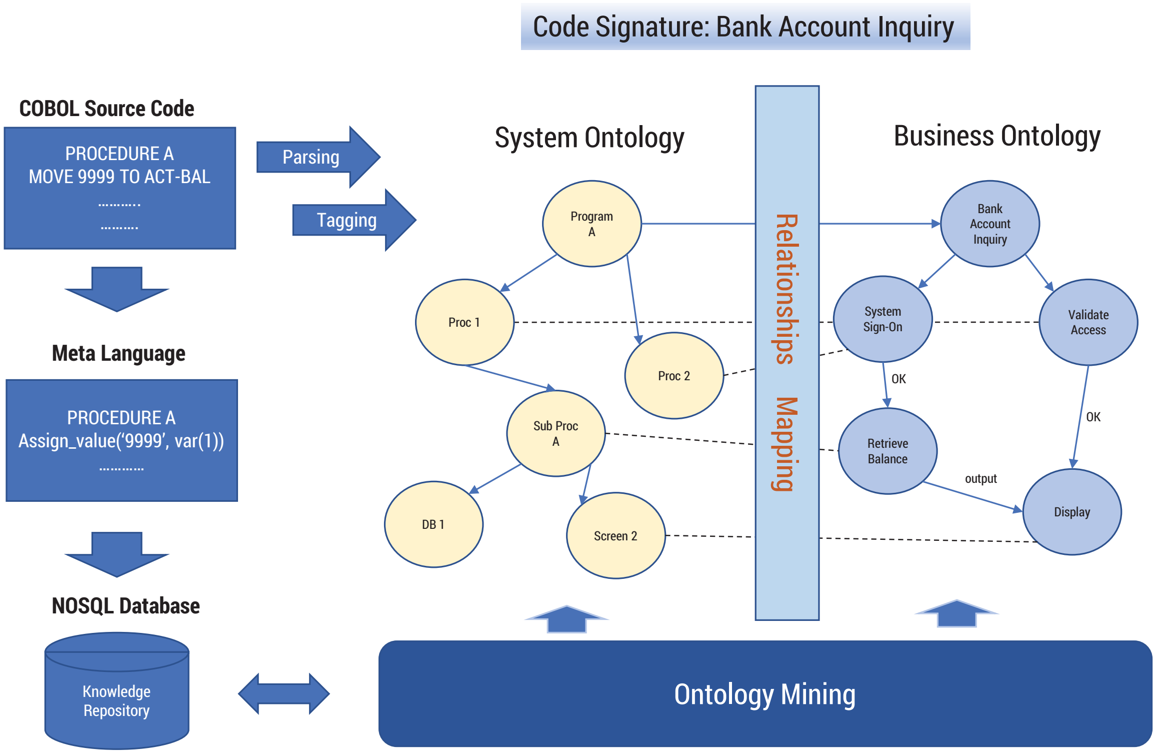


Figure 1 — Context diagram of ontology mining and code signature.

As shown in Figure 1, a system ontology serves as a graphical view of the program structure. Next to it is a business ontology that outlines the business functions and their relationships with individual system components.

Our past experience told us that the biggest challenge in a migration project is lack of relevant documentation. If we want to transform a program, we need to know what the source code means in business terms, but the purpose of a system ontology is to show the program structure without business knowledge. Where an ontological view of the system information coexists with a similar view of the business information, though, programmers can quickly identify all code dependencies from one single view. Furthermore, code with repetitive patterns can be translated to other languages in batches with some sort of domain-specific language by tracing through the ontology.

Legacy system migration always incurs major investment up front simply to understand the domain business problems and associated cost. With appropriate dynamic tagging, an ontological view allows programmers to uncover various details on the cost of migration, code complexity and dependencies, and so on. An ontology has unlimited applications when it is hooked up with continuous integration tools to support online queries for code reuse and generation.

Of course, a tool like BlueMorpho will not be able to solve all legacy system migration issues and automate the entire process flawlessly. What it can do is to optimize cost savings and improve agility in migrating systems to a modern platform. Machine intelligence will play a key role here. Even a 50%-60% cost reduction is a big incentive to kick off a modernization effort. The trend toward applying machine intelligence in software development is irreversible.

The Business Challenges of Legacy Modernization

Industry observers declared that mainframe computers would eventually be replaced when the PC was brought to market back in the 1980s. The same comment was heard with the rapid adoption of cloud computing over the past several years.[4] However, this does not seem to be happening — the mainframe is still deployed in many large corporations because of its reliability in heavy data crunching and transaction processing.[5] Ovum estimated that the current inventory of production COBOL running on mainframes is 150 to 200 billion lines of code.[6]

Yet the reliability and efficiency of mainframes do not come without cost. According to a 2007 Microsoft report, it "costs close to $50 per tpm-C (the TPC-C measure of business throughput in transactions per minute) for IBM System z9 configurations running z/OS compared to costs ranging from $0.77 to $3.84 per tpm-C for systems running Windows Server."[7] These figures indicate that there is a lot of room for cost reduction in migrating applications from mainframe to modern platforms. Time-sharing and pay-as-you-go models in cloud computing add the benefits of business agility and availability to the decision matrix.

While it is very clear that the mainframe is not going away anytime soon (the Micro Focus survey found that 71% of respondents believe businesses will continue to rely on COBOL-based applications for the next 10-plus years, while 24% put it at more than 20[8]), there are many legacy applications that have no major reasons to be run on a mainframe. It will be much more cost-effective to operate these non-mission-critical applications in modern technology stacks.

Although modernization can greatly reduce operating expenditures, many CIOs are still hesitant to pursue it. Following are the key factors that figure into their modernization planning:

- **Cost-effectiveness.** Many legacy systems are large and complex. It requires massive effort to understand, redesign, and reimplement such systems. However, the budget is always tight.
- **Business continuity.** Modernization activities must not adversely impact the company's productivity. Valuable time is often wasted as organizations wait for the migration process to complete. Switching to a modernized system should cause little or no service downtime, which is relatively difficult to achieve with existing methodologies and toolboxes.
- **Comparable system functions with enhanced value.** The modernized system must deliver the same functionality as the legacy system. There must be methods to ensure that a modernized system can also deliver added value to the business, either in cost savings or revenue generation.
- **Maintainability.** The implementation of the new system should adhere to a consistent design and coding convention, maximizing reuse and minimizing redundancy.

The Technical Challenges of Legacy Modernization

There are many ways to modernize legacy applications. The two most common approaches are to re-platform and to rewrite. For instance, COBOL programs can continue to run in a virtual machine without major code changes. This approach saves the mainframe operating expenses, but it does not address the talent issue.

Rewriting applications either entirely or partially is always the best long-term solution, but it poses many challenges:

- Many legacy systems either lack documentation or have obsolete documentation, which makes it very difficult to rewrite the application with the same business logic.
- Mainframe programs are mostly written in procedural languages like COBOL. Unlike programs developed in object-oriented languages, the sequential program structure and size of programs written in procedural languages always make it a challenge for programmers to extract the business logic behind them.

The goal of the BlueMorpho project is to use machine intelligence to mimic programmers rewriting a legacy system.

- It is a well-known problem in the mainframe world that programmers often cut and paste lines of code to avoid altering existing code. This adds to the difficulty of extracting business logic, as similar functions may exist in different places in the source code.
- Converting procedural languages like COBOL into OOP languages like Java or .Net requires a paradigm shift, even though today's static code analyzers are able to generate various kinds of structure diagrams.
- Today's conversion tools address only syntactical translation and not code semantics, making the output too complex — meaning that it executes much slower than expected.

Tackling the Problems with a New Approach

The goal of the BlueMorpho project is to use machine intelligence to mimic programmers rewriting a legacy system. While it was very clear from the start of the project that it would not achieve a 100% automated migration, any percentage of automation would still yield a lot of cost savings given the size of the mainframe legacy.

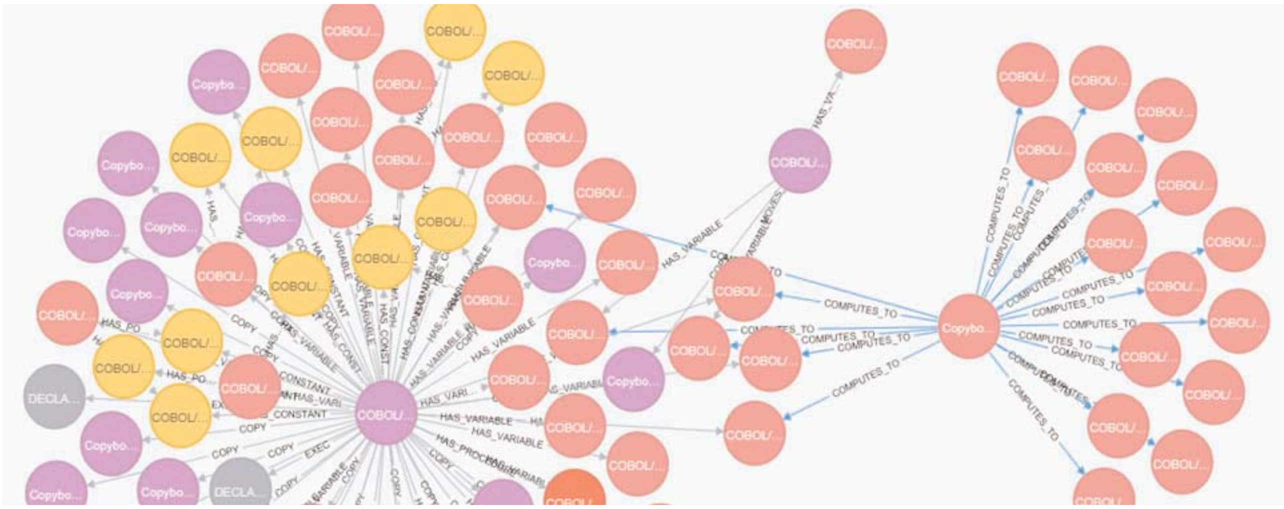


Figure 2 — System ontology view.

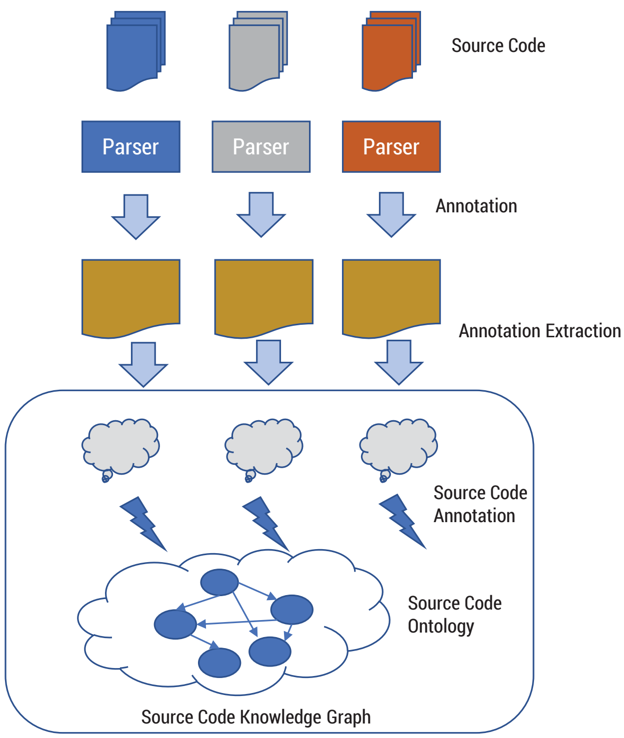


Figure 3 — Using ontology to document enterprise source code.

All manual modernization efforts must start with static code analysis followed by business rules extraction. Popular static code analyzers like Raincode or the IBM Rational suite generate static information in text and graphical formats only. It is not unusual to have to rescan code from time to time, as it may be changed during a lengthy project cycle.

BlueMorpho takes a very different approach. A custom-built COBOL parser generates a system ontology first. It stores the entire system architecture in the format of semantic triples persisted in a "graph database." The graph nodes (see Figure 2) represent mostly the procedure and function names, with various relationships linking them together. There are several benefits to doing this:

- Graphical presentation is always preferable to text.
- The view of nodes and relationships can be customized with simple SQL-like scripts, which allows programmers to locate the code segments relatively easier.
- Manual tags can always be added after the graph is generated, the implication being that the graph can keep evolving as an online documentation knowledge portal.
- The ontology serves as the core for subsequent machine learning of the source code.

In fact, an ontology can mostly be created automatically during source code parsing. If some lines of source code cannot be recognized, additional manual annotation using semantic triples will be required (see Figure 3). This allows BlueMorpho to generate a complete system ontology whenever machine intelligence is not applicable. An ontology is a schema-less design; in other words, every system may have its own design, making it difficult to create a universal standard. However, the process of parsing and mining source code should largely be applicable in different systems and programming languages.

Code Signature

BlueMorpho introduces a new concept, the code signature. A code signature is a way to identify a piece of business function using different algorithms. Compared to common object-oriented programming languages, COBOL is closer to natural language in syntax. Because COBOL is also procedural, programs will be relatively easier to read by a robot programmer.

A code signature can be as small as one line of code, such as "PERFORM CHECK-BALANCE," or code snippets with DO-WHILE loops and IF-THEN-ELSE statements. It can also be expanded to include the entire PROCEDURE or FUNCTION. It is very common for a high-level business function to include multiple tiers of business functions; thus, it is possible to have different code signatures uncovered in the same snippet of code. For example, "approval of payment" will include a series of validations on account balance, credit, authority, and counter-party information.

Uncovering the code signature from lines of source code offers many benefits:

- A code signature is business logic–related. It gives the business definition of code snippets and can augment the needs of additional inline documentation.
- A code signature can form a complex business ontology, which overlays the system ontology to exploit the lines of source code that perform specific business functions.
- Using a code signature as input, programmers will be able to discover redundant and duplicated lines of code with business semantics instead of just comparing the code syntax.
- Applying machine learning techniques to persisted code signatures enables better reuse of source code and automates the translation of COBOL code to other programming languages in business terms.

Machine Intelligence with Graph Database

Dynamic tagging of source code provides the foundation for generating system and business ontologies. Discovering code signatures from source code starts with a customized parser that reads the entire program structure and builds the nodes and relationships amongst them. In other words, this ontological view of the program structure provides the basic documentation and extension for graph mining. It offers a similar benefit to pre-processing a huge amount of unstructured text data and formatting it into a structural big data cluster for data mining.

The next step is to predict the business functions with structured procedure/function/variable names. Analysis of numerous programs in large companies has revealed that programmers quite commonly use a structured naming convention, which makes the prediction effort feasible on a certain scale. Paragraph names are usually a combination of alpha-numeric characters with a hyphen (or dash) as a separator. A typical example of a paragraph name — "CHK-ACT-BAL" — can be translated by a human.

