

# A DATA CENTRIC APPROACH TO UTILITY CHANGE DETECTION IN ONLINE SOCIAL MEDIA.

Anup Aprem

Dept. of Electrical and Computer Engineering  
University of British Columbia  
Vancouver, B.C., Canada  
aaprem@ece.ubc.ca

Vikram Krishnamurthy

Dept. of Electrical and Computer Engineering  
Cornell Tech, Cornell University, NY  
vikramk@cornell.edu

## ABSTRACT

This paper considers the problem of detecting changes in utility maximizing behaviour of agents in online social media. Such changes in utility maximizing behaviour in online social media occur due to the effect of marketing, advertising, or changes in ground truth. In contrast to traditional signal processing techniques, our approach is data-centric. We use the framework of revealed preference to detect the unknown time point (change point) at which the utility function changed. We derive necessary and sufficient conditions for detecting the change point. In addition, we provide an algorithm to recover the utility function before and after the change point. The results developed are illustrated on the Yahoo! Tech Buzz dataset. From the dataset, we obtain the following useful insights: First, the changes in ground truth affecting the utility of the agent can be detected by utility maximization behaviour in online search. Second, the recovered utility functions satisfy the single crossing property indicating strategic substitute behaviour in online search.

**Index Terms**— Social media, utility maximization, revealed preference, change point detection

## 1. INTRODUCTION

This paper deals with the problem of non-parametric change detection of utility maximizing behaviour in online social media. The problem we consider is fundamentally different to the theme used widely in the signal processing literature, where one postulates an objective function (typically convex) and then develops optimization algorithms. In contrast, the revealed preference framework, considered in this paper, is *data centric* - given a dataset, we wish to determine if it is consistent with utility maximization, and then detect changes in the utility function based on the observed behaviour.

The problem of non-parametric detection of utility maximizing behaviour is the central theme in the area of revealed preferences in microeconomics. Revealed preference in microeconomics, aims to answer the following question: Given a dataset,  $\mathcal{D}$ , consisting of probe,  $p_t \in \mathbb{R}_+^m$ , and response,  $x_t \in \mathbb{R}_+^m$ , of an agent for  $T$  time instants:

$$\mathcal{D} = \{(p_t, x_t), t = 1, 2, \dots, T\}, \quad (1)$$

Is the dataset in (1) consistent with utility-maximization behaviour of an agent? A utility-maximization behaviour (or utility maximizer) is defined as follows:

The full version of the paper [1] has been accepted to IEEE Transactions on Signal Processing

**Definition 1.1.** An agent is a utility maximizer if, at each time  $t$ , for input probe  $p_t$ , the output response,  $x_t$ , satisfies

$$x_t = x(p_t) \in \operatorname{argmax}_{\{p'_t x \leq I_t\}} u(x). \quad (2)$$

Here,  $u(x)$  denotes a locally non-satiated<sup>1</sup> utility function<sup>2</sup>. Also,  $I_t \in \mathbb{R}_+$ , is the total resource or the budget of the agent. The linear constraint,  $p'_t x \leq I_t$  impose a budget constraint on the agent, where  $p'_t x$  denotes the inner product between  $p_t$  and  $x$ .

Major contributions to revealed preference are due to Samuelson [2], Afriat [3], Varian [4], and Diewert [5] in the microeconomics literature. Afriat [3] devised a nonparametric test (called Afriat's theorem), which provides necessary and sufficient conditions to detect utility maximizing behaviour for a dataset. For an agent satisfying utility maximization, Afriat's theorem [3] provides a method to construct a utility function consistent with the data. The utility function, so obtained, can be used to predict future response of the agent. Varian [6] provides a comprehensive survey of revealed preference literature.

