A case study in complex systems evolution: consumer price obfuscation and mobile/cell phone tariff pricing

Les Hatton
CISM, University of Kingston*

September 3, 2005

Abstract

This paper takes a multi-disciplinary approach to consumer price obfuscation (confusion marketing) by combining global optimisation theory in the presence of uncertainties with a web-based client server architecture. Electronic networks make price information access easier and promote competition so it is interesting to speculate whether pricing becomes more complex to help suppliers keep ahead. Here, a case history of an Internet resource to find the cheapest mobile/cell phone tariff within supplied constraints is described. Data collected across approximately one thousand users present a compelling case that in the obfuscation arms race between consumer and supplier, the consumer is losing ground even against the background of static spending patterns.

1 Overview

The author's interest in this subject was first piqued by a visit to a high-street mobile phone [1] shop claiming to be independent of any of the tariff suppliers, in order to change tariffs for four mobile phones after experiencing unease about the costs of running the phones on their existing tariff, a business deal treating all four phones together. The salesperson went through some of the many possible combinations before recommending relatively emphatically one particular newly announced tariff which was "a really good deal". This is typical of the way consumer pricing is presented. In this case, the salesperson was simply and unconsciously associating the fact that a tariff was new with the idea that it must therefore be a good deal. As will be seen, the complexity of such tariffs in particular and Internet lubricated price competition in general is such that it is impossible to do any better without sophisticated modelling.

*L.Hatton@kingston.ac.uk, lesh@leshatton.org

[1]For US readers, a UK "mobile phone tariff" is synonymous with "cell phone plan". There are also strong similarities in the way they are priced but with some important differences. From here onwards in this paper, the phrase "mobile phone tariff" will be used interchangeably with "cell phone plan". Many other parts of the world use similar nomenclature and very similar pricing strategies.

The author wary of such unsubstantiated claims decided to take all the information home and calculate it based on the last two monthly itemised billings. The end product of several hours of patient calculation was that the salesperson was completely wrong and that one other and relatively old tariff the author tried as a control appeared significantly cheaper, (by a factor of around 30% !). Spurred on by this, the author decided to look at other tariffs before the realisation that there were then already over 80, each with many sub-options often changing rapidly with time rendering the whole thing simply infeasible by hand. Looking further afield, other areas particularly in financial services appeared similarly complex. At this point, the author decided to set up an experiment to measure the degree of obfuscation, estimate its rate of change and also estimate how much this affected the public at large. The experiment also acted as a feasibility study to see if the problem could be solved in a time considered reasonable in internet access terms as the browsing public has a notoriously short attention span.

Deliberate and accidental obfuscation It is a moot point just how much of the obfuscation in an area like mobile/cell phone tariffs is deliberate and how much is accidental. It is certainly to the benefit of the suppliers to obfuscate on the grounds that if there were ever a clear, unambiguous and easily worked out best deal, most people would take it given the global outreach of the web. News of a good deal can travel very quickly on an electronically connected social network. Deliberate obfuscation helps to muddy the water to avoid any decision being obvious and perhaps ensnaring a consumer to choose even a very over-priced alternative through inability to see through the layers of complexity. This might be considered somewhat cynical so it should also be noted that when things are this complex, accidental obfuscation is also likely when the supplier doesn't really understand the full complexity either. (There were a number of examples in the case history below where a 'special deal' was more expensive than the alternative it was supposed to replace and clearly accidentally so in some cases.)

It should perhaps be noted that accidental obfuscation is inevitable in complex systems designed by humans and is particularly visible in software controlled systems where there are few if any constraints on the imagination. Space does not permit pursuing this fascinating topic further but unnecessary complexity is a fundamental property of many modern IT systems and is major contributing factor to their high failure rate, ([1]).

Rate of change Another significant factor in obfuscation is the rate of change. As electronic connection becomes more ubiquitous and faster, the rate of change can be expected to accelerate. In fact, all modern social interaction appears to suffer from such acceleration. Today, the opportunity period for a new movie to make money is down to weeks if not days. In the web world, there is already an escalating arms race between pop-ups and pop-up blockers and scams involving identity theft evolve on an almost daily basis. In the study of mobile/cell phone tariffs reported here, exactly the same thing was observed and it rapidly proved necessary to update the tariffs or at least check they were up to date every month and it was a rare month when something did not need changing. The pace of change also accelerated near festive times such as Christmas where an almost weekly check had to be made to stay current. This accelerated rate of change coupled with the fundamental complexity caused by the presence of seven or so suppliers (in the UK) locked in deadly competition each offering bewildering numbers of options and add-ons means essentially that *only* a computational server using sophisticated mathematics was able to track them all and estimate the full effect of price obfuscation.

