Sales-force performance analytics and optimization

M. Baier
J. E. Carballo
A. J. Chang
Y. Lu
A. Mojsilović
M. J. Richard
M. Singh
M. S. Squillante
K. R. Varshney

We describe a quantitative analytics and optimization methodology designed to improve the efficiency and productivity of the IBM global sales force. This methodology is implemented and deployed via three company-wide initiatives, namely the Growth and Performance (GAP) program, the Territory Optimization Program (TOP), and the Coverage Optimization with Profitability (COP) initiative. GAP provides a set of analytical models to measure and optimize sales capacity and profitable sales growth. TOP develops a set of analytical models and methods for the analysis and optimization of assigning customers to sellers and other sales channels. COP provides additional recommendations on sales-coverage adjustment on the basis of an improved estimation of customer profit. We discuss these three programs in detail and describe how they work together to provide an analytics-driven transformation of the IBM global sales force to improve various sales metrics, such as revenue and cost.

Introduction

Transformation into a more effective and efficient enterprise often primarily concerns fundamental change to the business operations of an organization. The goal of such business transformation is to achieve broad alignment of activities related to people, processes, and technology with a more intelligent strategy and vision in order to satisfy long-term business objectives. As organizations transform and evolve over time, they must continually identify new ways to improve overall performance and deliver additional value to their customers. Advanced analytics, embedded into the core processes of an organization, are critical to achieving these transformational goals and providing the organization with deep insights and significant competitive advantage in the marketplace. Studies show that highly effective and efficient organizations use analytics for business transformation far more often than those exhibiting lower effectiveness and efficiencies (for example, see [1]).

IBM has committed to the use of analytics to enable enterprise-wide transformations in pursuit of ever improving performance, which in the context of this paper generally refers to profitable sales growth. With particular focus on increasing global sales-force productivity, three company-wide initiatives—Growth and Performance (GAP), Territory Optimization Program (TOP), and Coverage Optimization with Profitability (COP)—were created. The models and methods of each initiative are built upon a set of advanced analytics, some of which significantly expand upon previous mathematical techniques. In this paper, we present these initiatives and their impact on business transformation at IBM.

Sales performance improvement is an essential contributor to profitable revenue growth and a key enabler of the IBM business strategy [2]. To this end, the GAP initiative has developed models to comprehensively analyze and manage sales performance. GAP models can generate business insights to identify root causes for performance challenges (e.g., a distribution of skills that is less effective), compare performance across geographical regions, and project future performance on the basis of sales resource assumptions, where sales resources include alternative sales channels (e.g., inside sales) in addition to individual sellers. These models have been integrated into the Business Units, such as the Hardware, Software, and Services Units of IBM, to transform the way sales productivity is managed. Based on the insights driven from the GAP models, one current focus area for sales-force performance improvement is around optimizing the distribution of sales resources by Business Unit and geography. The ability to optimize the distribution of sales resources for a given set of clients and available opportunity can lead to sales performance gains and

Digital Object Identifier: 10.1147/JRD.2012.2216314

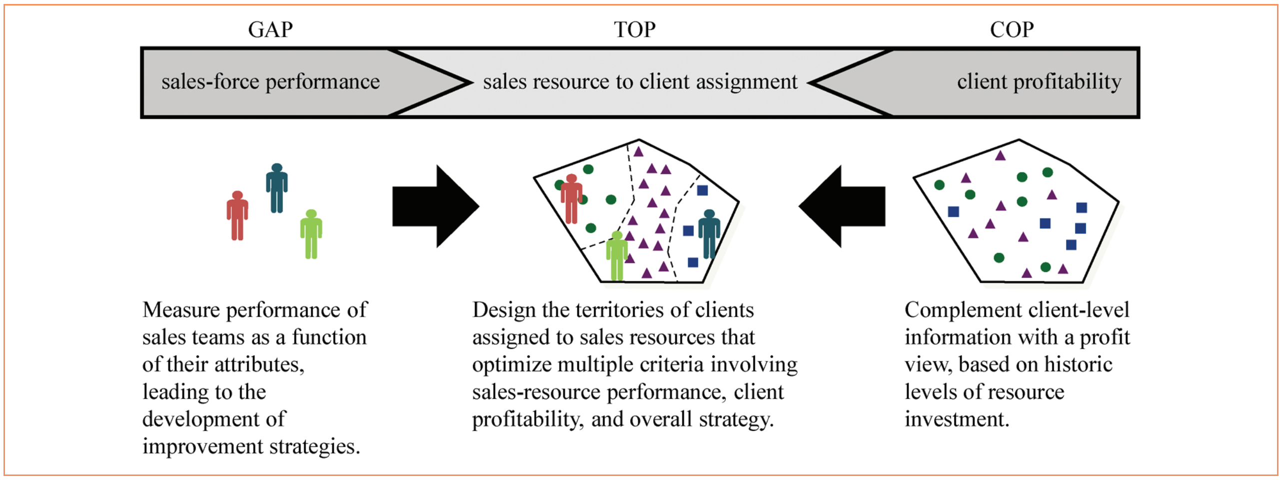

Figure 1

Integration of GAP, TOP, and COP within the IBM Sales Model to improve profitability and revenue growth.

profitable revenue growth, where a *client* refers to any business entity to which IBM sells its products typically not including an individual consumer.

