

The Effect Of Evaluation Questions And Maximizing The Continuous-Time Diffusion Model

Posted on September 20, 2019

Abstract

The typical viral model defines effective marketing users to maximize social networks accepting a product, including unlimited users, campaign budgets and time. In fact, there is a need for different advertising campaigns that show users of limited interest; users are convinced to spend costs, advertisers limit their budgets and expect their adoption to be maximized within a short period of time. Considering a user's problem, money and restrictions, we have met the problem as the problem to maximize AL modular in a continuous time diffusion model at the intersection of the Metroid and multiple backpack constraints. $E || + N || V, n = 0 (1 / \epsilon^2)$ randomization and influence user evaluation algorithm randomizing network estimation with ribs ($| V |$ nodes, $| e |$ values $O(n || |)$ as subprogram we developed algorithm with a portal-aggressive algorithm with for adaptation k backrest coefficient borders to best approach $/(2k + 2)$ using the algorithm to estimate the effect. Comprehensive millions of experimental network nodes show that the proposed algorithms provide the technologies and Latest in terms of efficiency and scalability.

Keywords: to maximize the impact of anticipation of the impact on the model spread over time, matroid, backpack

Introduction

Online social networks play an important role in promoting new products, disseminating news, success in political campaigns and the dissemination of technological innovation. In this regard, the problem of maximizing the effect (or viral marketing problems) is usually the following flavors: the social of each other in the network to determine the user effectively convincing the product, that will affect other users on the network and take maximum obese. This problem has been widely studied in literature and algorithms. Usually, this is represented by one product (e.g., an online social platform) that is endlessly interesting to the user and has an unlimited budget and infinite time. However, in fact, the host usually tends to be more severe:

- Different restrictions on the section. Some products can be distributed at the same time as the same set of social things. These products can have different characteristics, such as income and speed of spread.

- Time limits: advertisers are expected to take action at a certain time interval, and different products may have different timescales.
- User restrictions. Users of the social network, each of which can be a potential source, just want to disclose a small number of ads. In addition, users can group according to geographical location, and the audience can access the audience they want.
- Restrictions on the product. It will accept the cost users of the notifications, which the advertiser must pay, and often have a limited amount of money.

For example, Facebook (e.g., E. lessor) ads for various products, such as clothing, books or cosmetics, that have different attributes. While some products, such as clothing, for a short period of time, some people may have longer periods of time, such as books. In addition, Facebook limits limit the number of advertisements on the side panel of each user (typically less than five), and as a result, not all ads may be attributed to a number of users. Finally, all advertisers have a limited budget for advertising on Facebook, so no one can show up for a number of user layers. In our work, we take into account these practical and important requirements of the optimization problem.

In applying the interrelated models of a particular product, we recommend a variety of time and product constraints. We prefer to have continuous time differentiation models rather than discrete time models, which are used more frequently in previous studies. The reason for this is that there is an extra error between artificial separation in timebox in boxes. We can configure advanced settings, such as the size of the hopper, to be balanced with the cost of errors and calculations, but it is not easy to choose the parameters in the most appropriate way. Experimental comparison Extensive both synthetic and real data show that the result of a cut-off model of predictions is as accurate as effectiveness or continuous copies.

However, reducing the effect based on the differential models with continuous time to the upper level creates additional problems. First of all, in this case, maximize the estimate. The purpose of the problem is difficult to affect the effect of a graphic sample of the modeling problem, t. E. Calculate constant intensity variable limitation in graphic circular models. The exact answer can only be calculated in very specific cases. For example, Gomez-Rodriguez and Sholkopf (2012) show that using the theory of the ongoing Markov process, a problem can be completely solved when transferring functions are exponentially density. However, the computational complexity of this approach is often made independent of the size and intensity of the network. In addition, an extra non-core approach will be required to improve a wider approach that will help promote the calculation of calculation. Secondly, it is not clear how to filter forecasts and optimization algorithms based on continuous-time differentiation models to influence millions of nodes. Especially in the case of optimization, the procedure to assess the effect for different subclusters of selected nodes must be taken many times. For this reason, our first goal is to develop scalable algorithms that can predict the impact in web mode with millions of nodes.