Scope of this experiment The desired properties of the system necessary to carry out this experiment are as follows.

- It is able to model the overall complexity
- It is able to model this in an interactive response time considered reasonable on the web which is almost certainly less than around 20 seconds
- It is able to handle fundamental uncertainties in the usage patterns of mobile or cell phones
- It is able to adapt to the underlying rapid rate of change in complexity
- It is web-based to allow data to be acquired and analysed quickly
- It is secure enough to withstand malicious use, denial of service attacks and so on, (even though a free service was being offered, the need for this will become apparent later)

Assuming it is possible to implement this, the goals of the experiment would be as follows.

- To confirm the presence of consumer price obfuscation
- To quantify it
- To estimate whether it is changing with time

A final goal of the overall experiment would be to identify other similar areas amenable to the same kind of treatment. The branch of mathematics necessary to deal with this kind of problem is generically known as global optimisation. In the following sections, the design of such a system will be described.

1.1 Global optimisation strategies

In essence, there are two possibilities for finding a global minimum in an (N+1)-dimensional function $F(x_1, x_2, ..., x_N, t)$ representing the cost of a particular plan or tariff at a particular time t:-

- Exhaustive searching. Here, every possibility is explored and the smallest one found.

- Non-exhaustive searching. Here, usually for reasons of time constraints, a systematically chosen subset of the range of F is searched in the hope that the global minimum or at least one close to it can be found. In the case of consumer product pricing, this is complicated by the fact that the optimisation surface is discontinuous and sometimes even multi-valued due to the arbitrariness with which tariffs are priced. (Multi-valued behaviour occurs when the supplier does not state in which order discounts are applied leaving the calculation fundamentally ambiguous.)

A further complication in finding such a global minimum is that the estimates of the independent variables $x_1, x_2, ..., x_N$ which represent parameters such as the number of minutes spent calling a user on a different network or the number of text messages received per month are themselves uncertain for various reasons. For example, a user might not have itemised billing or perhaps can not be bothered to work it out or may have just had an unrepresentative month's use. This is discussed further below but the algorithm was adjusted to cater for it.

Which of the above two approaches is used depends entirely on how much computation can be afforded within the available time. First of all, the time dimension is fixed as the rate of change of tariffs is of the order of weeks rather than minutes, (although with an eye to the future, this may not always be the case and the author can foresee a time when deals have lifetimes of minutes or even seconds.). The 'available time' depends on the medium on which searching is done. On an internet server, users may not be prepared to wait more than a few seconds, whereas running locally on their own PC, they might be prepared to wait rather longer if the benefit was worth it.

Early versions of this project were written in TclTk, [2]. However, an exhaustive search of the complete range of F() for UK mobile phone tariffs could take as much as a minute even on a relatively high-end server, which is far too long for the interactive web-based service envisaged here. As a result, the range was initially searched using a *simulated annealing* approach, [3]. Simulated annealing originally emerged from statistical physics but has been used for many purposes since such as minimising the lengths of electronic chip connections. In simulated annealing, values for tne (N+1) dimensional independent vector $(x_1, x_2, ..., x_N)$ are chosen randomly and the function value F() calculated. If F() is bigger than the current minimum, it is rejected. If F is smaller than the current minimum, it might be rejected based on the value of $e^{-\phi T}$, where $\phi > 0$ is a constant and T is the 'temperature', a value which gradually diminishes during the minimisation run. The idea here is that there is quite a high chance that a value will be rejected at the beginning of the run even if it is lower than the current minimum. This is to prevent the algorithm getting stuck in a local minimum whilst missing a deeper one somewhere else. The probability of 'jumping out of a local hole' gradually diminishes as the algorithm progresses. Such methods are important when the (N+1) dimensional optimisation surface is subject to such arbitrary change as deals come and go. To deal with uncertainty, a further sophistication was added. Each trial was done several times with a randomly generated close grouping to attempt to divine the shape of the minimum and decide whether it should be rejected or not.