Another key factor in increasing revenue and improving sales performance concerns the sets of clients assigned to sales resources, where a collection of clients assigned to a sales resource is defined to be a *territory*. Extensive research indicates that creating higher-quality sales territories, based on quantitative metrics for scoring territories according to various sales evaluation criteria, is one established path to increased sales-force productivity (e.g., see [3]). To address the challenges of composing and allocating such high-quality territories, the TOP initiative has developed models and methods for the comprehensive analysis and optimization of the quality of sales territories. TOP models render business insights that facilitate better client assignments. Increased productivity is achieved by providing the means to quantitatively assess the quality of current and alternative territories. These models then form the basis of an optimization capability to determine the highest-quality territories that maximize revenue growth and sales performance across Business Units. To address the complexities and interdependencies of territory optimization in a large corporation with correlated Business Units, the TOP analytics are deployed as part of an integrated process that reoccurs on a semiannual schedule to provide information and support for the corporate sales planning cycle.

Yet another important input to consider for improving sales-force productivity is an estimate of future revenue opportunity at each of the IBM client accounts, such as that provided by the Market Alignment Program (MAP) [4]. COP builds on MAP and other existing IBM programs. The existence of such estimated revenue opportunity, however, does not guarantee actual future revenue, and furthermore, the historic levels of sales-effort investment for capturing such estimated opportunities have not always been optimal. In addition to the use of historical revenue data and future revenue opportunity forecasts for deciding on sales-force deployment, COP provides a complementary view based on profit, also at each of the IBM clients, in order to propose sales-force coverage adjustments by the identification of clients for which the investment in sales effort has not been justified, or where an increase in investment could be justified, on the basis of observing relative levels of revenue, opportunity, and profit.

Figure 1 depicts the primary focus of GAP, TOP, and COP and the interrelationships among them. In addition to the insights directly offered by each, GAP and COP provide important inputs to TOP with respect to seller and client characteristics, respectively, as indicated by the arrows in the figure. The small circles, triangles, and squares in Figure 1 represent various clients and their respective attributes, details of which are provided to TOP by COP. On the basis of these client attributes and the attributes of sales resources provided by GAP, TOP composes three territories consisting of the small circles, triangles, and squares and assigns each to one of the three available sellers.

To achieve the overall goal of increasing global sales-force productivity and improving profitable sales growth, advanced analytics have been developed to provide the foundation for the solution of each of the three company-wide initiatives. In analyzing and managing sales performance in GAP, we first model the sales capacity of the sales force using not just simple headcount, but also using individual seller ramp-up and ramp-down profiles (i.e., profiles involving information on the time required to reach or drop from full capacity) that incorporate human resource transitions (i.e., sellers moving in and out of job roles) [5, 6]. With sales productivity defined on the basis of the sales capacity model, we develop regression models to identify factors that affect performance. The regression models thus suggest avenues into the problem of sales-force optimization, which requires careful consideration of several variables. First, the proper metrics for optimization are selected, depending on the desired outcome. Second, the appropriate segmentation of analysis is determined, aligning to the level in the organization for which sales-resource decisions and assignments are made. Finally, since all geographies and clients have idiosyncrasies with different performance expectations and sales-resource requirements, we perform analysis for different groups of clients separately.

From an advanced analytics perspective, sales territories consist of disjoint partitions of the set of clients that are assigned to sales resources. The TOP problem, then, consists of determining the best allocation of clients to sales resources according to a set of evaluation criteria, where the latter is a formal representation of a comprehensive set of business principles, sales coverage strategies, and Business Unit growth initiatives to evaluate territory quality. This includes an iterative feedback loop that supports the continual refinement of evaluation criteria on the basis of business outcomes and growth opportunities. Previous solutions to the territory optimization problem have been limited to very few criteria and deterministic formulations such that the resulting linear integer program can be efficiently computed with standard optimization packages [3]. In contrast, TOP includes a more detailed and rigorous evaluation criteria, together with associated uncertainties and risks. The resulting stochastic nonlinear dynamic program is far too large and complex to be solved in an efficient manner with existing methods. Hence, on the basis of probabilistic properties of the underlying multidimensional stochastic process, we derive an efficient approximation algorithm whose solutions are proven to be nearly optimal.

One of the main technical challenges of advanced analytics for COP is establishing a reliable estimate of sales expense (i.e., the cost of the sales force directly linked to each of the IBM clients) and, with this, an estimate of profit for each client. We develop a simple analytical approach to estimate such expense on the basis of observed sales activity recorded in so-called *client relationship management* (CRM) tools. Subsequently, a simplified framework for combining historic profitability with revenue, historic revenue growth, and revenue opportunity is presented, which permits the identification of clients where qualitative adjustments of seller coverage can be proposed. We also outline a more advanced quantitative approach that seeks to estimate the optimal level of seller deployment adjustment with respect to maximizing profit on the basis of the estimated revenue opportunity by learning the appropriate functional forms that best describe the relationship between revenue and sales expense for various segments of clients.

Over the past few years, all three company-wide initiatives have been deployed to provide a set of integrated solutions for analytics-driven transformation within IBM. Starting in 2007, GAP analytics have been deployed across the various IBM Business Units and geographies, influencing key decisions from optimizing sales-resource investments to determining sales improvement strategies. Since the initial introduction of territory optimization at the end of 2008, TOP analytics have been similarly deployed throughout IBM to increase the sales-force productivity and improve the sales performance of more than 17,000 resources (i.e., sales representatives and other IBM channels), with territory optimization benefits expected to yield significant improvements in sales revenue annually. Starting as a pilot for a single Business Unit (IBM Software Group) and region (Italy) in late 2010, COP analytics have been used to produce worldwide results for IBM Software Group throughout 2011 with usage expected to continue to grow as COP becomes fully embedded in the sales process.