The technical implementation of BlueMorpho uses the Hidden Markov Model (HMM), which is a machine learning algorithm widely used in word segmentation and signal identification (see Figure 4). The core components of HMM are the observable states layer and the hidden states layer. In this example, the observable

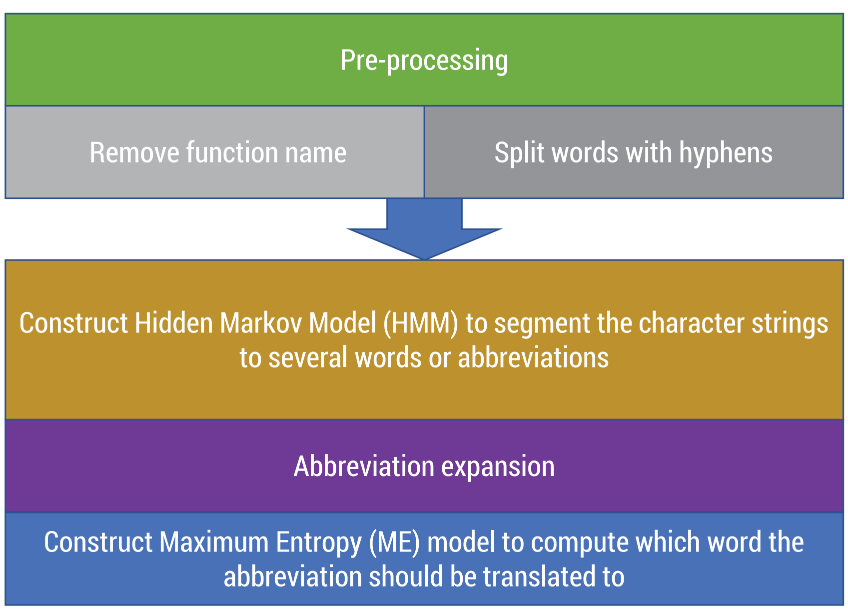


Figure 4 — Paragraph name prediction.

states are just the character strings waiting to be segmented, and the hidden states are the positions of the start, the inside, and the end of a word. Given an abbreviation like "TRANS," it might mean "TRANSMISSION," "TRANSACTION," or something else. We use the maximum entropy (ME) algorithm to make the decision. ME is used in most mainstream automatic translators to solve the word explanation problem. It uses context information to compute the probabilities when an abbreviation is translated to a word. In this case, if "TRANS" appears together with "ROLLBACK," then it is more likely to be translated into "TRANSACTION," whereas if "TRANS" appears together with "INFO," then "TRANSMISSION" would be a better choice, since information cannot be rolled back, but a rollback could indicate a transaction (in computer terms).

It is not uncommon to see bad naming conventions such as using mostly numeric characters in a paragraph name (e.g., A100023). This makes the above prediction effort almost useless, and a deep dive into the source code logic is required as a next step.

It is not the purpose of this article to dig into complex neural network algorithms. A simple illustration is that business logics are embedded in sequences of code execution with combinations of loop, conditional statement, and swapping data values. BlueMorpho adopts a semi-supervised training method to determine the business function of a code snippet and update the business ontology.

Assuming that source code has been sliced and diced with the generation of a system ontology, code snippets are extracted and tagged appropriately (see Figure 5). Tags will be used to update the GraphDB with node names and relationships. Neo4J was used in this project, but other semantic databases like Jena, Virtuoso, and Stardog can also be used. They are preferred for ontology persistence as they are schema-less, nonrelational databases. A GraphDB uses arbitrary object relations with no intrinsic importance. They are designed for storing huge semantic triples from the ground up with system performance in mind.

Source code being tagged will be converted into a customized metalanguage that is persisted in the database and linked to the ontology. Given the size of legacy applications, it is recommended to use a NoSQL database rather than a traditional relational database. This will be an iterative process until there are some foundations that can be used for the subsequent unsupervised training.

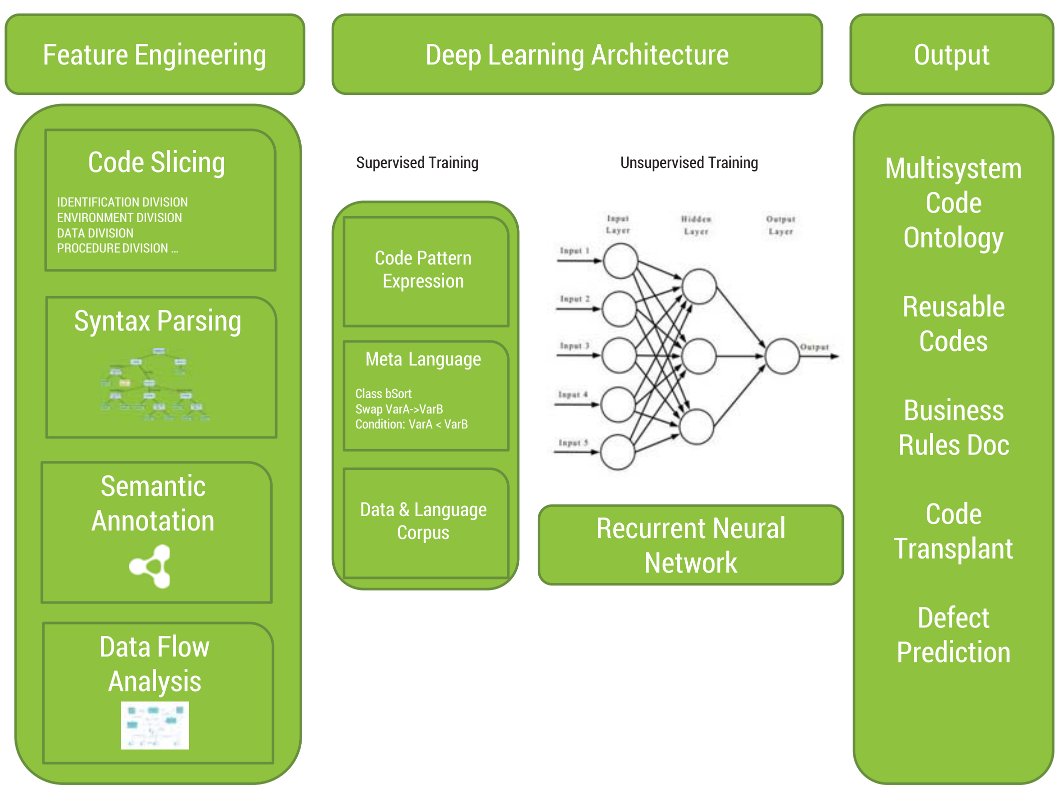


Figure 5 — Machine learning in BlueMorpho.

Determining whether a code snippet belongs to a business domain and/or carries out certain business functions depends on several factors (or cells) that are the hidden layer constructors of a neural network:

- Label (names of procedure, function, and variables)
- Loop (DO-WHILE and FOR-NEXT loops)
- Conditional statements (IF-THEN-ELSE)
- Data flow (track changes of variables in different programming modules)

Being able to translate code snippets into a metalanguage and persist them in a data store is important, as the neural network algorithms need to vectorize these factors before performing statistical calculation. A standard metalanguage to represent the code snippet syntax therefore must be created. This is different from storing the business logic of the code itself; rather, it is the syntax structure, which can be used to handle the classification task.

Common machine learning algorithms are claimed to be shallow. Most of the pattern recognition algorithms today require deep learning techniques with multiple hidden layers in a neural network. Neural networks have the common challenge of remembering long-term dependencies between multiple layers. In order to address this challenge, German computer scientists Sepp Hochreiter and Jürgen Schmidhuber invented the Long Short-Term Memory network (LSTM) in 1997.[9]

An LSTM is a special kind of recurrent neural network with the benefit of learning long-term dependencies (see Figure 6).[10] This capability is essential, as BlueMorpho has to look at different factors (loops, conditional statements, procedure/function names, data flows, etc.) to uncover the code signature. Each of the above constructors is a cell in the neural network layers. The gates are a way to optionally let information pass through. They are composed of a sigmoid neural net layer and a pointwise multiplication operation. Each of the constructors has its own calculation algorithms to control what information should be passed to the next layer until the end in order to uncover the code signature.

The Future As Summary

Artificial intelligence is rated by many CIO and CTOs as one of the top technologies that will disrupt our business models and day-to-day lives. Many machine learning and deep learning algorithms utilize raw data for calculation. BlueMorpho adopts a new approach by parsing the source codes into an ontology with graphic representation. This simplifies many steps in pre-processing data. Such an ontology empowers programmers to mine the source code that needs to be migrated and makes the business rules extraction process much easier.

My colleagues and I chose COBOL as the language for the initial research project because of its natural language-like syntax, the relatively larger size of programs written in it, and repetitive code patterns. These are the key elements in making algorithms work in most machine learning projects. However, there is no reason why similar algorithms will not work in other programming languages.

Automating the code transformation is always the end game. A business ontology unveils all the dependencies involved in accomplishing a particular business function. If similar business functions exist in the system ontology, writing simple scripts to "compare," "insert," and replace" code snippets will be relatively easy. A microservice architecture (MSA) offers another potential opportunity for developing a domain-specific language to automate this. Each system component in

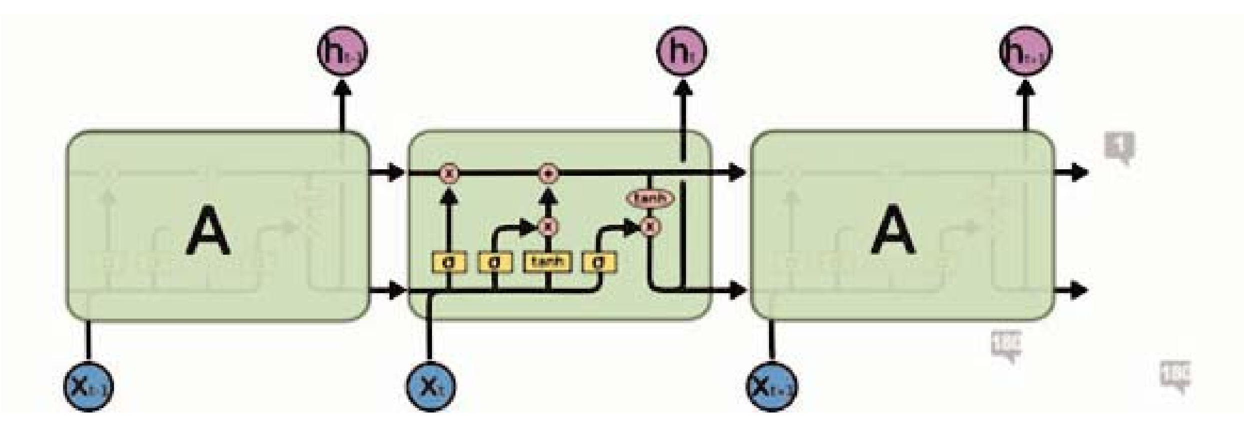


Figure 6 — Typical LSTM network. (Source: Olah.)

an MSA is intended to carry out one business function, and components are typically loosely coupled. Common machine learning classification algorithms may then play a key role in mapping a single business function to a microservice, yielding tremendous benefits in system scaling and maintainability.

Legacy system modernization is a costly and lengthy process. In the near future, there will be more and more machine intelligence–empowered tools to solve these problems, changing the way we develop software. The application of machine intelligence to legacy system modernization will know no limit but our own imagination.

Endnotes

[1]"The Business Value of Legacy Modernization." Microsoft Corporation, July 2007 (http://download.microsoft.com/download/E/9/7/E9734F87-C581-482A-AACA-2835DF48D40E/Business_Value_Legacy_Modernization.pdf).

[2]"Academia Needs More Support to Tackle the IT Skills Gap." Micro Focus, 7 March 2013 (www.microfocus.com/about/press-room/article/2013/academia-needs-more-support-to-tackle-the-it-skills-gap/).

[3]A semantic triple is a set of three entities (subject, predicate, object) that codifies the relationships between those entities. For example, "Albert knows Paul," "Paul married Elizabeth."

[4]Bice, Tom. "Back to the Future: Why the Cloud Won't Replace the Mainframe." *Wired*, April 2014 (www.wired.com/insights/2014/04/back-future-cloud-wont-replace-mainframe/).

[5]"3 Reasons Why Mainframes Still Matter." Advanced Software Products Group (ASPG) (http://aspg.com/reasons-why-mainframes-still-matter/#.WGAtEfArKUk).

[6]Barnett, Gary, and Mike Gilbert. "Legacy Renewal Strategies." Ovum, September 2006 (www.platformmodernization.org/microsoft/Lists/ResearchPapers/Attachments/18/Legacy_renewal.pdf).

[7]Microsoft (see 1).

[8]Micro Focus (see 2)

[9]Hochreiter, Sepp, and Jürgen Schmidhuber. "Long Short-Term Memory." *Journal of Neural Computation*, Vol. 9, No. 8, 15 November 1997 (http://dl.acm.org/citation.cfm?id=1246450).

[10]Olah, Christopher. "Understanding LSTM Networks." colah's blog, 27 August 2015 (http://colah.github.io/posts/2015-08-Understanding-LSTMs/).

Albert Ma is the Chief Innovation Officer of Insigma Hengtian Software, Ltd., with over 30 years of IT management and software development experience. In this role, Mr. Ma manages the company's R&D and innovation activities. He leads major research efforts in technical solutions in big data analytics, cloud computing, mobile, machine learning, and software engineering automation. Before joining Hengtian, Mr. Ma served for several years as CIO of the State Street Technology Center in Hangzhou, China. He was the core team member in setting up the IT operation locally and migrating strategic applications to the Hangzhou office. Mr. Ma graduated from the University of Queensland with a master's degree in technology management and an MBA from the University of Hull. He can be reached at amma@hengtianservices.com.

MORE IS MORE

A Robo-Advisor for China: Asset Allocation in Alpha UMa

by Jie Yang, Hanxi Ye, Yadan Wei, and Linqian Bao

There are 100 million retail investors in China. However, traditional financial advisors charge a lot, and not all investment advisors are trustworthy. For retail investors, it takes considerable time and knowledge to conduct portfolio management, and it is difficult for most retail investors to offset potential risks due to capital requirements in China. As a result, most of China's retail investors are not able to gain secured returns in the country's securities market.

Nobel Prize winner William Sharpe writes that "Properly measured, the average actively managed dollar must underperform the average passively managed dollar, net of costs."[1] Harvard economist John Y. Campbell also found the average investor will have an average before-costs performance equal to the market average. Therefore, the average investor will benefit more from reducing investment costs than from trying to beat the average.[2] Scholars saw similar results in China's fund market.[3] University of Chicago finance professor Eugene Fama's efficient-market hypothesis (EMH) also implies that it is impossible to "beat the market" consistently on a risk-adjusted basis since market prices should only react to new information or changes in discount rates (the latter may be predictable or unpredictable).[4] For all these reasons, investment advisors use globally diversified portfolios of index funds to invest passively for their clients.

With the development of technology, such investment is increasingly being enabled by robo-advisors, many of which have been established around the world. Robo-advisor companies like Betterment, Schwab Intelligent Portfolios, and Wealthfront provide financial advice or portfolio management online by employing algorithms such as modern portfolio theory to conduct portfolio management. Industry analysts predict that robo-advisors will have US $2 trillion in assets under management by 2020.[5]

Despite this growing trend, there was not a single truly algorithm-based robo-advisor in China. So we created our own robo-advisor, Alpha UMa, to help retail investors in China make reasonable investment decisions. In this article, we discuss how Alpha UMa goes about classifying asset classes and selecting exchange-traded funds (ETFs). We then introduce our portfolio optimization and rebalancing algorithms. Figure 1 shows the asset allocation process in Alpha UMa.