We will put accounts bordering user and product borders as the potential areas where maximization is

made. First of all, we note that the overall effect of the function of various submodulnoy functional products, and then User constraints and product borders are consistent with submodalities function locale. Informing us, evaluating any previous study of user and product constraints at the same time in different distribution networks, which are usually unknown, with various irregular costs. In particular, we initially made both product and user borders with an unlimited time window with uniform costs, which significantly reduced the specific type of our formulas. Similarly, a distribution problem is more than one product that can compete in an unlimited time window. In addition, it is assumed that more than one product is distributed on the same network. On the contrary, our products can usually allow different products to differ in different networks, which are not known in practice. The problem study on maximizing efficiency for a product subject to restrictions on a backpack on a two-way map identified between the marketing channels and potential customers was studied. Consumer accounts were included, but product restrictions were not ignored during the initial assignment, and the number of scientists studying the phenomenon of cross-sale (selling the first product increases the chances of selling the second product) and financial constraints on each product. However, users' constraints were not included, and the cost of each user was the same for each product. Therefore, our second goal is an effective algorithm for submodularity that can be considered at the same time developing both user and product constraints.

Influence Estimation

Gomez-Rodriguez et al. Developed a model of continuous diffusion. (2011), and then expressly thinks the problem is to assess the problem in terms of probabilistic graphic models. Given that efficiently efficient impact values for each node are very insignificant, we are developing a scalable algorithm for the predictions of the effects that millions of network nodes can deal with. The procedure for estimating the impact forecast as an important construction block for our maximum algorithm will be followed.

Networks Of Continuous Differences

A constant density diffusion model associates time along the edge of the transmission along the edge, unlike previous models, a specific time, with each edge combined with the function of transferring, that is, E , and the possibility of continuous infection of each rib. In addition, at the same time, time-consuming models differ in the sense that the events are in a progressive case that is repeatedly produced in shells, but the time of the events is generated directly from the function to translate in continuous time models.

Continuous independent cascade model. Thinking of the targeted web, $G = (V, E)$, we use an independent cascade model to shape the spreading process. The process starts with a series of contaminated source nodes (A), which initially found the unique contamination (idea, nozzle or product) as zero. Pollution is directly transmitted to neighbors with output edges from welded seeds.

The transition requires each edge of random waiting time (τ), taken from different independent double times (one at each end). Then, the infected neighbors offer the infection of their neighbors, and the process continues. We assume that the infected node is infected for each diffusion process. So, if the site is infected with some neighbors, it will be just a neighbor, a real parent, in which the unit is infected first. As a result, all of the process of non-imperative steering graph (SCM) is induced, although the network can be randomly targeted network.

Suitable conduct functions. It is for directional tubes ($j \rightarrow i$) conditional node density that is infected at time t | formal Dwarf communication function $f_{ji}(t_j | t_i)$. Suppose it is accepted that the variable change is: $f_{ji}(t_i | t_j) = f_{ji}(\tau_{ji})$, where $\tau_{ji} = t_i - t_j$ and causative: $\tau_{ji} < 0$ mura $f_{ji}(\tau_{ji}) = 0$. Its functions include parameters, such as exponential and Rayleigh functions, and use and predict non-parametric functions from cascading data.

The shorter way. In independent phase models, there is a useful function that we will have to use later when a sample of the edge's amortization time is taken into account; the time spent on the site is an infection with the edge node weight of the edges associated with the length of the path shorter (1) It represents the period.

Probabilistic graphic model for networks with constant discrimination time

Independent speed is a continuous-time model that is devoted to a series of graphic models to random random variables, i.e. time infection t in the support of a conditional independence structure in the communications network G . Despite the presence of the original track schedule. G -ringed rings, all transit (or cascade) processes stimulate a direct acrylic graph (DAG). On the steps, which are consistent with the specific GCM, we can simulate a common density using similar oriented graphics:

$$p(\{t_i\} | V) = p(t_i | \{t_j\} j \in \pi_i), (1)$$

In the case of each sum node, in the stimulated parents of the GCM, all members meet the conditional density when the parent node ($t_i | \{t_j\} j \in \pi_i$) takes into account the intake time. This is true due to the time of parent infection; it does not depend on other infected times that correspond to the local Markov property of the graphic orientation model. Note that the independent auto-movement model clearly indicates only the direction of the front of a two-way transfer function but cannot be determined directly with the conditional pneumatic density ($t_i | \{t_j\} j \in \pi_i$).