Each of the GAP, TOP, and COP initiatives has directly provided key insights and significant benefits to their respective sponsors and stakeholders. In addition, all three initiatives are integral to the IBM Sales Model—the corporate framework for embedding analytics into the sales-related strategic decision-making process, from opportunity identification to efficient sales resource deployment and incentive provisions to increase productivity and improve performance. Its main components occur on a semiannual schedule to provide information for the two major sales planning phases of the second and fourth quarter. An example is the tripartite process of TOP that takes into consideration both client account and sales-resource attributes. First, current territory qualities are assessed and actions for improvement are identified. Next, deduced modifications for optimization are processed. Finally, the review of outcomes concludes the process. GAP similarly contributes to the overall process by providing improvements through better sales-force development and attributes (e.g., capacity insights), while COP provides analogous improvements through better client account attributes (e.g., profitability insights). Alongside this globally standardized process, GAP, TOP, and COP further enable customized projects by request, such as local investigations into specific sales productivity-related issues.

The remainder of this paper is organized as follows. We first present a more detailed description of the technical problems and mathematical solutions for each of the three initiatives. Next, we discuss some of our experiences with the analytics-driven transformation of IBM based on these initiatives. Finally, we close with some concluding remarks.

Technical problems and analytical solutions

Growth and performance

GAP is associated with three main goals: 1) to develop models to understand historical sales performance, 2) to model and identify actionable factors that affect performance; and 3) to develop techniques for improving performance by affecting such factors. These goals are in line with the marketing literature on improving sales performance and productivity [7].

One measure of sales performance used is sales productivity, defined as the revenue per unit of sales capacity, where the sales capacity is an adjusted headcount. In any typical organization, there is considerable movement within the sales force, with sellers being hired, transferring, or leaving. Moreover, sellers have different ramp-up and ramp-down profiles after such moves. For example, a new university hire has an initial period of little or no sales capacity. During this period, the seller learns how to do a particular job or activity, which is followed by a period of gradually increasing capacity until reaching an average, steady-state capacity level. Although the headcount change is immediate, the capacity changes gradually. Therefore, to obtain an accurate measure of sales performance, it is essential to measure productivity as a function of actual sales capacity, rather than headcount, at any given time.

To understand historical sales productivity and revenue performance, GAP uses regression-based sales capacity and productivity models created as follows. First, sales capacity with respect to time is calculated using historical headcount, headcount dynamics (i.e., movement of sellers in and out of job roles), and ramp-up profiles. Ramp-up profiles quantify the post-event capacity of sellers and were developed via interviews with sales leaders. We are currently exploring the learning of such profiles from historical data by adopting a regularized linear programming formulation [8]. In addition, intrinsic, historical sales productivity and productivity growth are calculated by fitting an exponential curve to the sales productivity data with respect to time, after smoothing the seasonality of buying behavior in observed productivity using rolling averages.

Adjusting for seasonal trends and taking a longer time period view allows a truer picture of the underlying intrinsic productivity, and correspondingly sales-force performance, to emerge. Moreover, it allows for the development of regression models to predict future performance on the basis of historical and projected sales capacity (adjusted for productivity growth) and revenue. This, in turn, enables sales leaders to make accurate revenue projections to determine gaps for reaching targets, run scenarios to assess optimal actions needed to meet targets, and test viability of plans/targets relative to different assumptions (e.g., productivity improvements).

While productivity is a commonly used performance measure, it is not always appropriate. For example, it may only reflect the effects of a few large deals and mask a substantial number of lost opportunities. Similarly, revenue may be a less appropriate performance indicator in a growth market compared with number of clients acquired. Therefore, GAP includes other performance metrics, such as number of opportunities resulting in actual sales, average duration of opportunities, and number of opportunities identified. These performance metrics are at the seller, transactional, and aggregate (geography and Business Unit) levels. Regression models are also developed to determine the effect of various factors on these performance metrics, details of which are provided in reference [6].

The capacity and productivity models, along with the other performance metrics and models, are used by sales executives to identify geographies and Business Units for which performance challenges exist and provide insight into which factors have the largest effect. One factor that has emerged from such analysis and that has also been suggested independently by sales executives as a major determiner of sales performance is the composition of the sales force. Consequently, a current focus of GAP is to determine the optimal team composition within each sales geography and Business Unit. Team composition refers to the mix of sales resources (along with their characteristics and attributes) that together interact with a client.

Because only a small number of geographies and Business Units exist, there is insufficient data at this level for a rich and sound analysis. Hence, we try to answer this question at the client level and aggregate results to the higher level desired. The team composition analysis is performed statistically and does not take ramp-up dynamics into account because of data limitations; the available interaction and revenue data at the client level is over half-year time periods, not at a finer time scale that would permit dynamic analysis.

Two broad categories of sales job roles are considered, technical and nontechnical, to determine the ideal composition of a sales team that leads to high sales performance for a particular client. The answer may be different in different classifications of clients, including classifications based on geography, Business Unit, segment, and sector.

The first step is to define the sales team composition. The effort expended on a client is measured using full-time equivalents (FTEs) of sellers. For example, if seller A spent a quarter of her total effort on client i (with the remainder spent servicing other clients), seller B spent a fifth of his total effort on client i, seller C spent half of her total effort on client i, and no other sellers serviced client i, then the total effort expended by the sales force on client i is $0.25 + 0.2 + 0.5 = 0.95$ FTEs. Furthermore, if sellers A and B have technical job roles and seller C has a nontechnical job role, then the sales team composition at client i measured using the ratio of technical-to-nontechnical effort is $(0.25 + 0.2)/0.5 = 0.9$. Denote this team composition as $t_i \in [0, \infty)$, for all clients $i = 1, \ldots, n$.

The second step is to define a sales performance metric. One way to examine sales performance is through transactional revenue. However, this obscures the effect of the sales team because clients would produce revenue in a certain range with any sales team. The effect of the sales team is in modulating the point in the revenue range that is actually achieved. This general range of a client is captured through its aspirational revenue, an expected value of actual transactional revenue that may be achieved. Therefore, sales team performance is better revealed through the ratio of actual transactional revenue to aspirational revenue. Let this ratio for client i be $y_i \in \mathbb{R}$.