Simulated annealing is very effective when the number of possibilities to be explored dwarves those for which there is time to try and initially this met with considerable success. However, just as an experiment, the computational server was re-written in C and the resulting increase in speed was so dramatic that an exhaustive search could be carried out in about half a second which was entirely sufficient for an interactive web request[2]. As a result, global optimisation strategies were not considered further but in order to provide a good interactive web-browsing response in future, it may well become necessary as obfuscation increases, (which it did inexorably throughout the duration of this study). For a good general discussion of a number of alternatives to annealing, see [4].

Degree of uncertainty As has been mentioned already, the degree of robustness of the solution is an important part of any optimisation strategy with noisy and inconsistent data of the nature of mobile/cell phone use. The accuracy to which a user can estimate their own phone use can be subject to quite wide variations. In addition, some of the necessary parameters, for example the percentage of all calls made between different network operators which in some regimes like the UK attract different rates, is usually unknown. The network operators do not reveal this information on their itemised billing. To cater for this, a degree of uncertainty was introduced into the user interface in plain English terms by asking for the degree of confidence between "Very confident", "Reasonably confident" and "Not very confident". These were used to generate corresponding degrees of randomness in the searching engine. For example, a deep narrow minimum might be rejected at a low level of confidence but accepted at a high level of confidence.

2 The minimisation criteria and problem parametrisation

The minimisation goal of finding the cheapest mobile/cell phone tariff for a given level of confidence in the input data is quite simple to implement but understanding the problem parametrisation is considerably more complex.

Complexities in mobile/cell phone pricing Close inspection of phone pricing reveals an already staggering level of complexity which appears to be accelerating, a wonderful example of complex systems evolution through deals, offered and accepted. The current complexity includes but is not limited to the following:-

- There are typically between 5 and 10 suppliers of tariffs or plans in most countries. Most suppliers have multiple plans and in the UK this led to initially 60 or so which grew to over 80 during the year of the study. In the US, there were about 120 at the end of the study.

[2]Even so, this corresponds to around 200 million arithmetic operations which bears silent testimony to the scale of complexity already present in this problem.

- Each plan has numerous and frequently different options, add-ons and volume related discounts.
- The plans change about every 3 months with increased change near seasonal periods like Christmas where the imagination of tariff or plan architects literally knows no bounds. Some changes are pathological with entire tariffs being scrapped and replaced with new-look ones. Sometimes changes are subtle with details remaining the same but prices dropping by some percentage or there may be a change in the 'small print' affecting the tariff or plan. The natural by-product of this is that suppliers in most countries must support hundreds and in some cases thousands of legacy plans.
- The treatment of text or SMS messages, a massive growth area, is very different between suppliers. Some suppliers offer bundles of messages which must all be used or forfeited, others offer graded bundles with step function pricing. These have on occasions been so complicated that the suppliers have accidentally made mistakes so that a 'special deal' is more expensive than the option it is supposed to be replacing.
- Multiple user tariffs allow the users to share a pool of speaking or text message time.
- Some tariffs treat all network calls as being equivalent whilst others use price differentiation.
- Some tariffs offer a flat rate for the first N minutes of use and then discounted rates for all or some calls thereafter.
- Special services such as sports scores and internet use are priced somewhat differently to normal speaking or text calls but not always.
- Some tariffs are monthly based with a minimum period whilst others are pay per use.
- Some tariffs do price differentiation on distance between the two subscribers, (for example India and the USA). Others like the UK do not.
- New features come and go with sometimes completely differing pricing structures to the above, for example for e-mails.
- Some countries (like the USA) price differentiate between the cost of sending a text (SMS) message and the cost of receiving one.
- Roaming charges particularly across country borders vary dramatically and some tariffs attempt to cater specially for this.
- The time period deemed 'peak time' is frequently different between suppliers and some suppliers further differentiate between evening and weekends. This can also vary within the same supplier for different tariffs.
- There is a rapidly growing 'middle-man' business area where companies buy time from the suppliers and sell on solutions for multiple phones to corporations.

- Network suppliers do not always state how discounts work for inclusive messages or talk time leading to multiple valued functions. (Consider for example having 100 free talk minutes but not knowing against which particular pricing level they could be redeemed.)
- Network suppliers have very different 'roll-over' policies. Some allow all unused inclusive text messages and/or call time to be carried over to the next billing period and some allow some or none.

