Architecture, Implementation and Evaluation of a Software System for Field Experiments on the Impact of Relative and Absolute Rebates on Consumer Behavior

Roland Fassauer and Andrej Werner

Abstract—This paper discusses the implementation of a software system prototype to run online experiments regarding differences in the reaction of newsletter receivers to absolute and relative rebates. Therefore a split-run component was added to an online cashback system to gather experimental data on their behavior. The prototype was then evaluated from a functional point of view by executing and verifying several split-run tests. The usability was evaluated using a standard questionnaire approach (ISONORM 9241/110). Having passed both tests, we offer the prototype to the scientific community as a stable platform to run further experiments.

Index Terms—Cashback, consumer behavior, rebates, split-run test.

I. INTRODUCTION

While large amounts of users are quite common in the ecommerce business reality (web site visitors, online customers, newsletter receivers …), the scientific literature is short on field experiments with similar sized experimental groups [1].

The University of Leipzig and Andasa GmbH, an online cashback provider (http://www.andasa.de), conducted a government-funded research project to explore the (online-) consumer behavior regarding varying rebate offerings in newsletters. Therefore a software system to do field experiments with the user base of Andasa was implemented.

Using this system split tests with relative and absolute cashback values where run and data on the resulting consumer behavior (including orders) was gathered.

A. Consumer Behavior in Response to Rebates

The behavior of consumers in response to rebates is explained by the prospect theory [2]: Humans decide between two options by comparing expected future gains or losses with a reference situation (usually the status quo), and not by the value of the outcome. In this spirit they rather take small but save benefits instead of big potential wins coming with a risk of losses [3].

Various effects of the presentation of prices or rebates on consumer behavior are described in [4]-[9]. The studies show that discounts of absolute value perform better for products with a high price, while percentages produce more effects for offers of lower value.

There is no guidance for medium price levels. Unsurprisingly, higher rebates are more attractive.

The studies behind these results are often based on a small number of participants (usually below 500), where carried out using questionnaires and do not include real transactions. Therefore their external validity is restricted.

The presented software system addresses this aspect. It allows the fast setup of various field experiments to study the behavior of more than 125,000 participants in their real online environment, including the collection of their real transactions in online shops (orders, leads).

B. Online Cashback Programs

An online cashback program (implemented by a cashback (software) system), gives all registered users cash back on their purchases to increase the loyalty of the users to the participating shops. Cashback is a form of rebate [9], [10].

Cashback systems have to acquire data about the behavior of their users to fulfill their purpose: To grant revenue-dependent cashback they must know the transactions of their users. The handling of personal data is regulated in Germany by the data security law (Bundesdatenschutzgesetz (BDSG)) and the tele-media law (Telemediengesetz (TMG)), [11]-[13].

Andasa GmbH operates an online cashback system in Germany. With over 3000 participating online shops it collects substantial data on the behavior of its users, for instance selected shop, click time, buying time, or revenue. Doing so Andasa aggregates data on multi-shop consumer behavior. For more information about properties of the Andasa user group see [14].

The literature recommends the usage of data on consumer behavior from multiple shops to achieve better shopping-recommendations [15]. This is usually impossible because of the competition between shops. However, the described software system allows the observation of multi-shop consumer behavior, even including the resulting orders. This is possible because of the multi-shop transactional data collected by the Andasa cashback program.

The shop over-arching nature of the gathered behavioral data allows also the analysis of other interesting questions. For instance could the comparison of the impact on promoted shops or shop categories with the impact on non-promoted shops uncover spillover effects of online marketing campaigns [16], [17].

C. Success Metrics in Email Marketing

The success of e-mail marketing campaigns is measured...
through Online Marketing Controlling by a number of key technical metrics: The send-out volume is defined by the size of the used part of the available e-mail distribution list, the opening rate describes the part of opened emails of all received emails and the click-through-rate covers all receivers who clicked links in the received emails \([18]-[21]\).

The conversion rate \([8]\) describes the percentage of newsletter receivers which reach a predefined business objective (download, registration, order …). It is a key ratio to define the success of an online marketing campaign, linking the online marketing measurements with the business goals of a company \([19]\), \([23]\), \([24]\). Further information on the composition, content, measurement and handling of emails can be found for instance at \([18]\), \([25]\), \([26]\).

D. Split-Run Tests with Newsletters

Split-run tests \([27]\) are a standard procedure in online marketing \([28]-[32]\). Split-run tests measure the response of two different stimuli on two groups which are otherwise statistically identical.

For a split-run test using email marketing two variants of an email newsletter which differ only in one aspect (see Fig. 1, 2), are send out simultaneous to two randomly selected groups of receivers. The success is measured by a predefined online marketing ratio, for instance click-through rate, opening rate, or conversion rate \([33]\).

Every test must be regarded as a single, stand-alone experiment \([34]\), results from tests carried out at different times are not comparable: It must be assumed that all disturbances are independent of each other and take effect on the subjects with similar intensity \([35]\).

This is only valid, if the measurement of the dependent variable(s) is done in the experimental group and the control group simultaneously.

The statistical evaluation later on is done usually by ANOVA, using t-test or chi-quadrat-test. This requires t-distribution (or Gaussian distribution for \(n>30\)) and homogeneity of variance of the probed variable.