Thus, we have composition-performance pairs (t_i, y_i), $i = 1, \ldots, n$. We would like to understand the functional relationship, mapping t to y in order to develop guidelines for the sales organizations of various Business Units and geographies. We estimate this functional relationship using the Nadaraya-Watson estimator, a form of kernel regression [9, 10]. Other estimators from the applied data smoothing literature may also be used. Through this analysis, the team composition t that maximizes the estimate $\hat{y}(t)$ is a suggested team composition to achieve the best positive modulation in transactional revenue. Of course, this recommended team composition is subject to constraints induced by the actual job role staffing of the sales force. Importantly, not all clients respond to team composition in the same way.

Different groups of clients have different behavior in response to sales team composition. For example, a sales team with composition $t = 1.4$ may have better performance with clients in the industrial sector than clients in the public sector. Therefore, composition-performance functions $\hat{y}(t)$ are estimated for different groups of clients separately. This is easily achieved by using only the data points of the appropriate subset of clients in the estimate. To obtain a composition-performance function for North American clients in a specific segment, the (t_i, y_i) pairs only for those clients are utilized. Furthermore, rather than pre-specifying the groupings to consider, we can automatically discover the specification (based on features such as geography, sector, and coverage segment) of client groups with similar composition-performance behavior [11].

Territory optimization program

To address the question of territory quality, an initial element of any territory alignment process usually consists of determining the criteria for the evaluation of sales territories. This often involves four classes of objectives: balancing workload, balancing sales potential, minimizing disruption of relationships, and minimizing travel time and costs [3]. The TOP evaluation criteria include all four objectives with the appropriate levels of detail to capture the trade-offs within and among these objectives. In addition, TOP incorporates client attributes (e.g., industry, historical revenue, forecasts of revenue opportunities, coverage segments) and sales resource attributes (e.g., individual productivity, served channels, skills, areas of expertise) into the evaluation criteria. Interdependencies among paired client and resource attributes manifest further criteria, such as geographical considerations to reflect travel time and costs to cover a sales territory. Taking into consideration the unique requirements of each IBM Business Unit, the TOP criteria are fine-tuned for each of the different Business Units. Finally, criteria that consider the cooperation of multiple Business Units reflect measurements of efforts towards a coherent unified sales approach employed throughout IBM.

Each of the evaluation criteria represents a dimension of the territory alignment process. An initial set of dimensional attributes was developed based on pre-existing and independently defined sales coverage strategies. For example, the attempt of a Business Unit to increase client value through industry expertise can be measured by the number of distinct industries covered within the territory of a sales resource. The effectiveness of implementing this initial set of quantitative translations from business principles to dimensional attributes of territory quality was evaluated with respect to strategy alignment, which in turn has led to further refinement and augmentation of the dimensions of the territory alignment process over time. This evaluation and refinement of attributes includes periodic correlation analyses of individual resource productivity metrics (e.g., quota achievement) with territory quality (i.e., evaluation criteria representing dimensional attributes, such as revenue forecasts). One important aspect of this process has been an iterative feedback loop that provides information to facilitate the development of coverage strategies based on productivity findings at an individual resource granularity.

The next step in the TOP process is to optimize the composition and assignment of sales territories. More formally, the objective of territory optimization is to allocate multi-attribute clients among multi-attribute sales resources in order to optimize the evaluation scores taken over each dimension, $d = 1, 2, \ldots, D$, of the overall problem. The scoring of each dimension d is modeled by a utility function, f_d, based on Business Unit-specific criteria (e.g., revenue, cost, or profit). The utility function f_d severely penalizes any significant deviations from a range of reasonable values to reflect business preferences, and, in practice, f_d is a piecewise linear and quadratic function. The result is nonlinear functional forms for each dimension that are combined together through weights, w_d, assigned to each dimension. Then, the goal of the mathematical formulation is to determine an allocation that maximizes the weighted combination of utility functions, $\sum_{d=1}^{D} w_d f_d$, which renders an optimization problem falling within the realm of nonlinear mixed integer programs (because allocation of clients cannot be fractional). Solving even modest-size instances of such problems requires tremendous computing resources, which makes an exact solution of the IBM problem computationally prohibitive.

Our approach is consistent with Skiera and Albers [12] in that the objective function is designed to simultaneously include multiple considerations, such as profit maximization and fairness. However, Skiera and Albers consider a much simpler problem by modeling profit as a concave function of selling times that are allocated to sales resources, which results in a formulation solely of profit maximization as a convex program with integer variables. In stark contrast, our formulation includes simultaneous consideration of many more dimensions, each of which has more complicated structures than selling times. Hence, our objective function is not equipped with second-order monotonicity properties such as convexity or modularity. This in turn allows great flexibility in modeling the complex structures of each dimensional attribute and the multidimensional relationships among these evaluation criteria that arise within the IBM Sales Model. Moreover, our solution is much more efficient, which was a design requirement to support an interactive tool.

Furthermore, there are uncertainties and risks associated with many of the problem dimensions, something not considered in reference [12]. For example, the revenue forecast has considerable uncertainties related to client behavior, market conditions, and so on. Similarly, the travel time and cost of a sales resource has considerable uncertainties related to client characteristics, geographical distance, and the like. Risks are also associated with dimensions where investments are made, and the possibilities of perceived benefits may not be realized. In such situations where the uncertainties and risks are non-negligible, considering only average values for each dimension will render poor representations of the actual utility function values, large discrepancies among different territories, and poor alignment of sales resources. To cope with these difficulties, the utility functions f_d include characterizations of the uncertainties and risks associated with the dimensions d. In addition, the objective function of our optimization problem can be adjusted to obtain solutions that hedge against associated risks (see, for example, [13] for basic details on hedging).