This list is not exhaustive but should be sufficient to convince the reader that the complexity is already staggering and the chance of a salesperson giving reasonably correct advice is basically non-existent. At first glance, it would appear that the entire infrastructure is maintained by artificial complexity because if there was a simple and obvious comparison, most people would simply adopt it. In reality, the answer is probably much more complex. Increasing obfuscation may well be an inherent property of the way consumers and suppliers interact in a tightly-knit, highly efficient social network as provided by the Internet. There was certainly some evidence in this study that this was the case.

3 Architecture of a web service

The basic idea was to set up a simple web server which solicited enough target information from the user for a given level of use of a mobile/cell phone, either actual (an existing user) or planned (a new user) and some idea of how uncertain these estimates were in terms which the average user would not find too technical. In the end, a web service using PHP was constructed which acted as a client, driving a server written in C for efficiency which actually performed the calculations before passing them back to the PHP client for display. All the data was maintained in a simple database on the web server and downloaded periodically for analysis. The analysis was itself automated using simple shell scripts to extract, clean and plot the data.

Questions for the users Users were differentiated into two types:-

- Existing users: people who already had a phone and had a reasonable idea of how they used it. Such users would be solicited for their average monthly bill.
- New users: people without a phone who had only a vague idea of how they might use it and had no real idea of actual costs. Such users were not of course solicited for their average monthly bill.

In addition, both kinds of users were solicited for answers to the following questions:-

- Average number of minutes used per month for calls
- Average number of text (SMS) messages sent/received per month.
- Percentage of all calls made at peak times
- Percentage of all calls made to normal landlines

- Percentage of all calls made to the same network supplier
- Percentage of all calls made to a different network supplier
- Percentage of all calls made whilst roaming
- An estimate of the confidence in the above answers

Note that percentages were solicited as "None", "Some", "Half", "Most" and "All" to cater for any numerical phobia. Confidence was solicited as previously described.

It was found that initially 19 and later 22 variables was sufficient to give good robust parametrisation of the above complexity. These variables and their values were encoded as text files, one for each tariff with options for that tariff included within the text file. This allowed the tariffs to be updated easily and quickly to keep up with the rapid change in most cases. Even so, twice in the one year period, tariff changes forced upgrades in the way the optimisation engine actually handled the data as tariff/plan suppliers exercised their considerable marketing imagination.

3.1 Data handling

Before submitting to the computational engine, the data had to be edited. Some web-users (about 2% in the sample here) delight in entering crazy values for parameters. Obviously stupid ones were simply removed by comprehensive server-side checking. Furthermore, it was decided arbitrarily that anybody saving more than about 150 pounds (around 300 dollars) per month should probably think more about the danger of frying their brains than of saving money, so larger amounts were simply ignored rather than truncated. This removed another 4% or so.

3.2 Security

PHP was used to drive the computational engine. Numerous denial of service attacks were attempted even though this was a free service, another feature of electronically connected exchange systems which will only get worse. The service was fully check-pointed with mailers to show where and what users were doing so as the DoS attacks continued, the profiling was used to design counter measures. A combination of checking DNS values, putting up random words and asking users to input 2 randomly situated letters which were then compared with the internal PHP session management values and incorporating exponential queueing in the server response proved sufficient to keep these under control for the duration of the experiment. Such attacks are unfortunately however a fact of life with web-based services so vigilance must be maintained and the check-point mailers proved their worth again and again in identifying how attacks developed.

Malicious entries were occasionally entered into PHP fields to attempt to gain control of the server. These were protected against in the usual ways by checking all inputs for malicious content, attempts to execute arbitrary code on the server and attempts to include spoofed PHP scripts. Standard techniques for doing this are included with PHP, for example the clean() function which scans strings to be executed for a variety of attacks, and full use was made of these.

4 Results

First of all, the experiment clearly proves that the public at large is massively confused by mobile/cell phone charging and the average savings are extraordinary revealing the degree to which the public is vulnerable to such price obfuscation. Approximately 1000 unique users took advantage of the service over a 12 month period. Most of these involved the UK computational server but some early data from the US server is also shown for comparison.