Classifying Asset Classes

As we noted in the last issue of *Cutter Business Technology Journal,* "In general, the main goal of asset classification is to construct a diversified portfolio" in order to reduce investor risk.[6] The first step in asset allocation is therefore to divide an investment portfolio into different asset classes. As of 12 August 2016, there were 4,669 funds in the China's Open-end Fund Market.

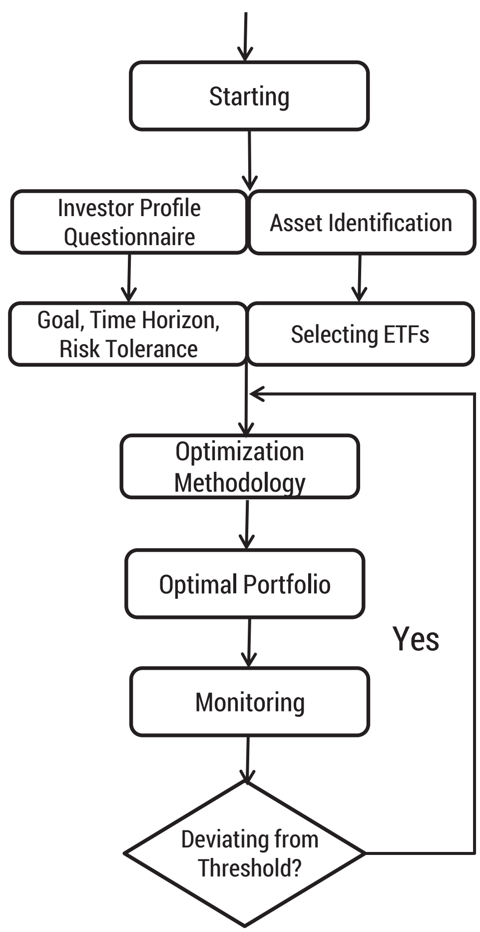


Figure 1 — Asset allocation process in Alpha UMa.

The China Securities Regulatory Commission (CSRC) divides the funds into six categories:

1. Stock fund
2. Hybrid fund
3. Bond fund
4. Money market fund (MMF)
5. Exchange-traded fund (ETF)
6. Qualified Domestic Institutional Investor (QDII) fund

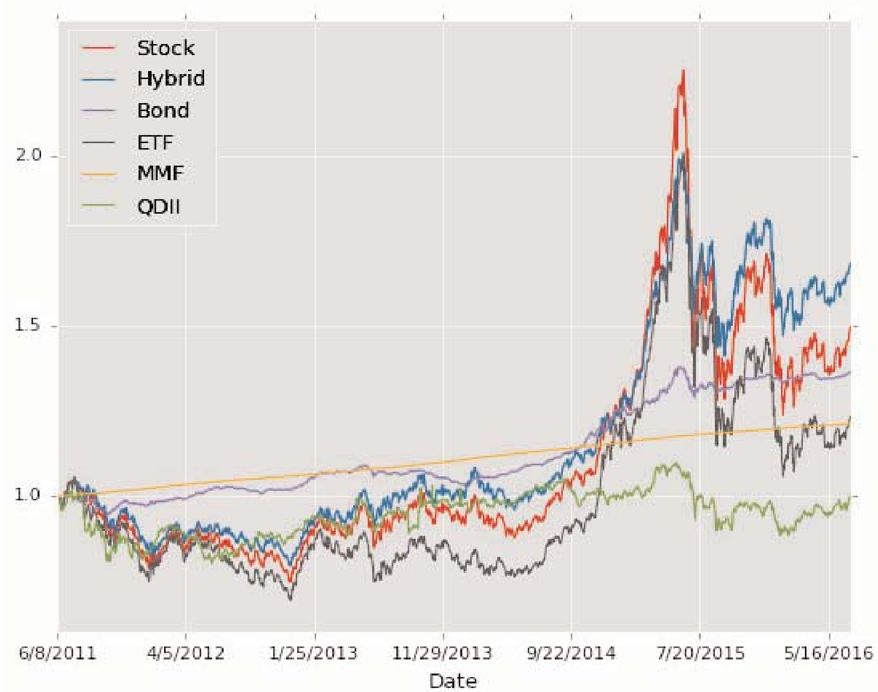
For each category, we can select one relevant index as the benchmark. Investors and fund managers can then compare the market performance of the specific fund with its corresponding benchmark return.

Figure 2 shows what the return in each asset category would be in 2016 for a dollar invested in June 2011. We can see from the figure that although the final return differs significantly across asset categories, some assets do show similar trends during some time periods. So we need to compute the correlation between any two benchmarks, which will yield the correlation matrix of the six indexes. As Table 1 shows, traditional assets — namely the stock fund, hybrid fund, and ETF — are highly correlated because stocks make up a large portion of the three indexes. MMF and QDII have very low correlations to other classes, thus making them a good tool for risk diversification.

Although QDII for Chinese investors is a good tool for investing overseas, correlations have been rising due to greater interconnectivity between global markets. As Figure 3 shows, events like the UK's leaving the European Union (or the 2008 US financial crisis, the European debt crisis, etc.) can hammer financial markets around the world.

To reduce the potential risk and construct a more diversified portfolio, we need to divide the QDII one step further. Based on quantitative and qualitative analysis, the stock market analysis firm Morningstar classifies Chinese QDII funds into the following: QDII Asia-Pacific ex-Japan Equity, QDII Greater China Equity, QDII Emerging Equity, QDII Global Equity, QDII Sector Equity, QDII US Equity, QDII Global Allocation, QDII Global Bond, QDII Commodities, QDII Others. To identify which asset classes to include in the portfolios, we consider these two factors:

1. Asset classes must be accessible through at least two liquid passive index funds.



We have the six benchmarks — namely H11021, H11022, H11022, H11024, H11025, and H11026 — to represent the six asset classes separately. Each asset class is represented by one line. The data resource is CSRC.

Figure 2 — Risk return profile of US $1 investment: 2011-2016.

2. Asset classes should be minimally correlated to achieve greater diversification benefits.

Thus, we can divide asset classes into China Stock, China Bond, China MMF, and China Commodity. Then, because we only have limited QDII funds, we divide QDII into four categories: QDII Greater China Equity, QDII Emerging Equity, QDII Developed Equity, and QDII Inflation-Protected Equity. This produces a total of eight asset classes in Alpha UMa.

Selecting Passive Index Funds

Mutual funds can make no claim of superiority over the market averages,[7] so most robo-advisors select open-ended tracking ETFs due to their low manager risk and low embedded costs. Although the ETFs have advantages on liquidity, we only have 140 ETFs in China as of today. What's more, there is a minimum investment requirement for ETFs that most retail investors cannot

	H11021	H11022	H11022	H11024	H11025	H11026
Stock	1.00	0.97	0.78	0.98	0.02	0.10
Hybrid	0.97	1.00	0.75	0.93	0.03	0.11
Bond	0.78	0.75	1.00	0.79	0.05	0.08
ETF	0.98	0.93	0.79	1.00	0.01	0.10
MMF	0.02	0.03	0.05	0.01	1.00	0.03
QDII	0.10	0.11	0.08	0.10	0.03	1.00

We use the data of the six benchmarks from 2011 to 2016 to calculate the correlation matrix. The data resource is CSRC.

Table 1 — Correlation matrix of asset classes.

Figure 3 — Global ETF performance after Brexit.

Asset Class	Selected Fund		
	Fund 1	Fund 2	Fund 3
China Stock	470068	512070	110020
China Bond	160720	161821	161119
China MMF	159002	159006	159004
China Commodity	000217	000930	000929
QDII Greater China Equity	000075	000071	110031
QDII Emerging Equity	160121	161714	
QDII Developed Equity	050025	040046	000834
		040047	
		040048	
QDII Inflation-Protected Equity	000885	000179	164815
		000180	

To make sure the funds' net asset value (NAV) is accurate and complete, the time period is 22 August 2013 to 1 August 2016. We have the fund codes in the columns labeled Fund 1, Fund 2, and Fund 3. Fund 1 indicates the passive index fund with the lowest cost. Fund 2 and Fund 3 are the alternative funds with the second and third lowest cost.

Table 2 — Asset allocation: passive index funds.

meet. Thus, we have decided to select only passive index funds in the portfolio. Since passive index funds closely match the performance of the index, we only need to consider the total cost of the funds in the specific asset class. The total cost formula is below:

$$\begin{aligned} Total\ Cost\ &=\ Tracking\ Error + Subscription\ Fee + \\ &Redemption\ Fee +\ Sales\ Charge +\ Expense\ Ratio \end{aligned}$$

Here tracking error indicates how closely a portfolio follows the index to which it is benchmarked. It is measured by the standard deviation of the difference between the portfolio and index returns:

$$Tracking\ Error\ =\ \sqrt{Var(r_p - r_b)} = \sqrt{E(r_p - r_b)^2 - [E(r_p - r_b)]^2}$$

where $r_p - r_b$ is the difference between the portfolio return and the benchmark return.

To calculate the cost, we used data from 1 July 2006 to 1 August 2016.[8] Then we selected passive index funds for the eight asset classes. In Table 2, Fund 1 indicates the passive index fund with the lowest cost. Fund 2 and Fund 3 are the alternative funds, with the second and third lowest cost. Sometimes, we have two to three funds in one grid, a situation that generally results from the difference in denominated currency. In practice, we can usually choose Fund 1 to represent the corresponding asset class. Then we can calculate the correlation matrix of the chosen funds.

As shown in Table 3, China Bond, China MMF, and QDII Developed Equity funds are highly correlated with each other. The strong correlations result from the low volatility of MMF and the steady increase of the US stock market. In general, the correlations among assets are relatively low, and these asset classes provide a level of diversification beyond that of traditional stocks, bonds, and cash.

	470068	160720	159002	000217	000075	160121	050025	270027
China Stock	1.00	0.60	0.67	-0.45	0.53	0.11	0.65	0.27
China Bond	0.60	1.00	0.97	-0.06	-0.17	-0.41	0.90	0.01
China MMF	0.67	0.97	1.00	-0.13	-0.04	-0.32	0.95	0.16
China Commodity	-0.45	-0.06	-0.13	1.00	-0.35	0.02	-0.08	-0.02
QDII Greater China	0.53	-0.17	-0.04	-0.35	1.00	0.85	0.10	0.65
QDII Emerging Equity	0.11	-0.41	-0.32	0.02	0.85	1.00	-0.14	0.63
QDII Developed Equity	0.65	0.90	0.95	-0.08	0.10	-0.14	1.00	0.37
QDII Inflation-Protected Equity	0.27	0.01	0.16	-0.02	0.65	0.63	0.37	1.00

Asset class is on the left ordinate axis, and the corresponding selected fund is on the upper horizontal axis. We use “Fund 1” to represent each asset class. The time period is 22 August 2013 to 1 August 2016.

Table 3 — Correlation matrixes of chosen funds.

Investor’s Profile Questionnaire

To match each client’s investment goal(s) and to determine the optimal mix of the portfolio, we ask clients risk questions and verify the consistency among the answers (see Appendix).

Investment Goal and Portfolio Sets

In Alpha UMa 1.0, we divide goals into three categories:

1. Build wealth
2. Build a rainy day fund for emergencies
3. Save for expenses (children’s education, health, travel, etc.)

For goal 1, all eight asset classes are included in the portfolio. For goal 2, only China Bond and China MMF are included. For goal 3, China Commodity is excluded from the portfolio.

Risk Spectrum

La Salle University researcher Michael J. Roszkowski and his coauthors show that combining risk-attitude (or behavior) related questions with objective questions will provide a more complete understanding of the investor.[9] In Alpha UMa, investors’ risk capacity is based on information about their financial situation, while investors’ willingness to accept risk is typically indicated by the level of volatility they’re comfortable with and other factors. Generally speaking, an investor’s risk capacity and risk willingness are independent of each other:

$$\text{Risk Tolerance} = \text{Willingness to Take Risk} + \text{Capacity to Bear Risk}$$

In our questionnaire, Question 3 and Question 4 are designed to determine investors’ risk tolerance. Then we combine investors’ risk tolerance with the asset allocation table recommended by McGill University finance professor Isabelle G. Bajeux-Besnainou and her coauthors,[10] which is shown in Table 4, to derive the ratios of bonds to stocks.

Allocating Assets

In last month’s article,[11] we pointed out that mean-variance optimal portfolios have high sensitivity to the variance-covariance matrix, resulting in extreme (“corner”) solutions. For example, the optimization portfolio always ordains large short positions when investors impose no constraints. And the optimization portfolio often prescribes corner solutions with zero weights in many assets when short positions are ruled out. In the optimal solution, those assets with positive pricing errors are significantly overweighted versus

Advisor and Investor Type	Percent of Portfolio			Ratio of Bonds to Stocks
	Cash	Bonds	Stocks	
T = 3 months				
γ				
8	67	15	17	0.87
5	51	21	28	0.76
2	-14	45	70	0.64
T = 1 Year				
γ				
8	55	28	17	1.6
5	39	33	28	1.17
2	-22	52	70	0.75
T = 5 Years				
γ				
8	10	73	17	4.18
5	22	74	28	2.65
2	-47	78	70	1.12
T = 10 Years				
γ				
8	-15	98	17	5.63
5	-25	97	28	3.48
2	-62	92	70	1.32

Here γ represents the risk aversion coefficient.

Table 4 — Asset allocations recommended by Bajeux-Besnainou et al.

those with negative errors. Besides, the risk-minimizing procedure tends to rely too much on assets with very low volatility relative to other assets rather than diversifying across a wide range of holdings.

To overcome these problems in mean-variance optimization and incorporate investors' views (i.e., the specific opinions investors have about asset returns) into this framework, we use the Black-Litterman (BL) model in Alpha UMa. The BL model provides not only the equilibrium market portfolio as a starting point for estimation of asset returns, but also a clear way to specify investors' views on returns and to blend the investors' views with prior information. The BL model's process is shown in Figure 4.