Overall Algorithm

Algorithms are to optimize low modular backpacks with different constraints to achieve coefficients about 1 - 1. Therefore, we can tempt

Equation (19) and $|V|$ Backpack

Borders, the problem, therefore L | a problem of submodular optimization. +

$|V|$ blocking. However, this naive approach does not fit for large-scale

Cases, since the length of these algorithms,

blocking. Instead, we prefer algorithms for lower modular optimization

k on the backpack and borders of matroid P , which gets the best ratio

This is the 1 with polynomial time algorithms. However, $P + 2k + 1$ is still not enough since our problem is $k = |L| P = 1$ can be small, but it can be great.

Here, we developed an algorithm to obtain the best rate approximation, using the following important comments regarding the structure of the tasks defined in the equation (19): Restrictions of the backpack on Z_i * different groups, Full sets of tires, and the purpose – the functions podmodulyarnyh size by the different groups,

Details of the BudgetMax algorithm are defined in Algorithms 1. BudgetMax different values of the threshold values of the p density and subroutine start to get answers for each p . Special product for a specific user and, ultimately, a solution with the maximum value of purpose. Amazingly, the algorithm offers a limit to the search area for a set of optimum costs. Subroutine data for constant solutions of p is used to determine the intensity threshold explained in Algorithm 2. Geometrically Inspired algorithm dimensions in the ratio of $1 + \text{factor}$ as far as the set (G) and the gain limit threshold (wt) will not be enough. They can be set identically to zero from the lazy evaluation intuition (Leskovec et al., 2007). In each wt , the sub-section selects all new elements that provide the following properties:

1. This can be done, and the density coefficient (the ratio of the limiting gain and the cost) exceeds the current density threshold;

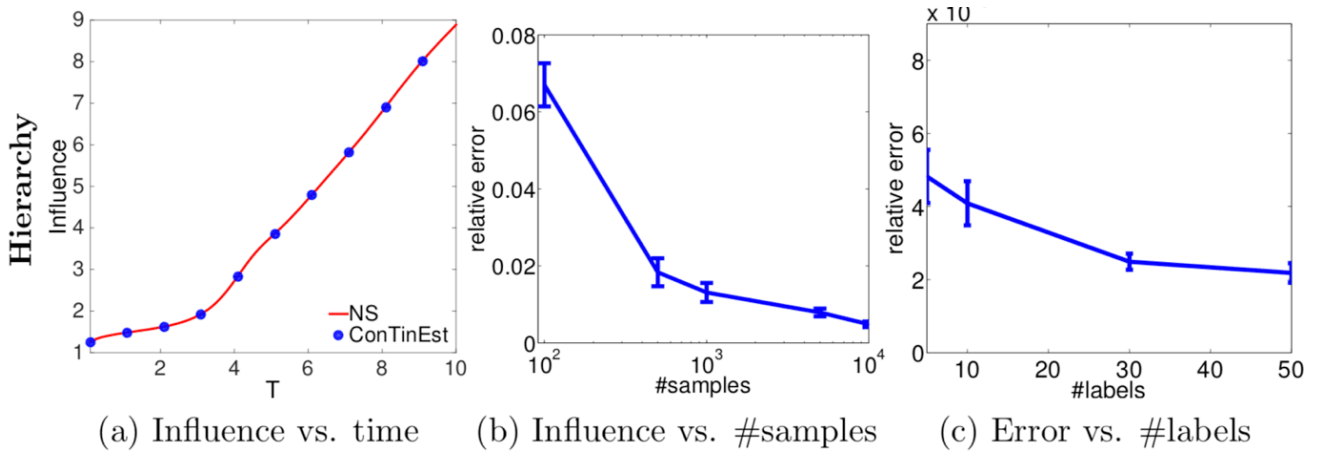
2. Edge increase

$f(G \cup \{z\}) - f(G)$ above the yield threshold threshold.

The term “density” is used based on the knapsack problem where the marginal returns are by weight and volume volume. High intensity means earning a lot without paying enough money. Soon, the algorithm does not take into account high-quality tasks and chooses what is possible repeatedly with a small victory too much.

Experiments on synthetic and real data

In this section, we assess the accuracy of the anticipated effect provided by ConTanEst and then examine the effectiveness of increasing the impact on synthetic and real networks, including ConTinEst, in the context of BudgetMax. Our approach shows that the quality of both speeds and solutions is better than the most modern modes.