Despite being originally developed in economics, there has been some recent work on application of revealed preference to social networks and signal processing. In the signal processing literature, revealed preference framework was used for detection of malicious nodes in a social network in [7, 8] and in demand estimation in smart grids in [9, 10]. [11] analyzes social behaviour and friendship formation using revealed preference among high school friends. In online social networks, [12] uses revealed preference to obtain information about products from bidding behaviour in eBay or similar bidding networks.

In this paper, we extend revealed preference framework to agents with dynamic utility functions. The utility function jump changes at an unknown time instant by a linear perturbation. Given the dataset of probe and responses of an agent, the objective is to develop a nonparametric test to detect the change time and the utility functions before and after the change.

Such utility change point detection problems arise in online search in social media. The online search is currently the most popular method for information retrieval [13]. There has been a gamut of research which links internet search behaviour to ground

<sup>1</sup>Local non-satiation means that for any point,  $x$ , there exists another point,  $y$ , within an  $\varepsilon$  distance (i.e.  $\|x - y\| \leq \varepsilon$ ), such that the point  $y$  provides a higher utility than  $x$  (i.e.  $u(x) < u(y)$ ).

<sup>2</sup>The utility function is a function that captures the preference of the agent. For example, if  $x$  is preferred to  $y$ ,  $u(x) \geq u(y)$ .

truths such as symptoms of illness, political election, or major sporting events [14, 15, 16, 17, 18, 19]. Hence, a change in the utility in the online search corresponds to change in ground truth or exogenous events affecting the utility of agent, such as the onset of disease or the announcement of major political decision. Detection of utility change in online social behaviour, therefore, is helpful to identify changes in ground truth and useful, for example, for early containment of diseases [15] or predicting changes in political opinion [20, 21]. Also, the intrinsic nature of the online search utility function motivates such a study under a revealed preference framework.

The problem of detecting a linear perturbation change in the utility function is motivated by several reasons. First, it provides sufficient selectivity such that the non-parametric test is not trivially satisfied by all datasets but still provides enough degrees of freedom. Second, the linear perturbation can be interpreted as the change in the marginal rate of utility relative to a “base” utility function. In online social media, the linear perturbation coefficients measure the impact of marketing or the measure of severity of the change in ground truth on the utility of the agent. This is similar to the linear perturbation models used to model taste changes [22, 23, 24, 25] in microeconomics. Finally, in social networks, linear change in the utility is usually used to model the change in utility of an agent based on the interaction with the agent’s neighbours [26]. Compared to the taste change model, our model is unique in that we allow the linear perturbation to be introduced at an unknown time. To the best of our knowledge, this is the first time in the literature that change point detection problem has been studied in the revealed preference setting.

The organization of the paper is as follows: In Sec. 2, we derive necessary and sufficient conditions for change point detection, for dynamic utility maximizing agents under the revealed preference framework. Section 3 illustrates the result on the Yahoo! Tech Buzz dataset. Concluding remarks are offered in Section 4.

## 2. UTILITY CHANGE POINT DETECTION

In this section, we extend the revealed preference framework and consider agents whose change in utility function can be modelled by linear perturbations. We state below Afriat’s theorem<sup>3</sup> for a dataset in (1) to satisfy utility maximization model in (2).

**Theorem 2.1** (Afriat’s Theorem [3]). *Given a dataset  $\mathcal{D}$  in (1), the following statements are equivalent:*

1. *The agent is a utility maximizer and there exists a monotonically increasing<sup>4</sup> and concave utility function that satisfies (2).*
2. *For scalars  $u_t$  and  $\lambda_t > 0$  the following set of inequalities has a feasible solution:*

$$u_s - u_t - \lambda_t p'_t(x_s - x_t) \leq 0 \quad \forall t, s \in \{1, 2, \dots, T\}. \quad (3)$$

<sup>3</sup>To the signal processing reader unfamiliar with this theorem, it can be viewed as a set-valued system identification method for an argmax nonlinear system with a constraint on the inner product of the probe and response of a system. Afriat’s theorem has several interesting consequences including the fact that if a dataset is consistent with utility maximization, then it is rationalizable by a concave, monotone and continuous utility function. Hence, the preference of the agent represented by a concave utility function can never be refuted based on a finite dataset.