Given the complexities related to both computation and uncertainty/risk, we use an approach that maps the original stochastic nonlinear mixed integer programming problem to a corresponding multiple, multi-dimensional stochastic knapsacks problem. Here, each knapsack corresponds to a sales territory, the multi-dimensional aspects correspond to the multiple dimensions of territory evaluation scoring, and the stochastic aspects correspond to the uncertainties and risks associated with each dimension. Different forms of related stochastic scheduling problems have received considerable attention in the research literature (e.g., [14]), as have several different instances of the stochastic knapsack problem (e.g., [15–17]).

To the best of our knowledge, virtually all of the previous work on stochastic knapsack problems has considered a single or one-dimensional knapsack. However, since territory optimization requires the allocation of multiple clients among multiple sales resources, both with multiple attributes, we investigate a significantly more general class of stochastic knapsack problems than has been previously considered. Each of the items available for allocation and each of the knapsacks into which the items can be placed consist of multiple dimensions. Here, the size of an item for each dimension is drawn from a general probability distribution, and the size of each dimension is fixed for every knapsack. The item size is realized only after an attempt has been made to allocate the item into one of the knapsacks. A value is associated with each item, which is realized upon successfully allocating the item into one of the knapsacks without capacity overflow. The value of the knapsack is a function of the summation of the values of all the items that have been allocated into the knapsack multiplied by a weight that is unique to the knapsack. The overall value of an allocation is the summation of the values of all knapsacks under the allocation. We consider a combination of strategies for sequentially allocating a set of items among a collection of knapsacks with respect to maximizing a functional of the overall value of the allocation, and establish fundamental properties of these strategies.

The optimization algorithm we implemented in TOP is a careful combination of both non-adaptive and adaptive strategies. In the non-adaptive (greedy search) step, each client is examined sequentially with respect to its impact on the score of existing territory assignments based on the change of the overall utility function for each possible allocation, and then the client is assigned to the territory that realizes the greatest overall improvement. Once the set of clients is allocated through this non-adaptive step, the subsequent adaptive (local search) step selects a small group of clients across different territories and explores whether the current assignment can be improved by solely adjusting the allocation of this selected group. This adaptive step is repeated until a tolerance constraint is satisfied. Based on numerous numerical experiments, we observe that our algorithm returns near-optimal results within a second, using a standard laptop computer, for typical IBM instances of the problem consisting of many thousands of variables and constraints. More importantly, we use optimal stopping of martingales and other probabilistic analysis to explore the structures of the underlying stochastic processes. We then demonstrate that the algorithm is a polynomial approximation for the original stochastic nonlinear mixed integer programming problem, and the solution provided by our algorithm is within a constant ratio of the optimal allocation. Additional details can be found in [18].

Coverage optimization with profitability

At the heart of COP is the estimation of client-level *profit*. In the context of COP, profit is defined as the *revenue* that the client generates, minus *costs* (defined here as costs of goods sold, including installation and other direct costs that are accounted for and linked to client-specific revenue) and minus *expense* (defined here as the expense of sales activities, including seller salary, commissions, and other payments that are part of direct "Sales, General, and Administrative" expenses), for a given time period, typically one year. Expense is not currently allocated at the client level in IBM financial systems, since there is no standardized company-wide formal tracking of the time a seller has spent with clients. Many tracking tools do exist and are in use throughout the company, but their main purpose is not to allocate and track the time a seller has spent with clients, but rather to keep records of seller-client interactions for the purpose of managing the relationship.

In order to estimate sales expense, we have re-purposed the data from a CRM tool used widely in IBM, which provides client overviews of each stage of the sales process. By maintaining information for each of these stages, the sales opportunities are tracked as they evolve from being noticed, to being validated, and ultimately to being converted into a "won" sale (and subsequently considered "fulfilled") or "lost" sale. Only time spent by sellers in true sales stages (even for those resulting in "lost" sales) are considered in estimating expense; their time spent fulfilling won opportunities are recorded as costs and these are allocated to clients.

The allocation of expense is based on distributing the time a seller has spent on any and all clients that the seller participated with during any of the sales stages. In summary, the time of a seller is distributed to clients based on the sales opportunities that they are linked with, as shown in the CRM records (for any client, there may be multiple sales opportunities, and one or more sellers linked to these). The entirety of the time of the seller (equal to one FTE) is allocated to opportunities on a fractional basis, such that her FTE allocation to all opportunities for all clients sums to one. The allocation of fractional FTEs is directly proportional to the value and the duration of the opportunity, but inversely proportional to the size of the sales team linked to it. For each client, the sum of all allocated FTEs represents the invested sales effort in the period being analyzed. To finalize the profit calculation, the client-level expense is estimated from the average expense per seller multiplied by the sum of its allocated FTEs. Subsequently, a value of client profitability is calculated. In order to account for normal fluctuations in revenue, for example, stemming from cyclical patterns in information technology expenditure, we use a multi-year profit estimate to fairly compare profit levels across clients.

The goal of COP is to further improve resource allocation decisions by *complementing* the current IBM usage of historic revenue and estimated future revenue opportunity data [4] with a new dimension, namely, client-level profit estimation. A very pragmatic framework was developed for identifying clients that can be considered as candidates for sales coverage adjustment, based on the relative levels along variations of these three dimensions. For each client, values are calculated for profit margin (defined as client profit normalized by revenue), as well as for historic revenue and opportunity growth. Here, historic revenue growth is defined by an estimate of a representative growth rate for a given period. Opportunity growth refers to the required growth rate for achieving the estimated revenue opportunity from current revenue levels.