II. THE PROTOTYPE

A. Existing Environment

The basic operational architecture of the Andasa System is shown in Fig. 3. The platform is based on Windows Server 2008 R2, Microsoft SQL Server 2008 R2 and Internet Information Server 7.5. The development environment consisted of Microsoft Visual Studio, C#, .NET, Git (git-scm.com/) and Assembla (www.assembla.com).
The Andasa software system consists of a data warehouse, a content management system, an e-mail-send-out- and tracking component and a data mining system. This setup must be extended by a split-run test component which is integrated with the existing parts. (See Fig. 4)

To understand the implementation of the new component it is necessary to describe the basic functions of the system:

After the registration at Andasa all user-specific data is stored in a table. Since the given cashback for each order depends on various parameters, it is calculated using a campaign based approach. For example there may be different cashback values for users coming directly to the website or from the newsletter. In this case there are different campaigns for orders triggered by the newsletter and through the website. If an order is placed, the information is stored in a table to calculate the cashback value for the user. This table is again linked with shop, booking and campaign tables.

B. Implementation and Usage of the Split-Run Test Component Prototype

The new component could fortunately employ the existing web back-office (Fig. 6) and the existing database schema with small extensions. Only programmed logic to adapt the content management (CMS)-component and the send-out-component was necessary to enable the simultaneous and evenly distributed send-out of multiple versions of a newsletter.

The construction of a split-run test with the component requires several steps. In the first step a (HTML-) template for the newsletter, containing the basic structure and fix elements, is created with the content management. For each element to be tested the template contains a placeholder. This placeholder will be replaced with a so-called Buildingblock when the newsletter is sent out.

In the next step, these Buildingblocks need to be defined. They also consist of HTML-Code.

Multiple Buildingblocks can be assigned to the same placeholder in the template. At the time of the actual send-out, the send-out component randomly selects one of the assigned Buildingblocks per receiver and replaces the placeholder with it. Which receiver got which Buildingblock is then stored in the database. For instance, to test relative against absolute cashback values in the newsletter, it is only necessary to define two Buildingblocks substituting a placeholder representing the cashback value.

The regular Andasa newsletter (HTML) is sent out usually once a week on varying days to all users registered though a double opt-in process (standard verification process: an email is sent to the registered email address and the user needs to click a link in this email to verify his address). It takes the e-mail component about six hours to complete the process of creating the newsletter variants and send it.

The open-, click- and buying- behavior of the newsletter receivers is tracked by the tracking component using standard web tracking mechanisms like Link-Tracking or HTML-Tracking. The data is also stored in the database. For each newsletter receiver the following data is collected: Opening date & time, which links where clicked when, selected shops and shop categories, orders (in which shop, date & time, revenue).

The stored data is then extracted and pseudonymized [12] (so the data can no longer be assigned to a real person) for analysis.

C. Evaluation of the Prototype

We evaluated the software prototype in two ways: first we conducted several split-run tests with absolute and relative cashback to prove the viability from a functional perspective. Later on we used questionnaires to verify the usability of the system.

In total we did 15 split-run tests with more than 125,000 newsletter receivers each and could show that all parts of the software work as intended. We could also conclude as a first insight from these experiments that the participants are significantly more likely to order if confronted with an absolute cashback value as long as there was a clear call-to-action [20]. See [36].
prototype scored an ISONORM value of 103.8 points, which reflects good usability.

It has to be considered that the users did know that they used a prototype, what may have influenced the scores. The reflected weaknesses in fault tolerance, the ability to self-explain and friendliness to learn (Fig. 6) are to be expected in a prototype and will be improved over time.

III. CONCLUSION

It could be shown that the implemented software system allows the execution of large scale online field experiments. It forms a stable base for future experiments with the andasa user group.

The validity of the results is limited to the basic population of the Andasa newsletter receivers and cannot be simply applied to the whole internet population. Nevertheless the software system can used to confirm behavioral effects in a controlled real live environment.

Looking forward to follow-up experiments, especially the analysis of the impact of varying cashback rates on consumer behavior looks promising for the optimization of marketing campaigns.

Further we extended the system with a recommendation component to individualize the newsletters for each receiver. This allows us to set up field experiments to study the impact of personalized on consumer behavior.

We openly invite the scientific community to actively participate in these experiments – please contact us!

REFERENCES


Roland Fassauer was born in Leipzig Germany, on July 24, 1971, he holds a degree (Dipl. Inf.) of Computer Science at the faculty of Computer Science at the TH Leipzig, Germany, 1996. Currently he works as chief technical officer at Andasa GmbH in Leipzig. As a serial entrepreneur he acquired experience in the design and integration of e-business and e-commerce solutions and business processes at various companies (Intershop, Pixaco, Hewlett-Packard, Infai, e.g.) and in different research projects (FABEL, InterLog Grid, Smarter Shopping). Current themes of work are the study and design of recommendation systems and e-commerce systems.
Andrej Werner was born in Waluiki Russia, on June 7, 1973, he received diplom Wirtschaftsinformatik (Diploma Business Informatics) at the Faculty of Economics and Management at the Martin-Luther-University Halle/Wittenberg in Germany, 1999, Dr. rer. pol. in the area business informatics at the Faculty of Economics and Management at the Leipzig University in Germany, 2008, business process management and engineering. Currently he works as a post-doc researcher at the Information Systems Institute of the Faculty of Economics and Management at the Leipzig University (Germany). Experience in the design and integration of e-business and e-commerce solutions and business processes he acquired at various companies (Intershop, JARDIX, InfAI e.g.) and different research projects (Value Chain Cockpit, 10000 Smart Meter, Smarter Shopping e.g.). Current themes of work are the study and design of knowledge-intensive service systems based on ICT in various industries. Dr. Werner is member of German Informatics Society (Gesellschaft für Informatik e.V. (GI)).