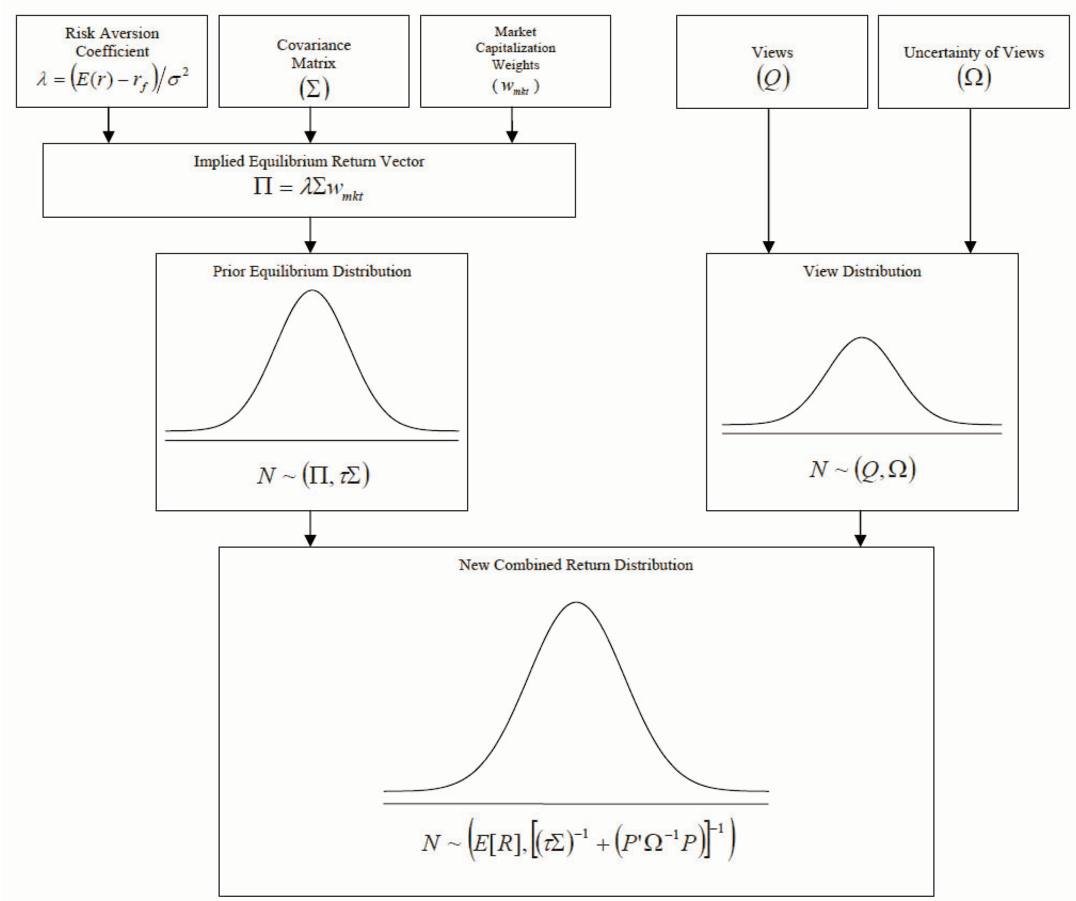


Figure 4 — Black-Litterman model: deriving new expected combined return vector, *E[R]*.

Using Bayes' theorem, we can derive the new expected combined return vector as follows:[12]

$$E[R] = [(\tau\Sigma)^{-1} + P^T\Omega P]^{-1}[(\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q]$$

where:

- $E[R]$ is the new combined return vector.
- τ is the weight-on-views scalar.
- Σ is the covariance matrix.
- P is the matrix that identifies the asset involved in the different views.
- Ω is a matrix that identifies the uncertainty in the views.
- Π is the implied equilibrium return vector.
- Q is the estimated return vector for every different view.

Computing the Implied Equilibrium Return

To derive the equilibrium return, we start from the quadratic utility function:

$$U = \omega^T\Pi - \left(\frac{\delta}{2}\right)\omega^T\Sigma\omega$$

where:

U represents investors' utility, which is the objective function during mean-variance optimization.

ω is the vector of weights invested in each asset.

Π is the vector of equilibrium excess returns for each asset.

δ is the risk aversion parameter.

Σ is the covariance matrix of the excess returns for the assets.

U is a convex function, so it will have a single global maximum. If we maximize the utility with no constraints, there is a closed-form solution. We find the exact solution by taking the first derivative of the utility function with respect to the weights ω and setting it to 0:

$$\frac{dU}{d\omega} = \Pi - \delta\,\Sigma\omega$$

Then we have the implied equilibrium excess return Vector Π:

$$\Pi = \delta\,\Sigma\omega$$

If you have no view to express, the BL model would tell you to hold the market portfolio and obtain market returns. In contrast, our view is a source of excess returns.

To calculate the implied equilibrium excess return, we have to know the risk aversion parameter δ, covariance matrix Σ, and vector of weights ω. We can get Σ from the historical data, and we can use the market scale of each asset class to express ω. To calculate δ, we have:

$$\delta = \frac{\boldsymbol{r} - \boldsymbol{r_f}}{\delta^2}$$

where:

r is the total return on the market portfolio.

r_f is the risk-free rate.

δ^z is the variance of the market portfolio.

Here the three variables are known to us. Then we can calculate the implied equilibrium excess return.

Specifying the Views

The BL model itself is just a view-combining engine; we cannot expect to generate excess returns from it. That is to say, if you have no view to express, the BL model would tell you to hold the market portfolio and obtain market returns. In contrast, our view is a source of excess returns. High-quality views can lead to superior performance for our portfolio.

In general, the BL model allows such views to be expressed in either absolute or relative terms. Consider three sample views:

- **View 1.** China Stock will have an absolute excess return (over risk-free rate) of 6.5% (confidence of view = 65%).
- **View 2.** QDII Emerging Equity will outperform QDII Developed Equity by 5% (confidence of view = 50%).
- **View 3.** China Stock and China Bond will outperform QDII Developed Equity and QDII Emerging Equity by 2% (confidence of view = 75%).

View 1 is an example of an absolute view, while Views 2 and View 3 represent relative views.

In practice, there are a lot of approaches to generating views and specifying the confidence level of views. Traditional institutional investors would employ financial analysts to conduct research on various industries and companies, but nowadays there are a lot of popular quantitative methods to generate views. For example, Morningstar's Thomas Idzorek has discussed approaches to determine the user-specified confidence level of views,[13] while Radford University finance professor Steven L. Beach and SEC economist Alexei G. Orlov have applied EGARCH-M models to generate views.[14]

Currently, Alpha UMa uses quantitative methods to generate views. According to the observed mean-reverting behavior of returns and the clustering of volatilities, we use AR-GARCH models to generate views and the corresponding uncertainty of views:

$$r_t = \mu + \varphi r_{t-1} + \epsilon_t$$

$${\sigma_t}^2 = \omega + \alpha_1 {\epsilon_{t-1}}^2 + \beta {\sigma_{t-1}}^2$$

$$\epsilon_t = \sigma_t z_t, z_t \sim N(0,1)$$

Of course, we need to conduct some statistical tests before generating our views, such as testing the stationarity of time-series data. A stationary process has the property that the mean, variance, and autocorrelation structure do not change over time. In other words, the first and second moments and autocovariance are time-invariant.

After completing all these tedious steps, we can take a look at the basic outline of our view-generating process:

1. Because the BL model assumes that views are independent, historical time-series data can be used to fit an AR-GARCH model for each of the selected funds with regression techniques.
2. We generate our views on a rolling basis. That is, for each day, we fit AR-GARCH models using the information on returns in the previous S days. So the obtained model parameters are also updated in every rolling period.
3. After we obtain the estimated model parameters for each day, we can substitute these parameters and returns into the model to get the predicted expected returns and volatilities for the next day.
4. Combining the predicted expected returns for all the selected funds, we produce the views. All the views here are absolute views, which are similar to View 1.

Before we can calculate the expected combined return vector *E[R]*, we still have to calibrate the scalar's value *(τ)*. When estimating the mean of a distribution, the uncertainty (variance) of the mean estimate will be proportional to the inverse of the number of samples. In Alpha UMa, we always have S samples on a rolling basis; hence τ is equal to $\frac{1}{s}$.

Thus, we have *E[R]* and calculate the optimal weight for our portfolio.

Monitoring and Rebalancing

Monitoring means tracking the performance of Alpha UMa each day to ensure it is working properly. With the prices moving, our portfolio might not stay optimized, so it is necessary to rebalance our portfolio. Even if we have high-quality views, Alpha UMa might not yield impressive performance without smart monitoring and rebalancing strategies, because transaction costs play an important part in the final profits.

Since the prices of our selected funds are varying, our views are also varying every day. Therefore, every day, our BL model would give new optimal weights of each fund in our portfolio for the next day. However, if we trade every day to meet the optimal weights, the transaction costs would not be affordable, thereby undermining our passive investing philosophy.

Instead of rebalancing the portfolio on a time basis, we decided to rebalance it on a threshold basis so that the strategy becomes more flexible. Our rebalancing strategy works simultaneously with our view-generating activity. That is to say, during our monitoring process, if we observe that the deviation of current weights from the predicted optimal weights exceeds the threshold, we rebalance our portfolio. Otherwise, we just hold the current portfolio to avoid transaction costs.

As shown in Figure 5:

1. We generate new views by using historical data to forecast the return and volatility for the next day.
2. Combined with the new market implied equilibrium return, we compute the new optimal weight every day.
3. We can calculate the excess weight, which is equal to the new optimal weight minus the current portfolio weight.
4. We rank the excess weight and determine whether the maximum or minimum excess weight exceeds the threshold: (a) If it does deviate from the threshold, then we are to compute the excess holdings of the fund with maximum excess weight at business day T. Once we successfully redeem the fund and receive

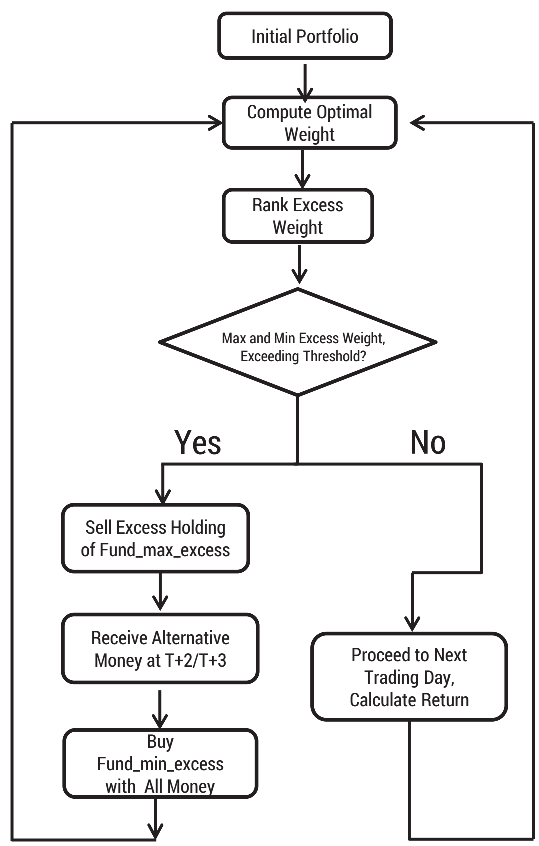


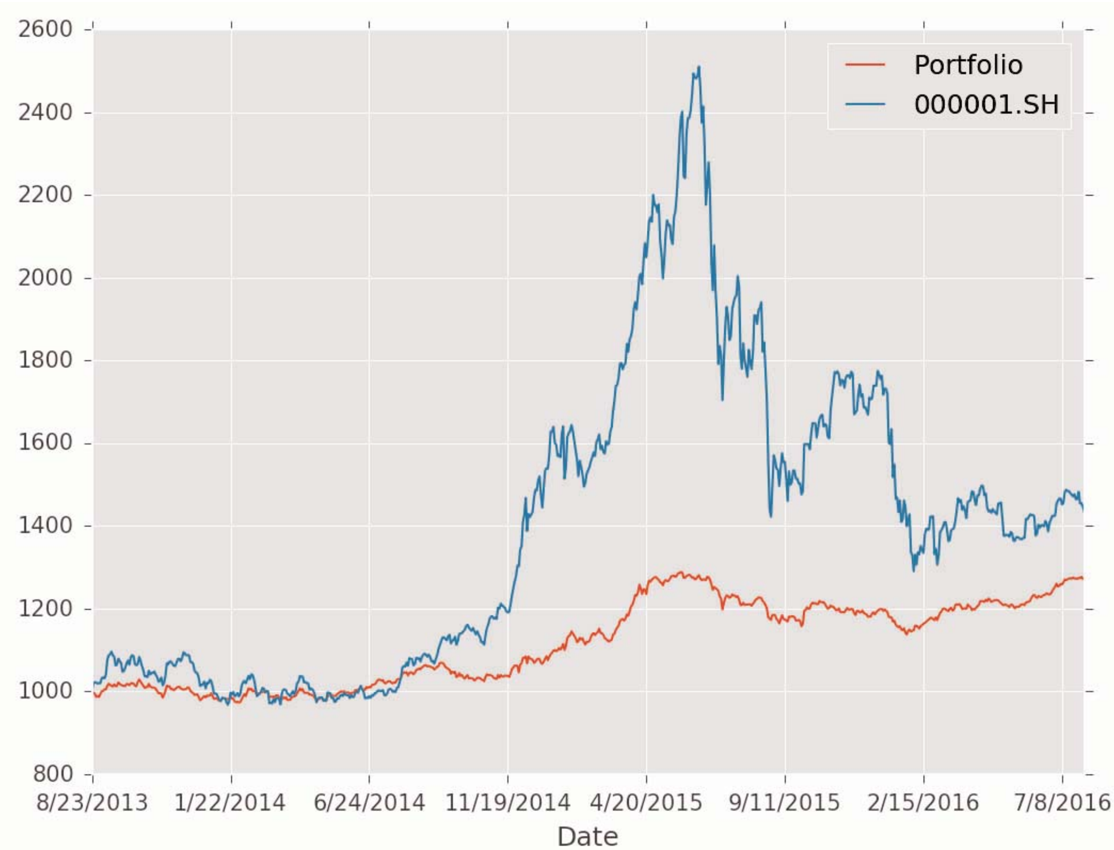
Figure 5 — Rebalancing process in Alpha UMa.

the money at business day T+2 or T+3,[15] we buy the fund with the minimum excess weight with all the money. (b) If it does not deviate from the threshold, we just proceed to the next trading day and calculate the current return.

In monitoring, we can also conduct pressure testing using a Monte Carlo method. In addition, there are other approaches to rebalancing a robo-advisor portfolio, such as using monthly cash-in flows to buy underweighted assets.

Evaluating Performance

Now we can use the historical data to do out-of-sample backtesting.[16] In this backtesting part, we have constructed a simplified portfolio that doesn't have constraints on the ratio of bonds to stocks. Instead, we limit the weight of each fund to be in the interval of [5%, 25%]. This means that the weight of each fund in the initial portfolio is no less than 5% and no more than 25%, thus making the initial optimized portfolio relatively diversified among the eight funds. We also set the threshold for each fund to be 30%. This implies that rebalancing is not needed except when the difference between the market weight and the newly optimized



	Portfolio	000001.SH
Annualized_Returns	0.1015	0.1549
Volatility	0.0798	0.2877
Sharpe_Ratios	0.8641	0.2877
Max_Drawdown	0.1176	0.4425

Figure 6 — Backtesting results of Alpha UMa.

weight of one fund in the portfolio is greater than 30%, thus making a balance between a sufficiently diversified portfolio and the high trading fee. To make sure each selected passive index fund has net asset value (NAV) data, the time starts from 23 August 2013 (before which some of the selected funds had not been set up).

As we can see from Figure 6, the neutral portfolio has an annualized return of 10.15%. Although the benchmark, the Shanghai Stock Exchange Composite Index, earns a higher annualized return, our portfolio's volatility is about one-third that of the benchmark. Thus, our portfolio's Sharpe ratio (the average return earned in excess of the risk-free rate per unit of volatility or total risk[17]) is about three times that of the benchmark. What's more, the maximum drawdown (MDD)[18] of our portfolio is only one-fourth that of the benchmark.