<sup>4</sup>In this paper, we use monotone and local non-satiation interchangeably. Afriat’s theorem was originally stated for a non-satiated utility function.

3. *A monotonic and concave utility function that satisfies (2) is given by:*

$$u(x) = \min_{t \in \{1, 2, \dots, T\}} \{u_t + \lambda_t p'_t(x - x_t)\} \quad (4)$$

4. *The dataset  $\mathcal{D}$  satisfies the Generalized Axiom of Revealed Preference (GARP), namely for any  $t \leq T$ ,  $p'_t x_t \geq p'_t x_{t+1} \quad \forall t \leq k-1 \implies p'_k x_k \leq p'_k x_1$ .  $\square$*

### 2.1. System Model: Dynamic utility maximization

Consider an agent who selects  $x$ , at time  $t$  to maximize the utility function given by:

$$u(x, \alpha; t) = v(x) + \alpha' x \mathbb{1}\{t \geq \tau\}, \quad (5)$$

subject to the following linear constraint  $p'_t x \leq I_t$ . Here,  $\mathbb{1}\{\cdot\}$  denotes the indicator function and  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$  denotes the  $m$ -dimensional linear perturbation vector. In contrast to the static utility function in (2), the utility function,  $u(x, \alpha; t)$ , in (5), consists of two components: a base utility function,  $v(x)$ , and a linear perturbation,  $\alpha' x$ , which occurs at an unknown time  $\tau$ . The utility function,  $u(x, \alpha; t)$  is assumed to be monotonic<sup>5</sup> and concave<sup>6</sup> conditioned on  $\alpha$ . We will restrict the components of the vector  $\alpha$  to be (strictly) greater than 0, so that the utility function,  $u$ , conditioned on  $\alpha$  is monotonic. The objective is to devise Afriat type inequalities to detect the time,  $\tau$  at which linear perturbation is introduced to the base utility function. Theorem 2.2 summarizes the necessary and sufficient conditions to detect the change in utility function according to the model in (5) and the proof is provided in the Appendix.

**Theorem 2.2.** *The dataset in (1) is consistent with the model in (5), if we can find sets  $\{v_t\}_{t=1, \dots, T}$ ,  $\{\lambda_t > 0\}_{t=1, \dots, T}$ ,  $\{\alpha_k\}_{k=1, \dots, m}$ , such that there exists a feasible solution to the following inequalities:*

$$v_t + \lambda_t p'_t(x_s - x_t) \geq v_s \quad (t < \tau) \quad (6)$$

$$v_t + \lambda_t p'_t(x_s - x_t) - \alpha' (x_s - x_t) \geq v_s \quad (t \geq \tau) \quad (7)$$

$$\alpha_i \leq \lambda_t p_t^i \quad (\forall i, t \geq \tau), \quad (8)$$

where  $\alpha_i$  and  $p_t^i$  are the  $i^{\text{th}}$  component of the linear perturbation  $\alpha$  and probe vector  $p_t$ , respectively.

Note that the inequalities in (6) to (8) closely resemble the Afriat inequalities in (3). The time step,  $\tau$  at which the inequalities are satisfied is the time at which the linear perturbation is introduced.

### 2.2. Recovery of minimum linear perturbation coefficients and base utility function

Computing the linear perturbation coefficients gives an indication of the severity of the ground truth or the effect of marketing and advertising. The solution to the following convex optimization provides

<sup>5</sup>Monotone utility function models the human preference: *more is preferred to less*.