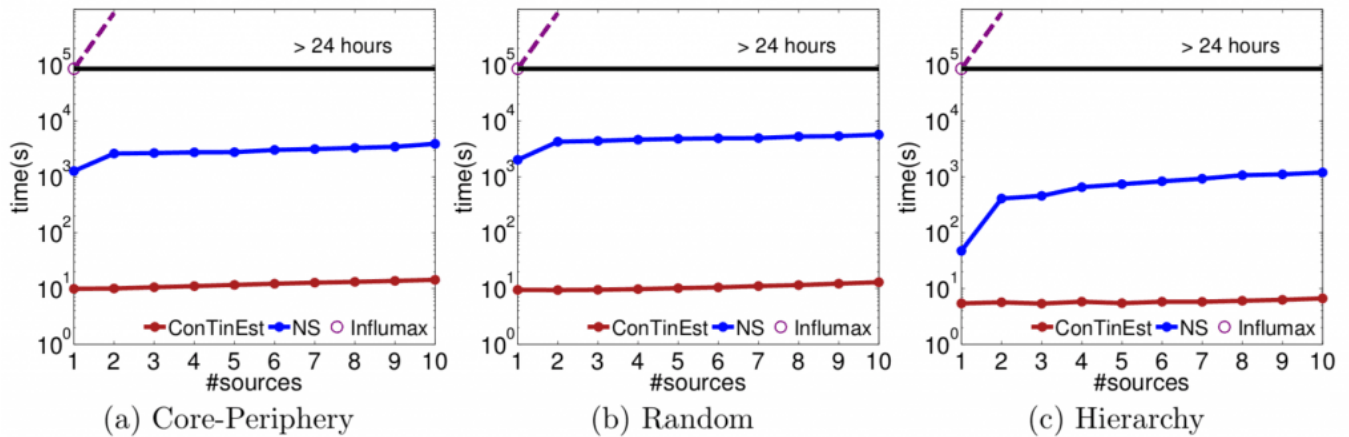
Core-environment forecast, random and forecast for hierarchical networks with 1,024 nodes and 2,048 fronts. Column (a) shows the estimated effect of NS (close to the predicted reality) and continues to increase the time window T ; Column (b) shows the relative error of ConTinEst with respect to sample number 5, randomly marked and $T = 10$; Column (c) shows the relative error of ConTinEst with respect to 10,000 random samples and the number of random tags with $T = 10$.

Adaptive Threshold Effects

We have investigated the accuracy of the algorithm of the threshold threshold adaptive value and the influence over time of the algorithm investigation with respect to the lazy evaluation method. Note that the execution of the slow evaluation and the duration of the work does not depend on δ since it is not related to it. Panel (a) shows the threshold threshold effect. As expected, the higher the value of δ , the lower the precision. However, our method is relatively robust for a particular choice since its performance is always above 90% relative accuracy, even in large dimensions. Panel (b) shows the execution time from the threshold threshold. In this case, the higher the value, the shorter the work time. In other words, the study confirms the perception that the time period and the distribution can shift the quality of the solution to δ .

Scalability

In this section, we begin with an assessment of the scalability of the proposed algorithms with respect to the classic maximization problem. We only have one product with the main flow restrictions for our users. It is comparable to Influmax and Naive Sampling (NS), which is the most modern method for continuous assessment of time and maximizing working time. For ConTinEst, we make 10,000 samples in the outer loop, each of which contains five random tags in the inner loop. We attach ConTinEst as a subroutine to the classic greedy algorithm. We also recruit 10,000 samples for NS. The first two experiments were performed with a single 2.4 GHz processor.



Experiments On Data

In this part, we first decide how well our proposed algorithm reflects the true impact on a series of global data. We will then evaluate the quality of the resources chosen to maximize the effect of different constraints. We used a series of MemeTracker public data with over 172 million news articles and blog posts from 1 million major sites and blogs.