Summing Up

In Alpha UMa, we tested whether the investment methodology of three leading US robo-advisors — Betterment, Schwab Intelligent Portfolios, and Wealthfront — would also work in China's mutual fund market. We used the Black-Litterman model as a basic framework for our robo-advisor and incorporated an AR-GARCH model to serve as a view generator in order to seek for excess returns. In constructing Alpha UMa, we first classified assets into eight asset classes and then selected eight passive index funds. Then we used our online platform, Wangfubao, to obtain investors' goals, time horizon, and risk tolerance. After that, we could compute the optimal portfolio with our optimization algorithms. Finally, we performed monitoring and rebalancing every day.

The backtesting we conducted shows that our simplified portfolio has an annualized return of more than 10%, which is a very good result in a turbulent market. Next, we plan to use machine learning algorithms to generate views of higher quality, resulting in potential higher returns for our clients.

Endnotes

[1]Sharpe, William F. "The Arithmetic of Active Management." *Financial Analysts Journal*, Vol. 47, No. 1, January-February 1991 (www.cfapubs.org/doi/pdf/10.2469/faj.v47.n1.7).

[2]Campbell, John Y., and Luis M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, 2002.

[3]Xiao, Xinrong, and Rui Yang. "Study on the Investment Behavior of Securities Investment Funds in China" (in Chinese) Science Press, 2009 (www.science-definition.com/whatis/Study_on_the_Investment_Behavior_of_Chinese_Securities_Investment_Fund).

[4]Malkiel, Burton G., and Eugene F. Fama. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance*, Vol. 25, No. 2, May 1970 (http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1970.tb00518.x/full).

[5]Epperson, Teresa, et al. "Hype vs. Reality: The Coming Waves of 'Robo' Adoption." A.T. Kearney, June 2015 (www.atkearney.com/documents/10192/7132014/Hype+vs.+Reality_The+Coming+Waves+of+Robo+Adoption.pdf).

[6]Yang, Jie, et al. "How Robo-Advisors Manage Investment Portfolios." *Cutter Business Technology Journal*, Vol. 30, No. 4, 2017 (www.cutter.com/article/how-robo-advisors-manage-investment-portfolios-495656).

[7]Bogle, John C. "The Economic Role of the Investment Company" (Senior Thesis). Department of Economics and Social Institutions, Princeton University, 1951 (https://pulsearch.princeton.edu/catalog/dsp017m01bm63k).

[8]The data resource is Wangfubao, a research project of State Street (Hangzhou). It is an online platform that helps end users simulate and analyze fund investments.

[9]Roszkowski, Michael J., Geoff Davey, and John E. Grable. "Insights from Psychology and Psychometrics on Measuring Risk Tolerance." *Journal of Financial Planning*, Vol. 18, No. 4, April 2005 (www.onefpa.org/journal/Pages/Insights%20from%20Psychology%20and%20Psychometrics%20on%20Measuring%20Risk%20Tolerance.aspx).

[10]Bajeux-Besnainou, Isabelle, James V. Jordan, and Roland Portait. "An Asset Allocation Puzzle: Comment." *American Economic Review*, Vol. 91, No. 4, September 2001 (www.aeaweb.org/articles?id=10.1257/aer.91.4.1170).

[11]Yang et al. (see 6).

[12]He, Guangliang, and Robert Litterman. "The Intuition Behind Black-Litterman Model Portfolios." Goldman Sachs Investment Management, December 1999 (https://faculty.fuqua.duke.edu/~charvey/Teaching/IntesaBci_2001/GS_The_intuition_behind.pdf).

[13]Idzorek, Thomas M. "A Step-by-Step Guide to the Black-Litterman Model," 2002-2005 (https://corporate.morningstar.com/ib/documents/MethodologyDocuments/IBBAssociates/BlackLitterman.pdf).

[14]Beach, Steven L., and Alexei G. Orlov. "An Application of the Black-Litterman Model with EGARCH-M-Derived Views for International Portfolio Management." *Financial Markets and Portfolio Management*, Vol. 21, No. 2, June 2007 (https://link.springer.com/article/10.1007/s11408-007-0046-6).

[15]We have to subscribe to and redeem passive index funds in the over-the-counter (OTC) market; thus, we have different trading fees and trading rules for different funds. The trading rule for QDII funds is T+2, which means our trading requests will be confirmed two business days after the request, and we receive the money at T+3.

[16]Statistical tests of a model's forecast performance are commonly conducted by splitting a given data set into an in-sample period, used for the initial parameter estimation and model selection, and an out-of-sample period, used to evaluate forecasting performance. Out-of-sample testing is a way to guard against curve-fitting. It is a good practice because we don't know how the market will go in the future. Backtesting is the process of testing a trading strategy on relevant historical data to ensure its viability before the trader risks any actual capital. A trader can simulate the trading of a strategy over an appropriate period of time and analyze the results for the levels of profitability and risk.

[17]"Sharpe Ratio." Investopedia (www.investopedia.com/terms/s/sharperatio.asp).

[18]A maximum drawdown (MDD) is the maximum loss from a peak to a trough of a portfolio, before a new peak is attained. MDD is an indicator of downside risk over a specified time period.

Jie Yang, Hanxi Ye, Yadan Wei, and Linqian Bao are Business System Analysts for the Wangfubao Project in State Street Technology (Zhejiang). Since 2010, Wangfubao has been dedicated to helping retail investors in China plan, save, and invest in the mutual fund market to reach their financial goals. Wangfubao tries to combine interactive visual analysis methods for high-dimensional time series data, collective intelligence of the investor community, and modern portfolio theory to build investment strategies and do asset allocation. The authors can be reached at JYang7@StateStreet.com, HYe@StateStreet.com, YWei@StateStreet.com, and LBao@StateStreet.com.

Appendix: Investor Questionnaire

Goal

1. My goal for this account is to:
   a. Build wealth
   b. Save for expenses
   c. Build a rainy day fund for emergencies

Time Horizon

2. The length of time I plan to invest before I begin withdrawals:
   a. Less than 3 months
   b. 1 year
   c. 5 years
   d. 10 years

Risk Tolerance

3. In the year I lost 20% of my investments/If I ever were to lose 20% or more of my investments in one year, I would:
   a. Sell everything
   b. Sell some
   c. Do nothing
   d. Reallocate my investments
   e. Buy more
4. My current and future income sources (such as salary and pension) are:
   a. Unstable
   b. Somewhat stable
   c. Very stable

Challenges and Success Factors in Global Test Outsourcing

by Xiaochun Zhu and Shanping Li

Nowadays, many companies contract their testing functions out to third-party IT outsourcing companies. This process, referred to as test outsourcing, is common in the industry, yet it is rarely studied in the research community. To bridge this gap, we performed an empirical study on test outsourcing with 10 interviewees and 140 survey respondents. We investigated various research questions and found that customer satisfaction expectations, tight project schedule, and domain unfamiliarity are the top three challenges testers face.

Overview

In test outsourcing, software testing is carried out by an independent organization, which can improve the quality of the applications and reduce risks through rigorous testing. With the financial industry's high demand for software quality, it is typical for financial organizations to outsource their software testing.

Several researchers have investigated test outsourcing. Lappeenranta University of Technology's Ossi Taipale and his coauthors showed that test outsourcing increases the efficiency and reduces the cost of software testing.[1] In the classic *Testing Computer Software*, Cem Kaner and his coauthors argue that product reliability will be better if independent test organizations conduct testing.[2] Unfortunately, despite the growing interest in outsourcing in general and test outsourcing in particular, there has been no study that comprehensively investigates the types, processes, and challenges of test outsourcing.

To learn about the test practices followed by developers involved in test outsourcing, we conducted an empirical study on a major IT company that has a large test outsourcing team working for global financial clients. We wanted to understand the different characteristics of test outsourcing and the different tools and techniques testers frequently use. To do this, we conducted one-on-one interviews with 10 senior QA managers and team leaders to get an in-depth understanding of the types, processes, and challenges involved in test outsourcing.

Study Methodology

In this section, we present our study methodology, which involved two parts: qualitative interviews and a survey.

The Interviews

Protocol

The interview format was semistructured and divided into three parts. In the first part, we asked some demographic questions, such as what experience the interviewee had in test outsourcing. In the second part, we asked some open-ended questions, such as what challenges test outsourcing teams faced. In the third part, we picked a list of topics related to test outsourcing and asked the interviewees to discuss topics that they had not explicitly talked about.

With the financial industry's high demand for software quality, it is typical for financial organizations to outsource their software testing.

Participant Selection

We conducted interviews with senior QA managers and team leaders at Insigma Technology, the second-largest IT outsourcing company in China.[3] With more than 6,000 employees, Insigma was ranked 24th on the International Association of Outsourcing Professionals' (IAOP) "Global Outsourcing 100" list in 2014.[4] A total of 10 people accepted the invitation. Six interviewees were from Insigma Hengtian (IH), an outsourcing company mainly for client companies from US and Europe. IH has more than 1,600 employees. The other four interviewees were from Insigma Global Service (IGS), an outsourcing company mainly for client companies

Interviewee	Working Experience	Role
P1	10 years	Head of the QA department
P2	10 years	Senior QA manager
P3	8 years	Senior QA manager
P4	8 years	Senior QA manager
P5	10 years	Project manager
P6	3 years	Project manager
P7	6 years	Project manager
P8	5 years	Team leader
P9	5 years	Team leader
P10	5 years	Team leader

Table 1 — Working experience and roles of the 10 interviewees.

within China. IGS has more than 400 employees. Eight interviewees were male, and two were female. In the remainder of the article, we denote these 10 interviewees as P1 to P10. Table 1 presents the working experience and roles of the 10 interviewees.

Data Analysis

After the interviews, we used a transcription service to transcribe the audio into text and grouped the text into different categories by performing a card sort.[5] We first extracted some keywords for each of the topics. Next, for each interviewee, we assigned responses to a topic if they contained the corresponding keywords. The process was repeated until all responses made by the interviewees were covered. Then we analyzed these topics and the related responses and grouped them into three different categories:

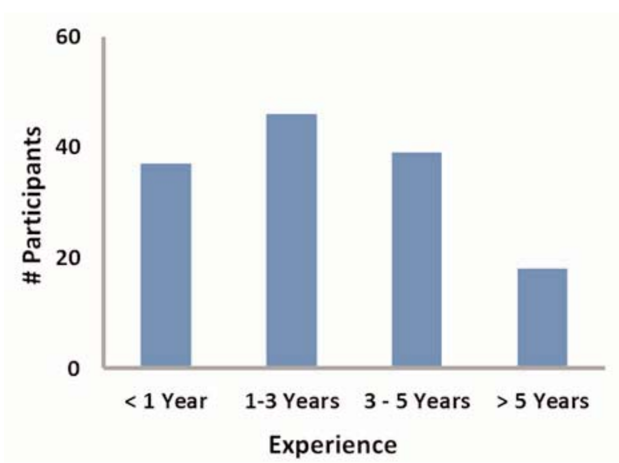


Figure 1 — Experience levels of survey participants.

1. General characteristics
2. Technical aspects of test outsourcing
3. Management aspects of test outsourcing

The Survey

Protocol

We designed a survey to validate hypotheses that we formulated based on the interviews. The goal of the survey was to quantify the qualitative results expressed by the 10 interviewees over a range of topics. We also asked respondents to fill in more specific multiple choice and open-ended questions to help us better understand test outsourcing. We asked respondents about ways to test outsourced projects, test types, types of test outsourcing projects they have participated in, test techniques, test support tools, and challenges. We also asked them to envision an automated testing tool that could help in test outsourcing projects. Finally, we collected demographic information from the respondents.

Participant Selection

We recruited respondents in the QA department of Insigma Hengtian and Insigma Global Service to participate in the survey. IH and IGS have more than 500 and 100 testers, respectively.

In total, we asked 428 testers to complete the survey, and 140 testers did so, yielding a response rate of 32.7%. Figure 1 presents the survey participants' years of experience.

Data Analysis

We examined the distribution of responses from the respondents. We linked interviewee comments with survey responses by referring to survey statements.

Study Findings

General Characteristics

Project Types

We refer to companies that outsource some of their IT functions as client companies. Test outsourcing can be categorized according to the IT capabilities of the client company. Some companies focus on areas other than IT, so they always contract their development and testing efforts out to an outsourcing company. According to the interviewees, especially P1 to P4, test outsourcing can be categorized into three types (see Figure 2):

1. **Basic.** The client company has a strong test team, the outsourcing projects are well documented, and the test specifications and test cases have been designed. The main task for test outsourcing professionals is to follow the instructions given by the client company and to test the projects according to the descriptions in the test specifications and test cases.
2. **Intermediate.** The client company has a test team, but the testers in the outsourcing company need to work together with the testers in the client company to establish the test plan, design the test cases, build the test system, and complete the whole test process.
3. **Advanced.** The client company has no test team, and most of their IT functions are outsourced. The testers in an outsourcing company need to help the client company establish the test plan, design the test cases, build the test system, and complete the whole test process.

Test Process

Not surprisingly, the test processes used in the different types of test outsourcing projects are also different. For a project of the basic type, the client company typically sends some senior testers to the test team in the outsourcing company, and they will organize some training sessions for the outsourcer's testers. The training sessions provide "an introduction of domain knowledge, background and basic operations of the projects, and test environment" (P1). They also provide "an explanation of the requirements and test case documents" (P1). After training, the outsourcing team's testers begin to test the system according to "the test cases, compare the outputs of the system with the expected outputs of the test cases, and report bugs found" (P4).

For a project of the intermediate type, testers in the client company and in the outsourcing company work together to design test cases and complete the whole test process. Similar to projects of the basic type, in the beginning some senior testers in the client company will train testers in the outsourcing company. Then they "work in a collective way to perform the test process; for example, testers in the outsourcing company may design test cases, and testers in the client company may review the test cases, and then they work together to execute the test cases and report bugs" (P3).

Projects of the advanced type are different from projects of the other two types. "A brainstorm session is commonly held to understand requirements from a client company" (P3). A lot of effort is spent on communication and discussion in the setup phase of the projects. Since the client company has limited experience in testing, testers in the outsourcing company also need to "design a detailed test plan" (P1). These testers then need to "train people in the client company to make sure the client company understands and accepts what they want to do" (P2). After that, these testers begin to design detailed test cases and run them to find bugs.