<sup>6</sup>Concavity of utility function models the human preference: *averages are better than the extremes*. It is also related to the law of diminishing marginal utility, i.e. the rate of utility decreases with  $x$ . One of the most widely used utility function is the *Cobb-Douglas* utility function which has the form  $u(x, y) = x^a y^{1-a}$ ,  $0 < a < 1$ , in 2-dimension.

the minimum value of the perturbation coefficients:

$$\min \|\alpha\|_2^2 \quad (9)$$

$$\text{s.t. } v_t + \lambda_t p'_t(x_s - x_t) \geq v_s \quad (t < \tau) \quad (10)$$

$$v_t + \lambda_t p'_t(x_s - x_t) - \alpha'(x_s - x_t) \geq v_s \quad (t \geq \tau) \quad (11)$$

$$\alpha_i \leq \lambda_t p'_t(\forall i, t \geq \tau) \quad (12)$$

$$\lambda_t > 0 \quad (13)$$

$$v_1 = \beta, \quad \lambda_1 = \delta, \quad (14)$$

where,  $\beta$  and  $\delta$  are arbitrary constants.

The equations (10) to (12) correspond to the revealed preference inequalities. The normalization conditions (14) are required because of the ordinality<sup>7</sup> of the utility function. This is because for any set of feasible values of  $\{v_t\}_{t=1,\dots,T}$ ,  $\{\lambda_t\}_{t=1,\dots,T}$ ,  $\{\alpha_k\}_{k=1,\dots,m}$  satisfying the constraints in Theorem 2.2 the following relation also holds

$$\beta(v_s + \delta) - \beta(v_t + \delta) - \beta\lambda_t p'_t(x_s - x_t) + \beta\alpha'(x_s - x_t) \leq 0.$$

Recall, the base utility function is the utility function before the linear change.

**Corollary 2.1.** *The recovered base utility function is given by*

$$\hat{v}(x) = \min_t \{v_t + \lambda_t \tilde{p}'_t(x - x_t)\}, \quad (15)$$

where

$$\tilde{p}_t^i = \begin{cases} p_t^i & t < \tau, \\ p_t^i - \alpha_i / \lambda_t & t \geq \tau, \end{cases} \quad (16)$$

where  $\{v_t\}$ ,  $\{\lambda_t\}$ ,  $\{\alpha_k\}$  are the solution of (9) to (14).

The recovery of the base utility function in (15) is similar to (4) in Afriat's Theorem, except that the probe has been "adjusted" for the linear perturbation coefficients.

### 3. REAL DATASET (YAHOO! BUZZ GAME)

In this section, we present an example of a real dataset of online search process. The objective is to investigate the utility maximization of the online search process and to detect change points or unknown time at which the utility has changed. The change points give valuable information on when the ground truths have changed.

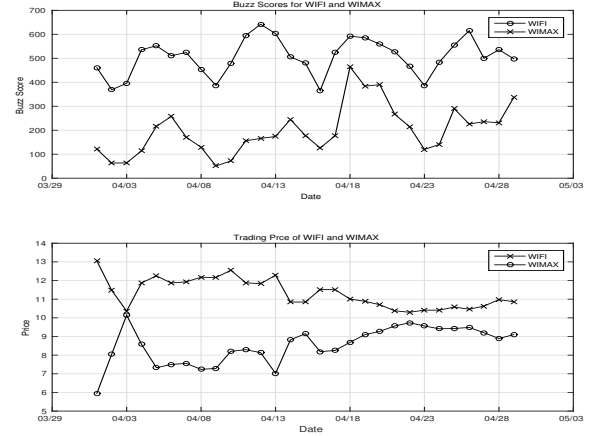
The dataset that we use in our study is the Yahoo! Buzz Game Transactions from the Webscope datasets<sup>8</sup> available from Yahoo! Labs. In 2005, Yahoo! along with O'Reilly Media started a fantasy market where the trending technologies at that point were pitted against each other. The players in the game have access to the "buzz", which is the online search index, measured by the number of people searching on the Yahoo! search engine for the technology. The objective of the game is to use the buzz and trade stocks accordingly. The interested reader is referred to [27] for an overview of the Buzz game. An empirical study of the dataset [28] reveal that most of the traders in the Buzz game follow a utility maximizing behaviour. Hence, the dataset falls within the revealed preference framework, if we consider the buzz as the probe and the "trading price"<sup>9</sup> as the response to the utility maximizing behaviour.