4.1 UK mobile phone tariffs

The distribution of savings for each of the two half years the service was run in the UK is shown as Figures 1 and 2. Taken across *all* users, something like 80% could have saved around 300 pounds (around 600 dollars) *each* per year. Furthermore, it can be seen that the savings increase as time goes by and the phone charging schemes become ever more complicated. This represents an extraordinary and increasing amount of money. The rationale behind using half years is that the average contract is annual so approximately half the users would have time to react to the data and change contract in time for the second half year. This is important in the formal analysis later.

Sanity check Testing a computational server for a problem of this complexity is itself no trivial task so as a sanity check to make sure these were not hypothetical, the author changed tariffs (but not supplier) for four phones according to the predictions made by the computational server. The cheapest tariffs with the same supplier were selected as changes within the same supplier can be done without penalty usually in the UK. Although not optimal, average savings since changing have been 45% on the original tariffs putting it squarely in the middle of the average savings of the two half years, so actual saving and projected saving were in close agreement verifying the underlying accuracy of the searching engine. (Other feedback from switching users corroborated this.)

4.2 Analysis of results

To analyse the results, all users will be included, not just the majority who actually saved money. The results are shown as Table 1. The second half year was not complete which accounts for the smaller number of users. The density of users per month was very similar throughout the experiment. The figures in brackets after the average values are the standard deviations. It is clear that savings have jumped by almost 10 pounds per month whilst the average spend has hardly changed at all.

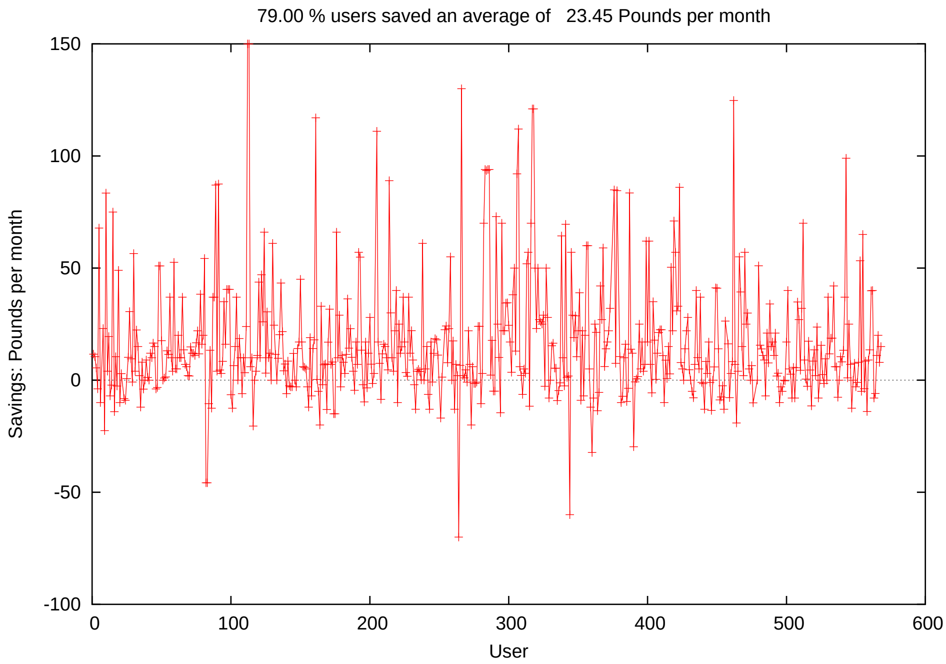

Figure 1: UK server: Savings per user, pounds per month, first half year (4Q03-1Q04)