Challenges

There are various challenges that can affect the success of test outsourcing projects, such as customer satisfaction expectations, time constraints, poor documentation, and domain unfamiliarity. All 10 interviewees agreed that customer satisfaction — understood here as high expectations from the customer for the outsourcing

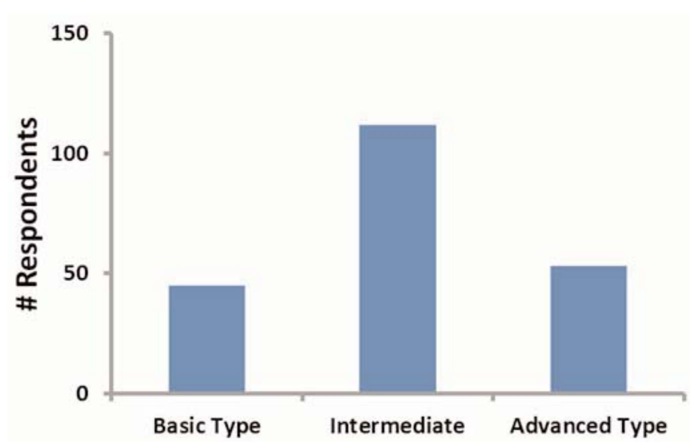


Figure 2 — Test outsourcing projects by type.

team to meet — is one of the biggest challenges for test outsourcing projects (P1 to P10). Furthermore, these customer satisfaction expectations vary across the different types of test outsourcing projects:

1. **Basic.** Since the client company provides all test documentation to the outsourcing company's test team, in order to ensure customer satisfaction, the team needs to "follow the test cases and plans provided by the client company, test the system based on the test cases, and complete the project on time" (P6). Sometimes, the test cases and test plans the client company provides may contain some problems. For example, some test cases may be "unreasonably designed or even wrongly designed and in conflict with the requirements documents" (P6), some test cases are "hard to execute and reproduce (e.g., testing concurrency modules in a system)" (P8), and some test plans "are too packed, which makes it hard to complete the project on time" (P9).

2. **Intermediate.** Since testers in the outsourcing company need to work closely together with testers in the client company, making the latter "recognize the professional level and ability of testers from the outsourcing company" (P6) is a challenge for projects of the intermediate type. As P6 stated: "Testers from both companies work much closer for projects of the intermediate type than projects of the other two types. Testers from the client company evaluate the abilities of testers from the outsourcing company. If the client thinks that the service level of the outsourcing company is not good, they will transfer to another outsourcing company. Thus, how to make the customer recognize the service and professional level of the outsourcing company is important for projects of the intermediate type."

3. **Advanced.** For projects of the advanced type, "since the customers have limited knowledge on software testing," the challenge is that "the testers need to make the customers understand what they want to do, and how they will accomplish it" (P7). To make the client company trust that the test team can do the work, "testers need to write more detailed documents and communicate with the customers well" (P8).

Eight of the 10 interviewees agree that domain unfamiliarity is another major challenge to the success of test outsourcing projects (P1-P5, P7, P8, P10). A tight project schedule is also a significant difficulty that affects the success of the projects (P1-P3, P5, P7, P9).

Figure 3 shows the challenges that testers in our survey faced. It is interesting to note that customer satisfaction (104 votes), tight project schedule (80 votes), and domain unfamiliarity (78 votes) are the top three obstacles cited. Notice that the number of votes (29) for "limited tool support" is comparatively low. P1 and P2 mentioned that in test outsourcing projects, testers prefer manual testing over automated testing; thus, the need for advanced tool support is not as important as the other challenges.

Technical Aspects of Test Outsourcing

Manual vs. Automated Testing

In test outsourcing projects, "testers prefer manual testing to automated testing" (P1). Automated testing can help reduce the amount of repetitive work in the testing

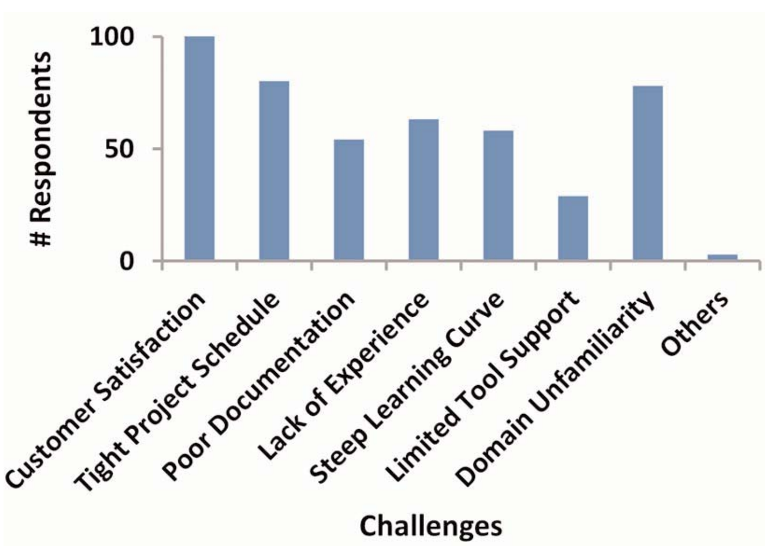


Figure 3 — Challenges testers faced in test outsourcing projects.

process, but automated testing is only used in a limited way in test outsourcing projects.

Figure 4 shows the testing methods used by survey respondents in their test outsourcing projects. Among the 140 survey respondents, 88 (62.9%) performed only manual testing, 8 (5.7%) performed only automated testing, 41 (29.3%) performed both manual and automated testing, and 3 (2.1%) performed no testing in past test outsourcing projects. (These last three respondents are senior managers, who focused instead on project management activities.)

Test Levels and Types

There are various test levels (unit testing, integration testing, system testing) and test types (functional testing, nonfunctional testing).[6] Not all of these test levels and types are frequently performed in test outsourcing projects.

Unit testing is an important test type. Delft University of Technology researcher Michaela Greiler and her coauthors found that unit testing plays an important role in the testing of plug-in systems.[7] However, in test outsourcing projects, unit testing is less often performed unless the customer requires it. Unit testing is often done by the developers who wrote the code, as they want to get the feedback as quickly as possible. Furthermore, unit testing will "add cost to the client company, and also increase testing time which causes delays to project completion" (P4). Since most of the test outsourcing projects have a "tight schedule and limited budget" (P4), to ensure projects are completed on time, clients prefer system testing to unit testing. For outsourcing projects of the basic type, there are often "no requirements for unit testing, since the client company may also perform unit testing within the company" (P1).

Supporting Tools

All 10 interviewees agreed that issue tracking tools are the most widely used tools as compared with other tools such as code inspection tools and functional test tools. All test teams need to report and manage bugs; the usage of issue tracking tools can help them "analyze bugs, and compute some statistics such as bug density and bug distributions" (P6). From the customer point of view, issue tracking tools can "help them to evaluate the quality of the system and better understand the status of the current system" (P8).

Test Case Design

For outsourcing projects of the intermediate and advanced types, testers need to design test cases. Eight of the 10 interviewees agreed that domain knowledge impacts test case design (P1, P3, P5-P10).

Previous studies show code coverage is an important metric to evaluate the effectiveness of test cases.[8] However, in test outsourcing projects, often code coverage is not deemed important for evaluating the design of test cases. Instead of code coverage, client companies care more about "requirements coverage"; that is, whether the generated test cases have covered all of the requirements.

Eight of the 10 interviewees agree that domain unfamiliarity is another major challenge to the success of test outsourcing projects.

Management Aspects of Test Outsourcing

Training, Knowledge Sharing, and Knowledge Transfer

In an outsourcing company, the turnover rate is high. Test teams will be affected if too many testers leave the team. Training, knowledge sharing, and knowledge transfer are three effective ways to reduce the risk of turnover-related knowledge loss, especially when the project team has a number of new testers.

Communication Skills and Client Relationship Management

All 10 interviewees agreed that communication skills are extremely important for an outsourcing company. As we described in the previous section, customer satisfaction is the biggest challenge in test outsourcing,

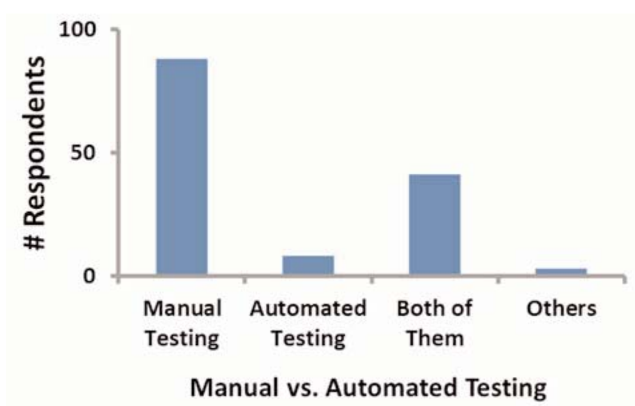


Figure 4 — Distribution of survey respondents who performed manual and automated testing.

and communication skill can help to reduce the risk of customer dissatisfaction.

Tester Experience

The experience level of testers is another important factor that affects the success of test outsourcing projects, especially for projects of the advanced type. In practice, a typical outsourcing company has only a small number of experienced testers and a large number of inexperienced testers. Thus, most of the time, experienced testers have to work across many different test projects at the same time.

Recommendations

Financial services organizations that are undertaking test outsourcing would be well advised to:

- **Understand the purpose of the outsourcing project.** Is the client company engaging the outsourcing team as a resource supplement, to complete a temporary testing task, or to help improve its testing capacity? Different purposes will lead to different types of outsourcing projects.
- **Build an efficient collaboration and management model between the client's project team and the outsourcing team.** There are three different collaboration and management models to consider: (1) the client company assigning its managers to manage the outsourcing team, (2) the outsourcing company managing the outsourcing team with supervision from the client company, and (3) the client company inviting a third-party company to manage the outsourcing team. Choosing the best approach will depend on the client company's needs and both companies' capabilities.
- **Leverage the outsourcing team's capacity to drive a joint development and testing model between the client company and the outsourcing company.** The outsourcing team should be not only an extension of the client's resource pool, but also a strategic capacity center on business and system knowledge as well as new technologies.

We hope our study will contribute to a better understanding of the state-of-practice of test outsourcing in the financial services industry. In the future, we plan to investigate more aspects of test outsourcing and reduce the threats to external validity by interviewing and surveying more people from more test outsourcing companies.

Endnotes

[1]Taipale, Ossi, Kari Smolander, and Heikki Kälviäinen. "Cost Reduction and Quality Improvement in Software Testing." *Proceedings of the 14th International Software Quality Management Conference (SQM '06)*. Southampton, UK, April 2006 (www.academia.edu/11365235/Cost_reduction_and_quality_improvement_in_software_testing).

[2]Kaner, Cem, Jack Falk, and Hung Quon Nguyen. *Testing Computer Software*. 2nd edition. John Wiley & Sons, 1999.

[3]"Awards and Recognition." Insigma, 2012 (www.insigmaus.com/about-us/awards-and-recognition/).

[4]"The 2014 Global Outsourcing 100." IAOP, 2014 (www.iaop.org/Content/ 19/165/3879).

[5]Spencer, Donna. "Card Sorting: A Definitive Guide." *Boxes and Arrows,* 7 April 2004 (http://boxesandarrows.com/card-sorting-a-definitive-guide/).

[6]Pezze, Mauro, and Michal Young. *Software Testing and Analysis: Process, Principles, and Techniques*. John Wiley & Sons, 2008.

[7]Greiler, Michael, Arie van Deursen, and Margaret-Anne Storey. "Test Confessions: A Study of Testing Practices for Plug-In Systems." *Proceedings of the 34th International Conference on Software Engineering*. IEEE Press, 2012.

[8]Andrews, James H., et al. "Using Mutation Analysis for Assessing and Comparing Testing Coverage Criteria." *IEEE Transactions on Software Engineering*, Vol. 32, No. 8, August 2006 (http://ieeexplore.ieee.org/document/1703390/); Gopinath, Rahul, Carlos Jensen, and Alex Groce. "Code Coverage for Suite Evaluation by Developers." *Proceedings of the 36th International Conference on Software Engineering*. ACM, 2014 (http://dl.acm.org/citation.cfm?id=2568278); Hutchins, Monica, et al. "Experiments of the Effectiveness of Dataflow- and Controlflow-Based Test Adequacy Criteria." *Proceedings of the 16th International Conference on Software Engineering*. IEEE Computer Society Press, 1994 (http://dl.acm.org/citation.cfm?id=257766).

Xiaochun Zhu is VP and the Director of Software Quality Assurance at Insigma Hengtian, a joint venture company of State Street and Insigma Technology. Dr. Zhu received his PhD in computer science in 2010 from Zhejiang University, one of the top three universities in China. His main research areas are quality assurance, software testing, and software engineering. He can be reached at joezhu@hengtiansoft.com.

Shanping Li has been with Insigma Hengtian since 2004, where he serves as CIO. Dr. Li has been a professor of Zhejiang University since 1998. He works at the College of Computer Science and the College of Software Engineering. Dr. Li's main research areas are software engineering, distributed computing, and the Linux kernel. He received his master's degree in 1987 and his PhD in 1993, both in computer science, from Zhejiang University. He can be reached at shan@zju.edu.cn.

Refining Reconciliation: A Machine Learning Approach to the Financial Industry's Toughest Task

by Zhuo Li and Jianling Sun

Introduction

Reconciliation can be defined as "the key process used to determine whether the money leaving an account matches the amount spent, ensuring the two values are balanced at the end of the recording period."[1] According to generally accepted accounting principles (GAAP), the purpose of account reconciliation is to provide accuracy and consistency in financial accounts. To ensure all cash outlays and inlays match between cash flow statements and income statements, it is necessary to carry out reconciliation. Theoretically speaking, there will be at least

$$\frac{n \times (n-1)}{2}$$

reconciliations for a financial workflow involving n accounts. Modern financial workflows contain more roles and steps with increased product complexity, which makes reconciliation the most time-consuming and painful component of the financial services industry. Big companies set up dedicated reconciliation teams to do this work on a regular basis — often daily or at least monthly. It is embedded in the workflow and accounts for "as much as needed" time to ensure process accuracy.

There have been efforts to automate the reconciliation process by building rule engines to match transactions and identify breaks. However, the rules engaged are traditionally derived from subject matter experts' experience and knowledge, which are not always available and may not cover all possible "match" and "break" scenarios. If any new issues are raised, significant human intervention will be needed to research their root causes and derive new rules. What financial services organizations need is an intelligent and automated rule-mining approach to find thorough and effective rules from massive data sets, which will reduce analysis cycle time, increase rule coverage, and ultimately benefit straight-through processing. Machine learning attracts our attention in this space for its capability to "learn" from massive data sets.

In this article, we examine different machine learning mechanisms and propose a maximally specific conjunctive approach to fitting massive data sets in the real world of reconciliation. Furthermore, we provide a balanced solution to address the high skewness in reconciliation data sets. (Skewness refers to the fact that the 80% or more of financial data matches take limited processing time, while the 20% or fewer "reconciliation breaks" account for most of reconciliation process time for research and resolution.) Our approach first uses a kNN (k Nearest Neighbor)[2] algorithm to do undersampling, which means reducing positive samples to make the data set more balanced. It then uses a decision tree algorithm to do rule mining. But before we delve further into our approach, let's meet the system that inspired our research efforts.

Modern financial workflows contain more roles and steps with increased product complexity, which makes reconciliation the most time-consuming and painful component of the financial services industry.

The Reconciliation System

Our research is based on a reconciliation system that receives and reconciles transactions from multiple data sources across different business areas like accounting, trade instruction (e.g., SWIFT [Society for Worldwide Interbank Financial Telecommunication] messages), custody, and others. Over 3 billion transactions covering almost all known transaction types flow through the system per year. A simplified relationship diagram is shown in Figure 1.