Depending on the combination of the relative levels of each dimension value for any client, a qualitative resource adjustment action is proposed. For example, the framework identifies clients as candidates for adding resources if all dimension values are "high" or as candidates for reducing resources if all dimension values are "low." The setting of appropriate thresholds is done by trial and error in order to obtain reasonable sets of clients for adjustment of coverage. Note that there may be other factors to consider, which were not modeled in this approach, that might override a decision to increase or reduce resource levels assigned to a client, such as strategic coverage decisions for "protecting" a client from a competitive presence. Also, note that this approach is qualitative in nature, since the recommended level for adding or reducing resources is not provided. In addition, in its current form, the framework does not distinguish between resource type (for example, technical versus nontechnical sellers), and treats all sellers the same.

The qualitative approach for leveraging profitability data in assisting sales coverage optimization described above lacks the numerical completeness and rigor for truly *optimizing* resource allocation. We are currently exploring several alternatives. One approach that has shown promising results is based on two steps. First, the data are clustered into numerically homogeneous groups, in order to find the best local fit of an appropriate functional form of expense vs. revenue. Second, for each client in each data cluster, we find the optimal level of expense that maximizes future profit (capping predicted revenue as a function of expense with estimated revenue opportunity) and identify the proper resource adjustment necessary (i.e., either increasing or decreasing seller expense) for such level of maximum profit.

An exponential relationship between revenue and expense approximates the "saturating" behavior of diminishing returns for increasing levels of expense. As cost tends to increase linearly with revenue, the corresponding profit function for the cluster will likely have a convex maximum in the region of interest, which represents the optimal level of expense. This optimal level may be lower than the predicted level required for achieving an expected revenue value.

Many refinements to this approach need to be explored, including, but not limited to, identifying appropriate clustering algorithms, automatically identifying cluster-specific improved functional forms, distinguishing between seller types, and subsequently proposing optimal seller team composition and size. Of particular interest is to apply/extend the GAP approach to team composition to optimize with respect to client profitability and profit margin, rather than just transactional revenue.

Experience

The combination of GAP, TOP, and COP analytics is being used extensively to help guide strategic decisions pertaining to the sales force across IBM. Specific areas that directly benefit from the use of these three aspects of the IBM Sales Model include sales resource planning, sales performance investigation, sales territory optimization, client profitability investigation, and client coverage optimization. During designated planning cycles, insights from GAP, TOP, and COP models and methods help guide planning decisions regarding sales resource allocation. The ability to project revenue based on sales productivity trends, expected Business Unit sales capacities, sales territory quality trends and expected client profitability allows for the modeling, analysis, and optimization of various sales resource plans. Taking into account these and other related factors, such as market opportunity, sales leaders use GAP, TOP, and COP analytics to make informed resource allocation decisions.

In particular, GAP models enable deep-dive investigations of business challenges by providing holistic analysis and management of sales performance. These deep-dive investigations occur across Business Units and geographies, each with a specific focus. One such deep-dive investigation focused on examining sales performance challenges in a particular Business Unit by using analytics to better understand root causes across several dimensions. GAP models were used to fairly measure sales-force productivity and examine the effects of various sales-force attributes on sales performance. Applying the aforementioned GAP models, it was determined that the sales force for this specific Business Unit had both a current sales productivity level and a long-term historical productivity trend lower than its peers. This challenge validated the need for further investigation of root causes. The impact of sales-force attributes such as seller skills, sales-force composition, and strategy for covering clients were examined in detail by utilizing GAP analytics. Preliminary findings indicated that the distribution of sales resources across key skill sets for this Business Unit yielded a higher proportion of sales resources with lower skill levels when compared to its peers. In addition, the analysis showed that adjustments to the sales-force composition along with an enhanced strategy for covering clients could result in improved sales performance for this Business Unit. Ultimately, the analysis led to a set of actionable recommendations to be implemented, measured, and managed over the coming months.

Since late 2009, the TOP initiative and optimization capabilities have been provided to operating units across IBM. After overcoming initial barriers and challenges of adaptation to some new concepts of optimization, stakeholders found the program to be essential in understanding and improving the territory assignments of a global sales force. Individual adjustments and local customization have proven to be key components of the success of the program, as diverse business organizations with heterogeneous priorities select dimensional attributes and weight their importance accordingly.

IBM-internal studies found that the described analytics-driven territory optimization increases seller productivity by an average of 7%, affecting the transactional revenue “in scope” (i.e., excluding non-transactional revenue streams such as software licenses or maintenance cost). Business Units have improved their average territory quality by 16% every optimization cycle during the first two years (measured on a qualitative scale, with 0 being the worst and 10 being the best score). Diminishing returns are experienced during repeated iterations of optimization, but the benefit still does not fall below a third of its initial level. This threshold is the result of ongoing external (e.g., clients and markets) and internal (e.g., reorganization) changes, which create a continuous flux in the attributes to be optimized.

Moreover, certain dimensional attributes exhibit a stronger correlation with sales performance than others; thus, certain attributes contribute less to the success of sales resources. Purity in the client segmentation of a territory (e.g., when a sales resource exclusively serves clients identified for competitive selling) as well as a low number of different industries (to drive industry expertise and, thus, client value) were continually identified to have the highest impact on the productivity of a sales resource. Attributes that reflect future states of analyzed dimensions (e.g., field-validated revenue forecasts instead of financial revenues) also are more strongly supported by statistical analyses, as planning happens for upcoming sales cycles and not current sales cycles.

As of late 2011, global COP results based on a three-year sales cycle (2008 to 2010) have been finalized for one of the Business Units of IBM and have been provided as additional input for 2012 coverage decisions for European as well as for growth market regions. So far, only the quantitative approach (based on resource adjustment action clustering of clients) is in use. Although the re-purposing of CRM data for allocation at the client level of sales expense has some precision problems, our results have provided sales teams with very useful additional information for resource deployment decision making, and have received positive feedback. Immediate practical uses of COP data include reprioritization of sales efforts and overall support of the sales territory design. The acquisition of results for additional Business Units is underway, and the additional data will allow strengthening of the quantitative approach outlined earlier.