<sup>7</sup>Clearly any positive monotonic transformation of  $u(x)$  in (2) gives the same response.

<sup>8</sup>Yahoo! Webscope dataset: A2 - Yahoo! Buzz Game Transactions with Buzz Scores, version 1.0 [http://research.yahoo.com/Academic\\_Relations](http://research.yahoo.com/Academic_Relations)

<sup>9</sup>The trading price is indicative of the value of the stock.

We consider a subset of the dataset containing only the "WIRELESS" market which contained two main competing technologies: "WIFI" and "WIMAX". Figure 1 shows the buzz and the "trading price" of the technologies starting from April 1 to April 29. The buzz is published by Yahoo! at the start of each day and the "trading price" was computed as the average of the trading price of the stock for each day. Choosing the probe and response vector for this



**Fig. 1:** Buzz scores and trading price for WIFI and WIMAX in the WIRELESS market from April 1 to April 29. The change point was estimated as April 18. This corresponds to a new WIFI product announcement. The change can also be observed due to the sudden peak of interest in WIFI around April 18.

dataset as follows:

$$p_t = [\text{Buzz(WIFI)} \quad \text{Buzz(WIMAX)}] \\ x_t = [\text{Trading price(WIFI)} \quad \text{Trading price(WIMAX)}].$$

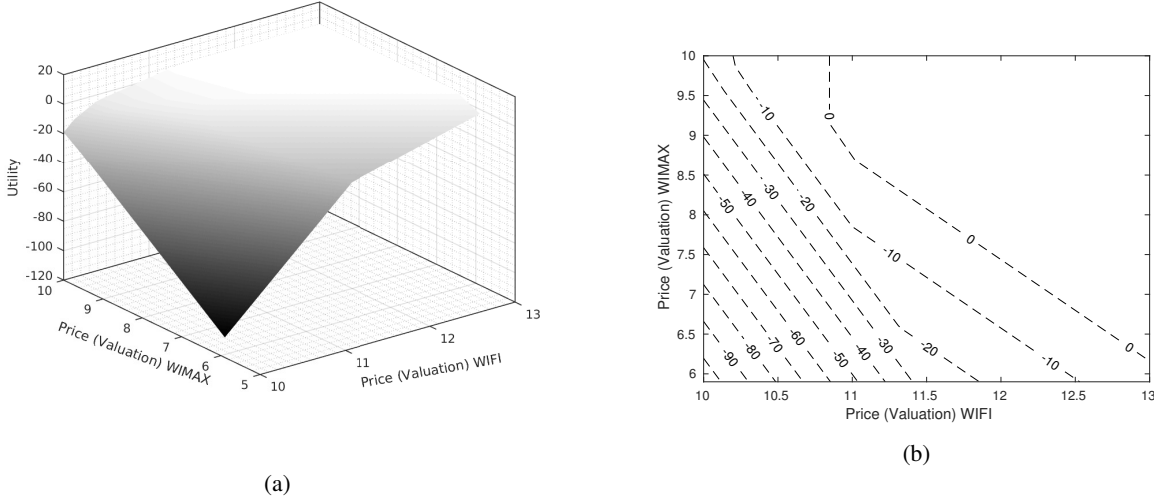
We find that the dataset does not satisfy utility maximization for the entire duration from April 1 to April 29, i.e. the Afriat inequalities (3) are not satisfied. However, we find that the dataset satisfies utility maximization from April 1 to April 17. Using the inequalities (6) to (8), that we derived in Sec 2, for the model in (5), we see that utility has changed with change point,  $\tau$ , set to April 18. This correspond to a change in the ground truth which affected the utility of the agents. Indeed, we find that the change point corresponds to Intel's announcement of WIMAX chip<sup>10</sup>.