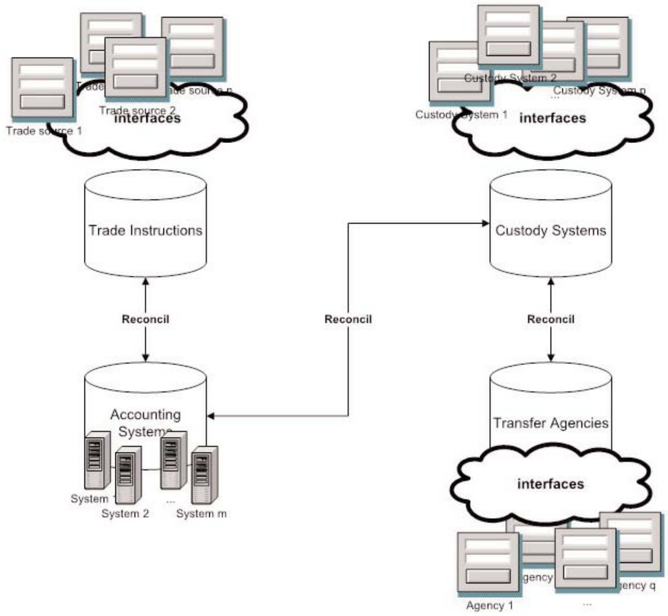


Figure 1 — Reconciliation general view.

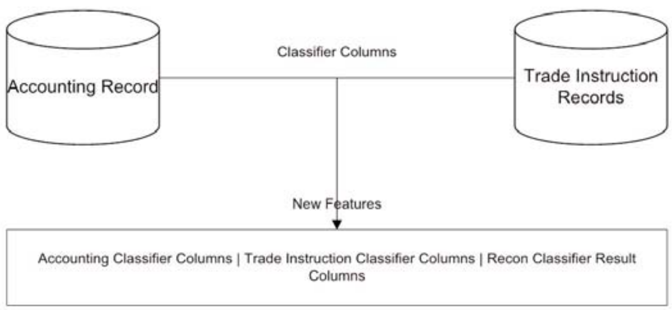


Figure 2 — Extended reconciliation training record.

The reconciliation rules the system currently uses are derived by domain experts and converted into executable libraries by IT professionals. Relying on these rules, the system identifies predefined transaction match or break patterns, intelligently tags them either with a match status or a break type, and/or automatically remediates identified breaks. However, due to the variety of transaction types involved, there are still thousands of reconciliation breaks reported daily, and a big team is required to analyze and process them. While generic reconciliation rules do exist, the intricacies of all the involved business lines mean that such rules are inappropriate for uncovered transaction types. There are still many possible scenarios that can't be understood and caught. To make things worse, it is a long process to locate the right expert to analyze these false alarms, propose a new rule to handle the breaks, and test it out in production. The goal of our research is to find an efficient way of mining correct reconciliation rules to shorten the analysis process and ultimately reduce the effort needed to analyze the reconciliation breaks.

Supervised Reconciliation Rule Mining

In trialing our proposed approach, we started with a training data set abstracted from a one-day production time window from both accounting and trade instruction data sources. The data set is supervised, which means a label is added to each transaction from it to indicate whether it is a match (meaning an accounting transaction can find a corresponding trade instruction transaction or vice versa) or a break (meaning no match can be found). This is achieved by flowing the transactions through the reconciliation system; a domain expert completes the review. The data structure is extended to include the label information as well (see Figure 2).

Rule mining is a classic data classification problem, and the decision tree[3] is a classic data classification algorithm.[4] It has the following pros and cons:

Pros:

- A derived decision tree can easily be converted to human-readable rules.
- It can be constructed once and used repeatedly, which is highly efficient.
- Derived rules can be integrated into the rule engine programmatically.

Cons:

- The training set must be labeled beforehand, which is not always feasible.
- When a training set is "unbalanced," the dominant label will cover up rules that only impact a small portion of the transactions.
- "Noise" data can easily introduce incorrect rules.

Speaking of unbalanced training sets, they are a serious challenge in real reconciliation processing. Most transactions from the accounting source and the trade instruction source can find their matched counterpart from the other side's data source in a straightforward way by matching fewer than a dozen transaction elementary fields such as client name, fund name, currency, trade date, and the like. But if we directly apply such real production data onto a decision tree as a training set, an unbalanced data set will most likely lead the decision tree to a biased result set, ignoring those minor transactions that actually contain the truly valuable knowledge we want to mine out.

To resolve this challenge, we introduced a customized kNN undersampling algorithm to derive a better balanced training set, filtering out more matched transactions while ensuring that targeted outliers (those transactions leading to reconciliation breaks) are still reserved. To achieve this, as mentioned above, we "supervise" the training set, labeling selected transactions either positive (matches) or negative (breaks). For the identified negative transactions, we find their nearest k positive transactions. Here the "distance" among transactions is measured by their similarities, and similarity is measured by data values from the transaction elementary fields mentioned earlier. This way, the size of the total positive transactions is controlled to be under k times the total negative samples, and most related positive samples are identified.

Once the kNN algorithm derives a balanced training set, we can apply the decision tree to produce detailed rules as follows (see Figure 3):

1. Group the unbalanced training set into positive samples and negative samples.
2. Starting with the negative samples, identify kNN positive sample records for each negative sample record.
3. Combine all negative sample records and all identified kNN positive sample records into a new balanced training set.
4. Using the new balanced training set as input for the decision tree algorithm, where the ratio between positive samples and negative samples is K:1:
   a. Transform the decision tree into classification rules.
   b. Apply the derived classification rules upon the unlabeled data to the labeled data set.
   c. Evaluate the result set evaluation.

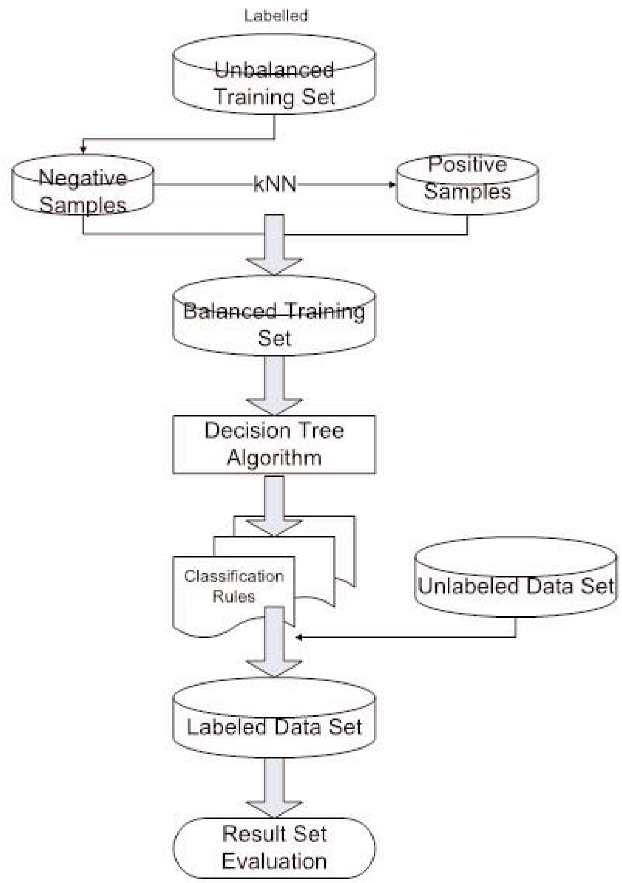


Figure 3 — Extended reconciliation rule-mining algorithm.

Evaluation Criteria and Experiment

Evaluation Criteria

We designed three quantitative criteria to evaluate the result set: rule concision, rule reliability, and forecast failure ratio.

Rule Concision: Con(R)

Rule concision determines the complexity of the decision tree. More complex design trees will face more challenges in transforming and validating rules. When rule numbers are the same, rule concision is related to the average depth of rules: the deeper the rule, the more complex it is. Assuming the rule set from a decision tree is *R*, the rule set's size (number of rules) is *N*, and the length of *ith* rule is *L(i)*, *Con(R)* can be defined as:

$$\mathrm{Con(R)} = \frac{\sum_i^N L(i)}{\mathrm{N} \cdot \log_2 N}$$

The simplest case is that every rule has the same length $log_2 N$, then *Con(R)* = 1; the worst case is every rule has a different length from 1, 2,…, N-1, then

$$Con(R) = \frac{(N+2)(N+1)}{2N \cdot \log_2 N} \approx \frac{1+N}{2 \cdot \log_2 N}$$

(see Figures 4a-4b).

Rule Reliability: Rel(R)

Rule reliability stands for the alignment level between rules and real business logic. Feature columns in the

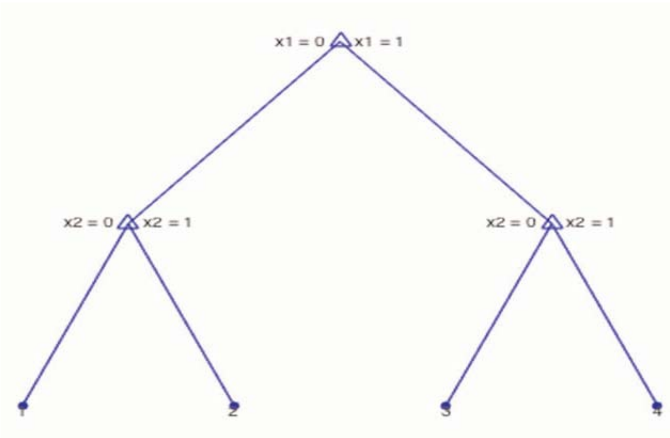


Figure 4a — Rule concision: the simplest case.

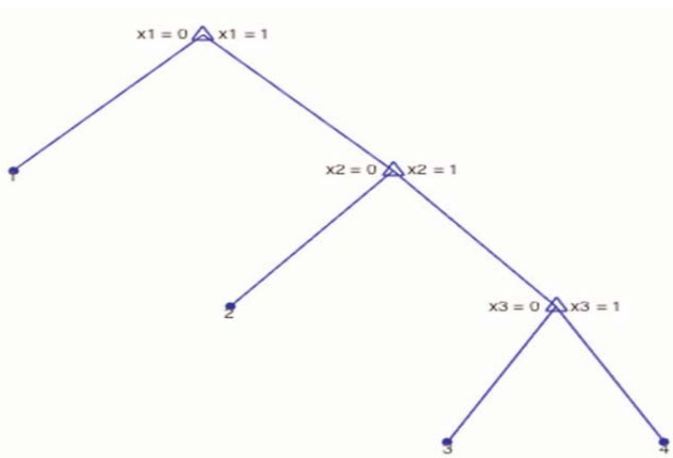


Figure 4b — Rule concision: the worst case.

training data set have two types: unilateral and bilateral. Unilateral columns indicate transaction type, and bilateral columns indicate value liaison status between accounting and trade instruction. The higher the unilateral columns stay in the decision tree hierarchy, the more reliable the rule set is. Assuming rule set R, rule set size N_R, unilateral column set C, size N_C, and *ith* unilateral column depth $L(i)$, then rule reliability *Rel(R)* is defined as:

$$Rel(R) = \frac{N_R \cdot \log_2 N_R}{\sum_i^{N_C} L(i)}$$

The higher *Rel(R)* is, the more reliable the rule is.

Rule Forecast Failure Rate: FFR(R)

As part of the evaluation process, derived rules will be applied against an unlabeled data set, and domain experts are engaged to conduct review to identify reasonable rules. Assuming rule set R, rule set size N_R, and identified reasonable rule number N_G, the Rule Forecast Failure Rate *FFR(R)* is defined as:

$$FFR(R) = 1 - \frac{N_G}{N_R}$$

A special note here: *FFR(R)* is only feasible against a balanced data set after kNN tailoring. This is because negative cases are too rare in unbalanced data sets, and a few forecast failures could bring excessive impact on the final decision tree.

Benchmark Experiments

The initial training set contains 35 columns in total; 12 are unilateral columns and 23 are bilateral columns. We applied an enhanced decision tree algorithm with kNN undersampling on it to evaluate the performance; the same training set is applied to the raw decision tree algorithm and enhanced decision tree algorithm, but with just random undersampling as the benchmark.

There have been many existing machine learning algorithms, including but not limited to kNN, decision tree, artificial neural network (ANN),[5] support vector machine (SVM),[6] naive Bayes,[7] and more. We won't be able to cross-compare every single algorithm here, but there is one critical reason we chose to extend the decision tree algorithm for our experiment: it can provide an explicit inductive reasoning path why "learning" gives you the result of "match" or "break," which can later on be converted into reconciliation rules for review and integration with the reconciliation system in a straightforward way. An algorithm like ANN might have better performance in specific scenarios, but it can't output "explainable" rules for expert review. This means it can't be adopted in a reconciliation system because the completeness of rules is not assured from a logical perspective, and even experts can't judge a rule's thoroughness without knowing the logical reason behind it.

Raw Decision Tree

A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including observed probability event outcome, utility, and so on. As briefly mentioned above, it is specifically useful in machine learning for scenarios where an explicit reasoning process is expected, because a decision tree can be linearized into decision rules[8]

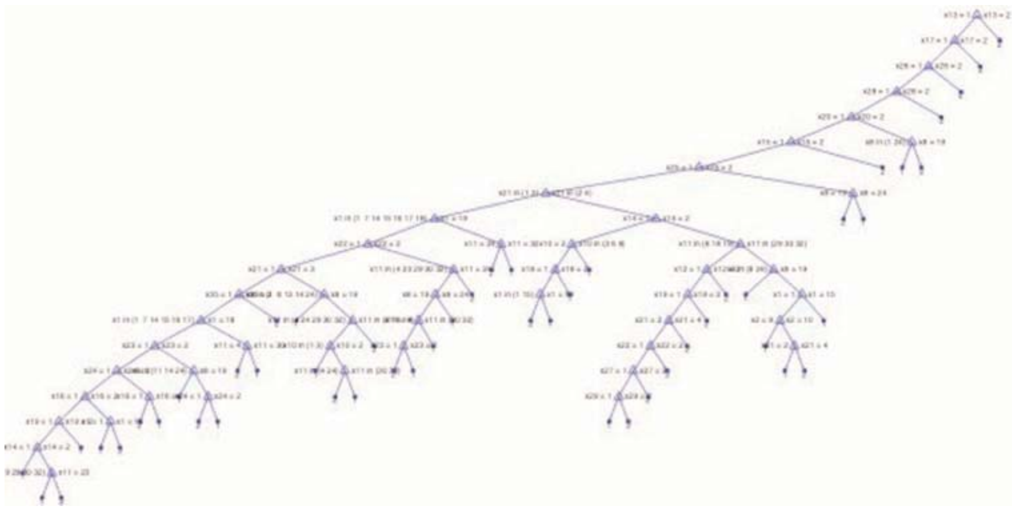


Figure 5 — A decision tree of unbalanced data.

where the outcome is the contents of the leaf node, and the conditions along the path form a conjunction in the if clause. However, as its outcome quality has a high dependency on training data set distribution among different categories, and reconciliation data is a highly skewed sampling set, a decision tree's performance is not expected to be high in raw format (i.e., without any pre-processing to resolve the data skewness challenge). We validated this hypothesis in the following experiment.