Then, we consider the budgetMax accuracy as follows. First of all, each group is divided equally into the training set and test series, and then we study an ongoing time pattern and a split-time sample for each group using the training sets. As we have already, we prefer NetRate with an exponential transfer function for continuous time models and for models with a specific time, we are studying the probability of infection with the use of the method and Netrapalli Sanghavi (2012), which sets the stride length. One. Secondly, we run a BudgetMax using a continuous model of time and a time-consuming model. We specify BudgetMax using the greedy (non-overlap) method, which is a short-term model. Since we do not have any real knowledge of the cost of each node, we are focusing on our experiments at a monotonous cost. Once we have been appointed in the study networks, we must evaluate the performance of the two techniques using cascades in the test set, if a pair of a group of nodes (i, j) are given to us, let us know $C(j)$. Brake cluster that group I began, with a group j. Then, take the average number of nodes after the j for each step in j (j) as a proxy for the average effect due to the appointment of j-groups. Finally, the selection effect is the average effect of each pair of group nodes in the solution. In our experiments, we accidentally chose 128 minutes as our target users.

This applies to four factors: the T. The greedy window (IC) and random distribution, but (a) the number of products, (b) the product budget, (c) the budget for the user, and (d) that BudgetMax receives a distribution that causes the largest spread in the Test data and usually provides an average improvement of 20 percent.

The study makes a qualitative analysis of the actual distribution of the groups (things or events of the

world, red) and settings (black). Here are some examples that can be expected intuitively: “Movant Wall Street” «japantoday.com» Fukushima assignment of a nuclear disaster, or “finance.yahoo.com.” Designed In addition, since we are dealing with themes and Real-world events with different main spreads of nets, the selected nest nodes are only very popular with media such as nytimes.com or cnn.com, but also more moderate, often specialized or local settings. freep.com or localnews8.com.

Conclusion

We studied the effect of evaluation questions and maximizing the continuous-time diffusion model. First of all, we recommend ConTinEst, a randomization algorithm to estimate the efficiency of scale up to millions of nodes, thus greatly improving the accuracy of the forecast in terms of accuracy than previous methods. When we are in a routine to the productive impact of large networks, our goal is to optimize the impact of different types of products (or information) in the time of their distribution networks based on different practical constraints: Different products can be diffused differently; In a specific time window, the effect is only considered; For each user that can be encouraged to just a small number of products, and all budget products are limited to the campaign, which is expensive to deploy to the user. We are creating a new formulation to maximize submodulnuyu at the matrix crossing restrictions and group feedback boundaries, and then we develop an adaptive greedy algorithm portal, which we call Bud-getMax, which is effective in guaranteeing a reasonably priced approach. Experimental results show that the proposed algorithm is better than the other scalable options of its synthetic and true data set. There are some interesting open problems. For example, when the effect is evaluated using ConTinEst, its error is a random one. How does this affect the function of the module? Is there an optimization algorithm with greater accuracy for random error? These issues are reserved for future work.

References

- Odd Aalen, Oernulf Borgan, and Håkon K Gjessing. Survival and event history analysis: a process point of view. Springer, 2008.
- Ashwinkumar Badanidiyuru and Jan Vondrák. Fast algorithms for maximizing submodular functions. In SODA. SIAM, 2014.
- Eytan Bakshy, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. Everyone’s an influencer: Quantifying influence on Twitter. In International Conference on Web Search and Data Mining (WSDM), pages 65–74, 2011.
- Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. Influence maximization in social networks: Towards an optimal algorithmic solution. arXiv preprint arXiv:1212.0884, 2012.
- Wei Chen, Yajun Wang, and Siyu Yang. Efficient influence maximization in social networks. In

Knowledge Discovery and Data Mining (SIGKDD), pages 199–208. ACM, 2009.

Wei Chen, Chi Wang, and Yajun Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In Knowledge Discovery and Data Mining (SIGKDD), pages 1029–1038. ACM, 2010a.

Wei Chen, Yifei Yuan, and Li Zhang. Scalable influence maximization in social networks under the linear threshold model. In International Conference on Data Mining series (ICDM), pages 88–97. IEEE, 2010b.

Wei Chen, Alex Collins, Rachel Cummings, Te Ke, Zhenming Liu, David Rincon, Xiaorui Sun, Yajun Wang, Wei Wei, and Yifei Yuan. Influence maximization in social networks when negative opinions may emerge and propagate. In SIAM International Conference on Data Mining (SDM), pages 379–390. SIAM, 2011.

Wei Chen, Wei Lu, and Ning Zhang. Time-critical influence maximization in social networks with time-delayed diffusion process. In AAAI Conference on Artificial Intelligence (AAAI), 2012.

Aaron Clauset, Cristopher Moore, and M.E.J. Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101, 2008.