Also, by minimizing the 2-norm of the linear perturbation, we find that the recovered linear coefficients which correspond to minimum perturbation is  $\alpha = [0 \ 5.9]$ . This is inline with what we expect, a positive change in the WIMAX utility, due to the change in ground truth. Furthermore, the recovered utility function,  $v(x)$ , is shown in Fig. 2a and the indifference curve (or contour plot) of the base utility is shown in Fig. 2b.

The recovered base utility function in Fig. 2a satisfy the single crossing condition<sup>11</sup> indicating strategic substitute behaviour in online search. The substitute behaviour in online search can also be

<sup>10</sup><http://www.dailywireless.org/2005/04/17/intel-shipping-wimax-silicon/>

<sup>11</sup>Utility function,  $U(x_1, x_2)$ , satisfy the single crossing condition if  $\forall x'_1 > x_1, x'_2 > x_2$ , we have  $U(x'_1, x_2) - U(x_1, x_2) \geq 0 \implies U(x'_1, x'_2) - U(x_1, x'_2) \geq 0$ . The single crossing condition is an ordinal condition and therefore compatible with Afriat's test.



**Fig. 2:** Fig. 2a shows the recovered utility function  $v(x)$  using (15). Indifference curve of the recovered utility function, is shown in Fig. 2b. The indifference curve suggest the substitution behaviour in online search.

noticed from the indifference curve in Fig. 2b. This is due to the fact that WIFI and WIMAX were competing technologies for the problem of providing wireless local area network.

#### 4. CONCLUSION

In this paper, we derived data driven methods for change point detection. We extended Afriat's theorem for agents with dynamic utility function. The online search process was considered as an illustrative example of an agent with dynamic utility function, whose utility is affected by ground truths. The main result, from Theorem 2.2, is deriving necessary and sufficient conditions to compute the change point at which the utility function jump changes by a linear perturbation. In addition, we provided an algorithm for detecting the unknown change point and recovering the utility function before and after the change point. The results were applied on the Yahoo! Tech Buzz dataset and the estimated change point corresponds to the change in ground truth.

#### Appendix: Proof of Theorem 2.2

**Necessary Condition:** Assume that the data has been generated by the model in (5). An optimal interior point solution to the problem must satisfy:

$$\nabla_{x_t^i} v(x_t) = \lambda_t p_t^i \quad (t < \tau) \quad (17)$$

$$\nabla_{x_t^i} v(x_t) + \alpha_i = \lambda_t p_t^i \quad (t \geq \tau) \quad (18)$$

At time  $t$ , the concavity of the utility function implies:

$$u(x_t, \alpha, t) + \nabla_{x_t} u(x_t, \alpha, t)'(x_s - x_t) \geq u(x_s, \alpha, t) \quad \forall s. \quad (19)$$

Substituting the first order conditions (17) and (18) into (19), yields

$$v(x_t) + \lambda_t p_t'(x_s - x_t) \geq v(x_s) \quad (t < \tau) \quad (20)$$

$$v(x_t) + \lambda_t p_t'(x_s - x_t) - \alpha'(x_s - x_t) \geq v(x_s) \quad (t \geq \tau) \quad (21)$$

Denoting  $v(x_t) = v_t$  gives (6) and (7). (8) comes from the fact the utility function  $v(x)$  is monotonic increasing.