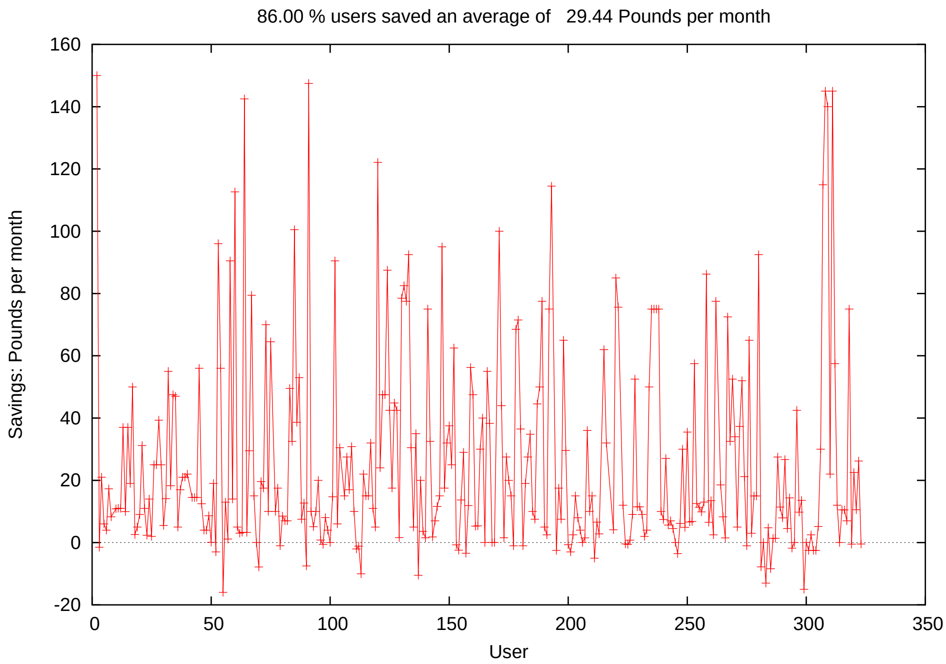

Figure 2: UK server: Savings per user, pounds per month, second half year (2Q04-3Q04)

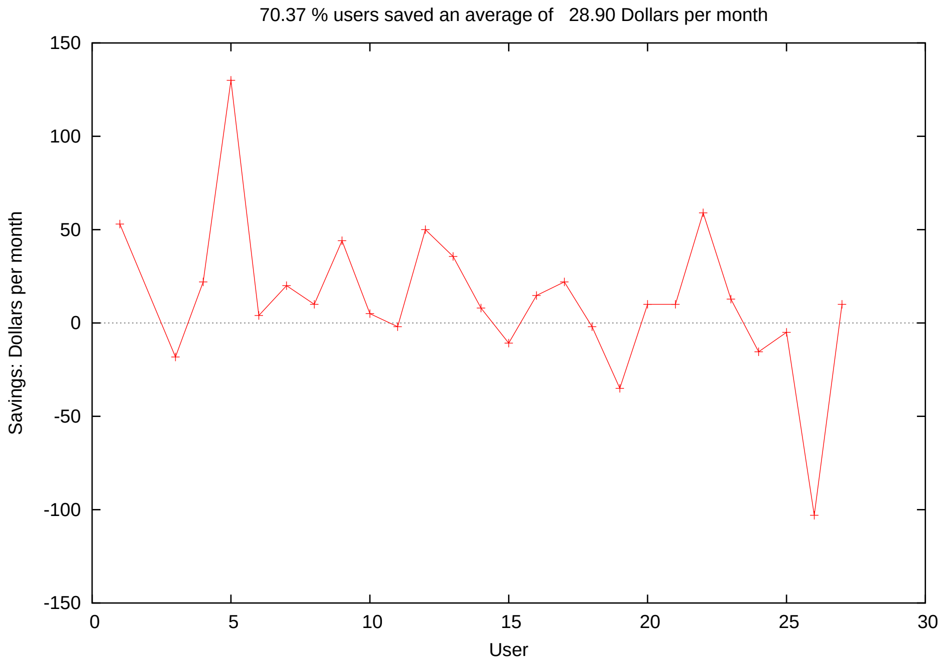

Figure 3: US server: Savings per user, dollars per month, fourth quarter 2003

Although apparently obvious in the table, the object of this analysis is to test the average spend and the average savings for changes with time using formal statistical methods.

Half year	Average actual spend per month	Average saving per month
First half year (545 users)	40.02 (29.67)	16.33 (27.08)
Second half year (306 users)	40.87 (30.68)	25.77 (32.07)

Table 1: Average actual spend and calculated savings per month per user

First the average spend will be tested for change. The following hypotheses will therefore be made:-

- H_0 The null hypothesis, the average spend in each half year came from the same population.
- H_1 The alternative hypothesis, the average spend in each half year did not come from the same population.