As the current reconciliation system's rules are mostly matching rules, the overwhelming majority of the records in the initial training data set have a "matched" label: 139,092 records are labeled matched, and three records are unmatched. We added 120 unmatched records into the data set to highlight the exceptional part, but the ratio between matched and unmatched is still 1159:1, which is highly unbalanced (see Figure 5).

The decision tree in the figure contains 51 rules (25 of which are good), average rule depth is 15.51, *Con(R)* = 2.73, and *Rel(R)* = 1.2857. As the break sample data is too sparse, the forecast failure rate *Rel(R)* is not reliable and is not calculated. From this result we can tell that a highly unbalanced training set yields too many rules with a high level of complexity (high rule depth and high *Con(R)*) and a low level of decisiveness/reliability (low *Rel(R)*).

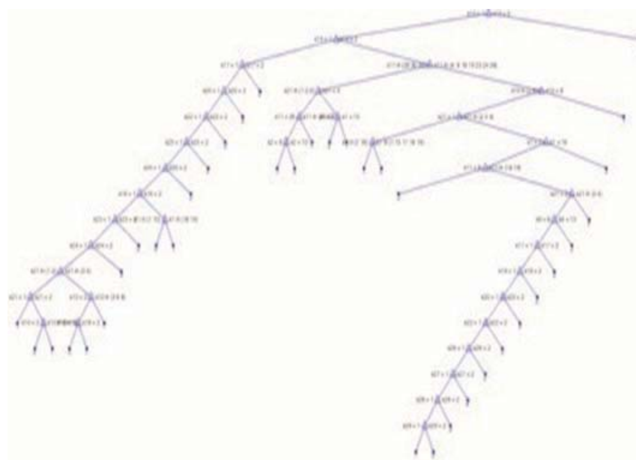


Figure 6a — Positive data reduced to 10%.

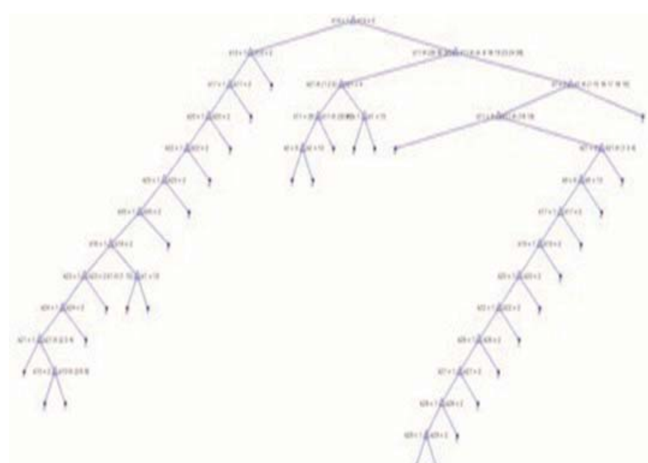


Figure 6b — Positive data reduced to 1%.

Enhanced Decision Tree with Random Undersampling

To improve result set quality, the training data set must be undersampled to lead the data set to a better balance between positive data samples and negative data samples. We first tried to randomly pick positive sample data following certain ratios to reduce positive data to a certain level — 10%, 1%, or 1‰ (meaning to

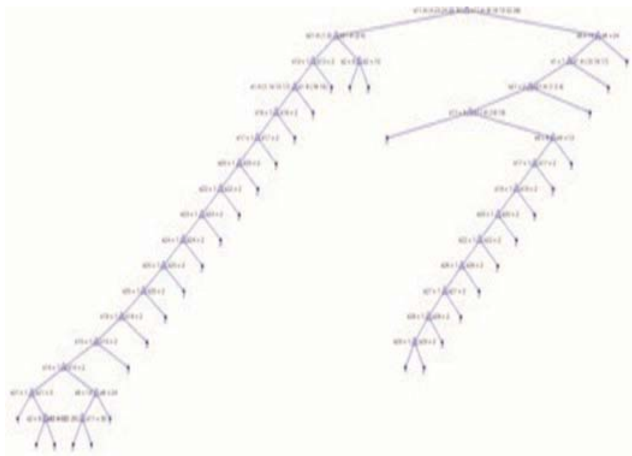


Figure 6c — Positive data reduced to 1‰.

randomly pick 1 positive sample from every 10 positive samples, 100 positive samples, or 1,000 positive samples, respectively). The decision tree algorithm was applied to each new data set (see Figures 6a-6c). Further criteria details are shown in Table 1, including the results from the previous raw decision tree test.

Looking at Figures 6a-6c, it's easy to tell that the 10% and 1% random sampling have less rule depth and fewer rules, although there are still more positive samples than negative ones. Remember that the original positive/negative ratio is 1159:1, so 10% undersampling means 115.9:1, and 1% undersampling means 11.59:1. However, we also observed that when the undersample ratio is 1‰ , the performance is worse — deeper rule levels, lower reliability, and higher forecast failure ratio. This suggests that when we undersample too much, a lot of useful data will be excluded.

Enhanced Decision Tree with kNN Undersampling

Based on our observations from the random undersampling described above, we want to achieve a proper balance between keeping too many positive samples (which overwhelms negative samples such that critical break rules are not discovered) and trimming off too many positive samples (with the result that fundamental match rules are not discovered). We introduced the kNN algorithm for this purpose.

In order to make sure all break rules are included, we use all negative samples as origin points and find kNN data points from positive samples to construct a balanced data set. Here the distance is calculated based on Euclidean distance using the previously mentioned 12 unilateral fields as coordinate values. Below is an indicative formula for the distance calculation:

$$\mathrm{d(x,y)} = \sqrt{\sum_{i=1}^{k} (x_i - x0_i)^2}$$

Here x_i stands for the positive sample's relative field value, while $x0_i$ stands for the origin negative sample's relative field value. Different k values lead to different positive/negative sample ratios. We used 3, 5, 8, and 30 k values (see Figures 7a-7d). Further criteria details are shown in Table 2.

Comparing the two tables, we can see that the kNN approach generally has less complexity, a smaller number of rules, higher reliability, and a very low forecast failure ratio. k=5 and k=8 have the same criteria values and better performance than k=3 and k=30. k=3 has the worst FFR(R), as it omitted too many samples. k=30 has bad reliability and concision because it includes too many positive cases, which take more weights in the final rule sets.

Table 3 compares the three mechanisms we covered in our experiment. It is easy to tell that for the unbalanced training data set, kNN undersampling has the best performance with proper k factor value settings. We also observed a common trend in k factor value tuning. Regardless of a data set's positive/negative ratio, when you start from k=1, the result rule set's quality increases as the k value increases, but once it hits a critical point, the result rule set's quality decreases slowly.

The Way Forward

In this article, we proposed an enhanced decision tree algorithm with kNN undersampling to resolve the problems that arise from an unbalanced training data set. Our experiment shows it performs well in terms of the result rule set's reliability, concision level, and forecast failure rate compared with a traditional decision tree algorithm. This means that with proper training data selection and pre-processing, machine learning could find reconciliation rules with good accuracy and rationality, which can reduce the time spent by domain experts in reconciliation data review and analysis. Ultimately, this could save effort and cost spent on reconciliation processing.

Sampling Ratio	Rule Number N	Con(R)	Rel(R)	FFR(R)
10%	37	1.58	3.0590	0.20%
1%	31	1.54	3.4130	0.22%
1‰	34	2.35	2.7034	2.58%
Raw Decision Tree	51	2.73	1.2857	N/A

Table 1 — Rule concision, reliability, and forecast failure rate for each data set.

K Factor Value	Rule Number N	Con(R)	Rel(R)	FFR(R)
3	35	1.89	1.9305	0.17%
5	36	1.46	3.5804	0.00%
8	36	1.46	3.5804	0.00%
30	71	2.73	1.2857	0.00%

Table 2 — kNN extension benchmark.

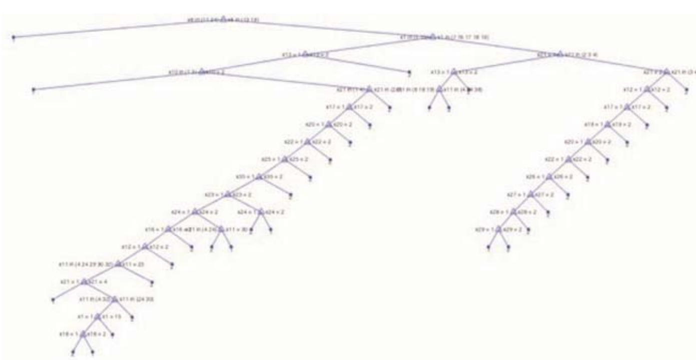


Figure 7a — Decision tree when k=3.

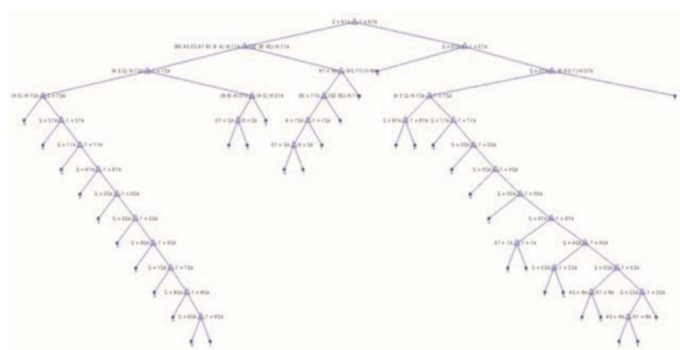


Figure 7b — Decision tree when k=5.

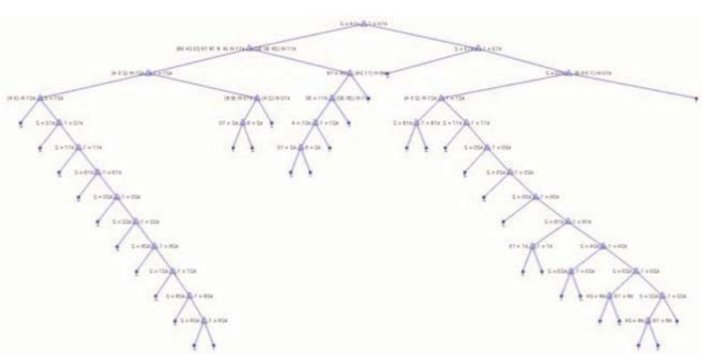


Figure 7c — Decision tree when k=8.

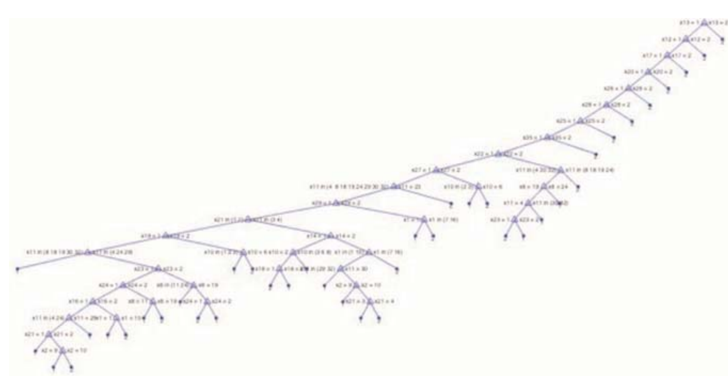


Figure 7d — Decision tree when k=30.

Mechanism	Rule Number N	Con(R)	Rel(R)	FFR(R)
Origin Decision Tree	51	2.73	1.2857	N/A
Random Undersampling (1%)	31	1.54	3.4130	0.22%
kNN Undersampling (k=8)	36	1.46	3.5804	0.00%

Table 3 — Experiment results benchmark.

However, our whole experiment happened in a supervised environment where all feature columns were defined and training data sets were labeled by domain experts beforehand. Furthermore, the training data was limited compared to the real volume in production. As a next step, we plan to address the problem of non-supervised rule learning with bigger data volume and more transaction types. Among other objectives, we hope to:

- Find an efficient approach to labeling massive training data sets (of at least a weekly or monthly time window) to achieve the most accurate results with the least manual intervention.
- Design an efficient approach to identify feature columns across different transaction types (data patterns).

We hope the non-supervised approach will further improve the rule-mining process by reducing the dependency and demand on domain experts' time and efforts, while keeping the outcome rules at a high level of quality regarding accuracy and completeness.

Endnotes

[1]"Reconciliation." Investopedia (www.investopedia.com/terms/r/reconciliation.asp).

[2]Larose, Daniel T., and Chantal D. Larose. "*k*-Nearest Neighbor Algorithm." Chap. 7 in *Discovering Knowledge in Data: An Introduction to Data Mining*. 2nd edition. John Wiley & Sons, 2014.

[3]Safavian, S. Rasoul, and David Landgrebe. "A Survey of Decision Tree Classifier Methodology." NASA Technical Report NASA-CR-188208, September 1990 (https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19910014493.pdf).

[4]Xiaohu Liu, S.L. "Optimization in Decision Tree Algorithm" (in Chinese). *Journal of Software*, No. 9, 1998.

[5]Dominic, Stephen, et al. "Genetic Reinforcement Learning for Neural Networks." *Proceedings of the International Joint Conference on Neural Networks*. IEEE, 1991 (http://ieeexplore.ieee.org/document/155315/).

[6]Cortes, Corinna, and Vladimir Vapnik. "Support-Vector Networks." *Machine Learning*, Vol. 20, No. 3, September 1995 (http://dl.acm.org/citation.cfm?id=218929).

[7]McCallum, Andrew, and Kamal Nigam. "A Comparison of Event Models for Naive Bayes Text Classification." *Proceedings of the AAAI-98 Workshop on Learning for Text Categorization*. AAAI, 1998 (http://www.cs.cmu.edu/~knigam/papers/multinomial-aaaiws98.pdf).

[8]Quinlan, John Ross. "Simplifying Decision Trees." *International Journal of Man-Machine Studies*, Vol. 27, No. 3, September 1987 (http://dl.acm.org/citation.cfm?id=50008).

Zhuo Li is Managing Director in State Street Corporation's Application Development and Maintenance Department. Dr. Li has more than 15 years' experience in financial technology R&D focusing on transaction processing, data mining, and visualization. He holds a PhD in computer science and an MS degree in management engineering from Zhejiang University. He can be reached at Zhuo_Li@statestreet.com.

Jianling Sun is a Professor and Doctoral Supervisor at Zhejiang University, with more than 30 years' experience in advanced technology R&D. His recent research focuses are data mining, machine learning, and blockchain. He holds a PhD in computer science. He can be reached at sunjl@zju.edu.cn.

About Cutter Consortium

Cutter Consortium is a unique, global business technology advisory firm dedicated to helping organizations leverage emerging technologies and the latest business management thinking to achieve competitive advantage and mission success. Through its research, training, executive education, and consulting, Cutter enables digital transformation.

Cutter helps clients address the spectrum of challenges technology change brings — from disruption of business models and the sustainable innovation, change management, and leadership a new order demands, to the creation, implementation, and optimization of software and systems that power newly holistic enterprise and business unit strategies.

Cutter pushes the thinking in the field by fostering debate and collaboration among its global community of thought leaders. Coupled with its famously objective "no ties to vendors" policy, Cutter's *Access to the Experts* approach delivers cutting-edge, objective information and innovative solutions to its clients worldwide.

For more information, visit www.cutter.com or call us at +1 781 648 8700.