**Sufficient Condition:** We first construct a piecewise linear utility function  $\mathcal{V}(x)$  from the lower envelope of the  $T$  overestimates, to approximate the function  $v(x)$  defined in (5),

$$\mathcal{V}(x) = \min_t \{v_t + \lambda_t \tilde{p}_t'(x - x_t)\}, \quad (22)$$

where each coordinate of  $\tilde{p}_t$  is defined as,

$$\tilde{p}_t^i = \begin{cases} p_t^i & t < \tau \\ p_t^i - \alpha_i / \lambda_t & t \geq \tau \end{cases} \quad (23)$$

To verify that the construction in (22) is indeed correct, consider an arbitrary response,  $\hat{x}$ , such that:  $p_t^i \hat{x} \leq p_t^i x_t$ <sup>12</sup>. We need to show  $\mathcal{V}(\hat{x}) + \alpha' \hat{x} \leq \mathcal{V}(x_t) + \alpha' x_t$ .

First, we show that  $\mathcal{V}(x_t) = v_t \forall t$  as follows: From (22), for some  $m$ ,

$$\mathcal{V}(x_t) = v_m + \lambda_m \tilde{p}_m'(x_t - x_m),$$

If,  $m \geq \tau$ ,

$$\begin{aligned} \mathcal{V}(x_t) &= v_m + \lambda_m \tilde{p}_m'(x_t - x_m) \\ &= v_m + \lambda_m p_m'(x_t - x_m) - \alpha'(x_t - x_m) \\ &\leq v_t + \lambda_t p_t'(x_t - x_t) \\ &= v_t \end{aligned}$$

If the inequality is true, then it would violate (21). Using similar technique, we obtain, if  $m < \tau$ ,  $\mathcal{V}(x_t) = v_t$ . Hence,  $\mathcal{V}(x_t) = v_t$ .

Next, we show  $\mathcal{V}(\hat{x}) + \alpha' \hat{x} \leq \mathcal{V}(x_t) + \alpha' x_t$ . If,  $t \geq \tau$ ,

$$\begin{aligned} \mathcal{V}(\hat{x}) + \alpha' \hat{x} &\leq v_t + \lambda_t \tilde{p}_t'(\hat{x} - x_t) + \alpha' \hat{x} \\ &= v_t + \lambda_t p_t'(\hat{x} - x_t) - \alpha'(\hat{x} - x_t) + \alpha' \hat{x} \\ &= v_t + \lambda_t p_t'(\hat{x} - x_t) + \alpha' x_t \\ &\leq v_t + \alpha' x_t = \mathcal{V}(x_t) + \alpha' x_t \end{aligned}$$

The inequality holds, similarly, for the case  $t < \tau$ . Therefore, we can construct a utility function that is consistent with the data satisfying the model in (5).  $\square$

<sup>12</sup>In microeconomic theory,  $x_t$  is said to be "revealed preferred" to  $\hat{x}$ . Since  $x_t$  was chosen as response for the probe  $p_t$ , the utility at  $x_t$  should be higher than the utility at  $\hat{x}$ .