The data will be analysed using the z-test for the difference of means in a population, [5]. This states that the following statistic is approximately distributed as N(0,1).

$$z = \frac{\overline{X_1} - \overline{X_2}}{(\frac{(s_1)^2}{N_1} + \frac{(s_2)^2}{N_2})^{\frac{1}{2}}} \tag{1}$$

where $\overline{X_i}, s_i$ and N_i are respectively the sample means, standard distributions and number of samples for each of the half years. Substituting the numbers from the table above yields,

$$z = \frac{40.87 - 40.02}{(\frac{(29.67)^2}{545} + \frac{(30.68)^2}{306})^{\frac{1}{2}}} \simeq 0.39 \tag{2}$$

There is a probability of about 0.35 that this could have occurred by chance so there is no basis for rejecting H_0 and it must be concluded that there is no basis to reject that spending is static. However repeating the same calculation for the average saving using corresponding hypotheses yields:-

$$z = \frac{25.77 - 16.33}{(\frac{(27.08)^2}{545} + \frac{(32.07)^2}{306})^{\frac{1}{2}}} \simeq 4.35 \tag{3}$$

which is very highly significant, (the probability of this occurring by chance is $\lesssim 4.10^{-6}$).

The conclusion is emphatic. There is no evidence to suggest that the average spend has changed significantly but there is compelling evidence that the potential savings increased in the same period. The implication here is that over the two half years, the suppliers made spectacular gains in the arms race against the consumer whilst average spending remains static. Now it is certainly true that obfuscation increased in this period as three additional parameters had to be added to differentiate between the suppliers, (for example, the definition of peak period became more variable, starting and ending at different times with different suppliers) and in the UK alone the number of basic pricing plans increased by around 30%. Given therefore that the global outreach of the web means that the information is there for all to find and that the experiment ran long enough for users to switch from expiring contracts, the only conclusion seems to be that the increased level of obfuscation very effectively prevented the consumers from finding genuinely better deals.

4.3 US cell phone plans

The picture in the US appears very similar as can be seen by looking at Figure 3. The US site was only run for a short period as it proved too difficult to keep up with the vagaries of phone charging in two countries with the time the author had available, particularly as the Christmas period encroached. However, the data collected, some of which is shown, suggested a very similar picture in the US with around 70% saving an average of around 360 dollars per year each. As with many consumer products, the US follows an approximate "dollar for pound" rule when compared with the UK, (which is known, not without justification as "Treasure Island" amongst suppliers because of the high costs of many products and services.)

5 Conclusions

An important aspect of networked exchange is the ebb and flow of information between consumers and suppliers as the consumers try to find the best deal while the suppliers try to maximise their profits. As the consumer world becomes ever more competitive and electronically efficient, options for the same service increase and complexity increases even more dramatically. In this case history, the arms race greatly favours the supplier and the effect on the average consumer is that they finish up paying much more than they need simply because of the level of obfuscation.

This paper also demonstrates clearly by formal statistical analysis that consumer price obfuscation is growing as electronic networked exchange systems become more efficient, that the consumer is currently losing ground and finally that it is entirely feasible to track this complexity and to control it to the consumer's benefit using standard mathematical techniques deployed in novel ways as servers which anyone can access. The techniques are equally relevant to other areas in which consumers are vulnerable to overly-complex pricing regimes such as in many financial services.

As far as feasibility goes, the total resource time necessary to build this web-based system was around 20 days engineering and test time but the global optimisation engine could easily be reconfigured to handle other areas of price obfuscation, (the basic techniques were borrowed from a project in computational physics the author was working on in parallel). In the case of mobile and cell phones, about another day or two per month was necessary to keep up with pricing changes. This could not be done automatically because the web data were not accessible in any standard form such as XML so the data had to be mined largely by hand. This was feasible in this case because there are not that many suppliers. In other areas such as loans, credit card deals, utility deals and share trading where there are usually many suppliers, the data mining necessary to feed the searching engine would likely be a much bigger undertaking. However the overall feasibility and value is amply demonstrated here as is the accelerating nature of the price obfuscation arms race.

References

[1] Royal Academy of Engineering (2004) *The challenge of complex IT projects*, Royal Academy of Engineering report, London ISBN 1-903496-15-2

[2] Tcl Tk http://www.tcltk.org/

[3] Kirkpatrick S, (1984) *Journal of Statistical Physics*, vol 34, pp 975-986

[4] Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. (1999) *Numerical Recipes in C*, Cambridge University Press ISBN 0-521-43108-5

[5] Spiegel M.R. and Stephens L.J. (1999) *Statistics, 3rd edition, (Schaum series)*, McGraw-Hill, New York.