In the future, the full realization of benefits around sales-force optimization requires that COP be fully embedded in the sales process, complementing the estimated client-level revenue history and future opportunity provided by mature IBM programs like MAP [4] with the estimated client-level historic profit as an important input for optimizing seller territories within TOP and further improvement of other resource deployment decisions. It is expected that COP will become fully embedded as part of the input of TOP via MAP, reaching maturity within several years.

Summary and conclusion

In this paper, we have described three business transformation initiatives integral to the IBM Sales Model, each created to improve various aspects of strategic sales decision making through advanced analytics. The IBM Sales Model has clearly benefited from senior executive support and sponsorship, which is essential for the success of large-scale business transformation. Its initiatives have been prioritized at the highest levels of the business, including senior sales executives, the chairman, and the chief executive officer.

Acknowledgments

We thank our fellow team members associated with the Growth and Performance, Territory Optimization Program, and Coverage Optimization with Profitability company-wide initiatives. We also thank the following executives for their support: Chid Apte, Dan Bennewitz, Brenda Dietrich, Martin Fleming, Shiva Kumar, Ross Mauri, and Linda Sanford. Finally, we acknowledge the important contributions of all IBM employees and executives involved in the IBM Sales Model, especially the invaluable support from the participating Business Units and geographies, Sales Resource Allocation community, Business Transformation/Information Technology organization, and other members of Business Performance Services and Research.

References

1. S. LaValle, M. Hopkins, E. Lesser, R. Shockley, and N. Kruschwitz. (2010). Analytics: The new path to value. *MIT Sloan Manage. Rev.* [Online]. *52*, pp. 1–24, Reprint 52180. Available: http://www-05.ibm.com/es/solutions/sis/download/IBV_Analytics_the_new_path_to_value_Nov_3_GBE03371-USEN-00.pdf
2. D. K. Taft. (2010, Dec.). IBM's six principles for successful enterprise IT transformation. *Eweek* [Online]. *27(21)*. Available: http://www.eweek.com/c/a/IT-Infrastructure/IBMs-Six-Principles-for-Successful-Enterprise-IT-Transformation-811411/
3. A. Zoltners and S. Lorimer, "Sales territory alignment: An overlooked productivity tool," *J. Pers. Selling Sales Manage.*, vol. 20, no. 3, pp. 139–150, 2000.
4. R. Lawrence, C. Perlich, S. Rosset, J. Arroyo, M. Callahan, J. M. Collins, A. Ershov, S. Feinzig, I. Khabibrakhmanov, S. Mahatma, M. Niemaszyk, and S. M. Weiss, "Analytics-driven solutions for customer targeting and sales-force allocation," *IBM Syst. J.*, vol. 46, no. 4, pp. 797–816, Oct. 2007.
5. J. D. Shaw, "Turnover rates and organizational performance," *Organ. Psychol. Rev.*, vol. 1, no. 3, pp. 187–213, Aug. 2011.
6. M. Singh, D. Bhattacharjya, L. Deleris, D. Katz-Rogozhnikov, M. Squillante, B. Ray, A. Mojsilovic, C. Barrera, D. Kakarania, J. Richard, A. Saha, and J. Fu, "The growth and performance diagnostics initiative: A multi-dimensional framework for sales performance analysis and management," *Service Sci.*, vol. 3, no. 1, pp. 82–98, 2011.
7. B. Skiera and S. Albers, "Prioritizing sales force decision areas for productivity improvements using a core sales response function," *J. Pers. Selling Sales Manage.*, vol. 28, no. 2, pp. 145–154, 2008.
8. K. R. Varshney, M. Singh, M. Sharma, and A. Mojsilovic, "Estimating post-event seller productivity profiles in dynamic sales organizations," in *Proc. IEEE Int. Conf. Data Mining—Workshop Data Mining Service*, Vancouver, BC, Canada, Dec. 2011, pp. 1191–1198.
9. E. A. Nadaraya, "On estimating regression," *Theory Probab. Appl.*, vol. 9, no. 1, pp. 141–142, 1964.
10. G. S. Watson, "Smooth regression analysis," *Sankhya Ser. A*, vol. 26, no. 4, pp. 359–372, Dec. 1964.
11. K. R. Varshney, M. Singh, and J. Wang, "Decision trees for heterogeneous dose–response signal analysis," in *Proc. IEEE Statist. Signal Process. Workshop*, Ann Arbor, MI, Aug. 2012, pp. 916–919.
12. B. Skiera and S. Albers, "COSTA: Contribution optimizing sales territory alignment," *Marketing Sci.*, vol. 17, no. 3, pp. 196–213, 1998.
13. M. Baxter and A. Rennie, *Financial Calculus: An Introduction to Derivative Pricing*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
14. M. L. Pinedo, *Scheduling: Theory, Algorithms and Systems,* 3rd ed. Berlin, Germany: Springer-Verlag, 2008.
15. J. Kleinberg, Y. Rabani, and E. Tardos, "Allocating bandwidth for bursty connections," in *Proc. ACM Symp. Theory Comput.*, El Paso, TX, 1997, pp. 664–673.
16. A. Goel and P. Indyk, "Stochastic load balancing and related problems," in *Proc. IEEE Symp. Found. Comput. Sci.*, 1999, pp. 579–586.
17. B. C. Dean, M. X. Goemans, and J. Vondrak, "Approximating the stochastic knapsack problem: The benefit of adaptivity," *Math. Oper. Res.*, vol. 33, no. 4, pp. 945–964, 2008.
18. Y. Lu and M. S. Squillante, "On approximations for multiple multidimensional stochastic knapsacks," *Perform. Eval. Rev.*, vol. 38, no. 2, pp. 45–47, Sep. 2010.