## A. REFERENCES

- [1] A. Aprem and V. Krishnamurthy, "Utility Change Point Detection in Online Social Media: A Revealed Preference Framework," *IEEE Transactions on Signal Processing (Accepted)*, 2016.
- [2] P. Samuelson, "A note on the pure theory of consumer's behaviour," *Economica*, pp. 61–71, 1938.
- [3] S. Afriat, "The construction of utility functions from expenditure data," *International economic review*, vol. 8, no. 1, pp. 67–77, 1967.
- [4] H. Varian, "The nonparametric approach to demand analysis," *Econometrica*, vol. 50, no. 1, pp. 945–973, 1982.
- [5] W. Diewert, "Afriat and revealed preference theory," *The Review of Economic Studies*, pp. 419–425, 1973.
- [6] H. Varian, "Revealed preference," *Samuelsonian economics and the twenty-first century*, pp. 99–115, 2006.
- [7] V. Krishnamurthy and W. Hoiles, "Afriat's test for detecting malicious agents," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 801–804, 2012.
- [8] M. Barni and F. Pérez-González, "Coping with the enemy: Advances in adversary-aware signal processing," in *Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2013, pp. 8682–8686.
- [9] W. Hoiles and V. Krishnamurthy, "Nonparametric demand forecasting and detection of energy aware consumers," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 695–704, March 2015.
- [10] W. Hoiles, V. Krishnamurthy, and A. Aprem, "Pac algorithms for detecting nash equilibrium play in social networks: From twitter to energy markets," *IEEE Access*, vol. 4, pp. 8147–8161, 2016.
- [11] S. Currarini, M. O. Jackson, and P. Pin, "Identifying the roles of race-based choice and chance in high school friendship network formation," *Proc. of the National Academy of Sciences*, vol. 107, no. 11, pp. 4857–4861, 2010.
- [12] R. K.-X. Jin, D. C. Parkes, and P. J. Wolfe, "Analysis of bidding networks in ebay: aggregate preference identification through community detection," in *Proc. of AAAI workshop on PAIR*, 2007.
- [13] J. Strebel, T. Erdem, and J. Swait, "Consumer search in high technology markets: exploring the use of traditional information channels," *Journal of Consumer Psychology*, vol. 14, no. 1, pp. 96–104, 2004.
- [14] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant, "Detecting influenza epidemics using search engine query data," *Nature*, vol. 457, no. 7232, pp. 1012–1014, 2009.
- [15] H. A. Carneiro and E. Mylonakis, "Google trends: A web-based tool for real-time surveillance of disease outbreaks," *Clinical Infectious Diseases*, vol. 49, no. 10, pp. 1557–1564, 2009.
- [16] X. Zhou, J. Ye, and Y. Feng, "Tuberculosis surveillance by analyzing google trends," *IEEE Trans. on Biomedical Engineering*, vol. 58, no. 8, pp. 2247–2254, Aug 2011.
- [17] A. Seifter, A. Schwarzwald, K. Geis, and J. Aucott, "The utility of google trends for epidemiological research: Lyme disease as an example," *Geospatial Health*, vol. 4, no. 2, pp. 135–137, 2010.
- [18] J. A. Doornik, "Improving the timeliness of data on influenza-like illnesses using google trends," Tech. Rep., 2010.
- [19] L. Wu and E. Brynjolfsson, *The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales*. University of Chicago Press, May 2014, pp. 89–118.
- [20] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welp, "Predicting elections with twitter: What 140 characters reveal about political sentiment," *ICWSM*, vol. 10, pp. 178–185, 2010.
- [21] B. K. Kaye and T. J. Johnson, "Online and in the know: Uses and gratifications of the web for political information," *Journal of Broadcasting & Electronic Media*, vol. 46, no. 1, pp. 54–71, 2002.
- [22] A. Adams, R. Blundell, M. Browning, and I. Crawford, "Prices versus preferences: taste change and revealed preference," Mar 2015. [Online]. Available: /uploads/publications/wps/WP201511.pdf
- [23] D. L. McFadden and M. Fosgerau, "A theory of the perturbed consumer with general budgets," National Bureau of Economic Research, Tech. Rep., 2012.
- [24] D. J. Brown and R. L. Matzkin, "Estimation of nonparametric functions in simultaneous equations models, with an application to consumer demand," 1998.
- [25] D. Fudenberg, R. Iijima, and T. Strzalecki, "Stochastic choice and revealed perturbed utility," *Econometrica*, vol. 83, no. 6, pp. 2371–2409, 2015.
- [26] G. C. Chasparis and J. Shamma, "Control of preferences in social networks," in *Decision and Control (CDC), IEEE Conf. on*, Dec 2010, pp. 6651–6656.
- [27] B. Mangold, M. Dooley, G. Flake, H. Hoffman, T. Kasturi, D. Pennock, and R. Dornfest, "The tech buzz game [stock market prediction]," *Computer*, vol. 38, no. 7, pp. 94–97, July 2005.
- [28] Y. Chen, D. M. Pennock, and T. Kasturi, "An empirical study of dynamic pari-mutuel markets: Evidence from the tech buzz game," in *Proc. of WebKDD*, 2008.