Received January 16, 2012; accepted for publication February 13, 2012

Moritz Baier *IBM Corporate Headquarters, Armonk, NY 10504 USA (moritz.baier@us.ibm.com)*. Mr. Baier is a Senior Strategy Consultant with the IBM Business Performance Services (BPS) Group, which develops, incubates, and deploys strategic growth and optimization initiatives across Business Units in IBM. He received a B.S. degree in computer science from Berlin School of Economics and Law in 2009. He subsequently joined BPS, where he has focused on developing and deploying global sales-force optimization initiatives, with a particular focus on leading the design and implementation of business intelligence solutions.

Jorge E. Carballo *IBM Corporate Headquarters, Enterprise On Demand, Armonk, NY 10504 USA (j.e.carballo@us.ibm.com)*. Dr. Carballo is a Senior Managing Consultant in the Business Performance Services Group, which identifies, develops, and deploys performance-focused initiatives across Business Units in IBM, based on innovative application of advanced analytics. He received a Ph.D. degree in structural engineering from Stanford University and an M. Eng. degree in structural engineering from the Universidad Nacional Autónoma de México. He joined IBM in 2007 and has led global strategic analytics-based initiatives in support of productivity improvement.

Alice J. Chang *IBM Corporate Headquarters, Armonk, NY 10504 USA (changa@us.ibm.com)*. Ms. Chang is a Senior Strategy Consultant with the Business Performance Services Group, which identifies, develops, and deploys strategic growth and optimization initiatives across Business Units in IBM. She received a B.A. degree in cognitive science with specialization in computational modeling from the University of California, Berkeley, in 2004 and an M.B.A. degree in strategic management and marketing from The Wharton School at the University of Pennsylvania in 2011. She subsequently joined Business Performance Services, where she has focused on global sales-force analytics programs.

Yingdong Lu *IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (yingdong@us.ibm.com)*. Dr. Lu received a B.S. degree in mathematics from Peking University and Ph.D. degree in operations research from Columbia University. He subsequently joined IBM at the Thomas J. Watson Research Center, where he has worked on stochastic processes and applications.

Aleksandra (Saška) Mojsilović *IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (aleksand@us.ibm.com)*. Dr. Mojsilovic manages the Predictive Modeling and Optimization Group in the Business Analytics and Mathematical Sciences Department at the Thomas J. Watson Research Center. She received her Ph.D. degree in electrical engineering from the University of Belgrade, Serbia, in 1997, and has worked at Bell Laboratories (1998 to 2000) and IBM Research (2000 to present). Her main research interests include multidimensional signal processing, pattern recognition, computer vision, and machine learning. She is an author of more than 70 publications, books, and book chapters and holds 14 patents. She received a number of awards for her work, including the IEEE Young Author Best Paper Award, European Conference on Computer Vision Best Paper Award, INFORMS Wagner Prize, IBM Market Intelligence Award, and IBM Outstanding Technical Achievement Awards. Dr. Mojsilovic is a member of INFORMS and a Senior Member of the IEEE.

M. Jonathan Richard *IBM Corporate Headquarters, Armonk, NY 10504 USA (mjrichar@us.ibm.com)*. Mr. Richard is a Managing Strategy Consultant with the Business Performance Services (BPS) Group, which develops, incubates, and deploys strategic growth and optimization initiatives across Business Units in IBM. He received a B.S. degree in economics and finance from Boston College in 1999 and an M.B.A. degree from MIT Sloan School of Management in 2009. He subsequently joined BPS, where he has focused on developing and deploying sales improvement and strategic planning initiatives.

Moninder Singh *IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (moninder@us.ibm.com)*. Dr. Singh is a Research Staff Member in the Business Analytics and Mathematical Sciences Department at the Thomas J. Watson Research Center. He joined IBM in 1998 after receiving his Ph.D. degree in computer and information science from the University of Pennsylvania. His research interests include the general areas of artificial intelligence, machine learning, and data mining.

Mark S. Squillante *IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (mss@us.ibm.com)*. Dr. Squillante is a Research Staff Member in the Business Analytics and Mathematical Sciences Department at the Thomas J. Watson Research Center, where he leads the Stochastic Processes and Applications Group. He received a Ph.D. degree from the University of Washington. His research interests concern mathematical foundations of the analysis, modeling, and optimization of complex stochastic systems. He is an elected Fellow of ACM and IEEE, and the author of more than 200 technical papers and more than 30 issued or filed patents. His work has been recognized by The Daniel H. Wagner Prize (INFORMS), eight best paper awards, nine keynote/plenary presentations , and nine major IBM technical awards. He serves on the editorial boards of Operations Research, Performance Evaluation, and Stochastic Models.

Kush R. Varshney *IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (krvarshn@us.ibm.com)*. Dr. Varshney is a Research Staff Member in the Business Analytics and Mathematical Sciences Department at the IBM T. J. Watson Research Center. He received a B.S. degree in electrical and computer engineering from Cornell University in 2004 and S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology in 2006 and 2010, respectively. He subsequently joined IBM at the T. J. Watson Research Center, where he has worked on statistical signal processing, machine learning, and business analytics. In 2011, he received two IBM Research Technical Accomplishment Awards for business impact of outsourcing analytics and for software sales analytics. He is an author or coauthor of eight journal papers and 20 conference papers. Dr. Varshney is a member of Eta Kappa Nu, Tau Beta Pi, the Institute of Electrical and Electronics Engineers, and the Institute for Operations Research and the Management Sciences.