# COMBATING ETHICAL ISSUES IN UNIVERSITY ADMISSIONS USING TECHNOLOGY

By

MICAEL S. GEORGE JR. * CAIO V. SOARES ** JUAN E. GILBERT ***

* Undergraduate Student, University of Connecticut.
** Ph.D. Candidate, Computer Science and Software Engineering, Auburn University.
*** Centered computing, School of Computing, Clemson University, Clemson.

*ABSTRACT*

*Obtaining diversity among admitted applicants is often a challenging task for most post secondary institutions. As a proposed solution to this challenge, Applications Quest (AQ) was created. AQ addresses the dilemma of how to achieve diversity while still upholding institutional academic standards and objectives while adhering to the law. A steady rise in admission applications, however, has lead to steep competition for admission slots. This in turn, has compelled many students to act less than ethical when completing their applications, often lying to try to gain an unfair advantage. So, this study investigates if applicants could "game" the application process, to their advantage, under AQ. New applicants are provided with information about the current applicant pool and details of how AQ works. They are then instructed to complete a true application, with their factual information, and a fake application, in an attempt to game the system. This study finds that even when provided with the aforementioned information, the null hypothesis that "a student will not be able to increase their chances of being admitted to an institution by lying on their application in an effort to game the system" cannot be rejected, thus furthering the case for the use of AQ in the application selection process.*

*Keywords : Admissions, Diversity, Application Process, Applications Quest, Unfair Advantage.*

## INTRODUCTION

Since the landmark decisions of the Supreme Court of the United States on affirmative action, in the Regents of the University of California v. Bakke (Regents of the University of California v. Allan Bakke, 1978), Grutter v. Bollinger (Barbara Grutter v. Lee Bollinger, et alJeffrey Lehman, Dennis Shields, Regents of the University of Michigan, and The University of Michigan Law School, 2003), and Gratz v. Bollinger (Jennifer Gratz and Patrick Hamacher v. Lee Bollinger, et alJames J. Duderstadt, and the Board of Regents of the University of Michigan, 2003), schools have been left with the dilemma of exactly how to use race and ethnicity in the admission application process without giving preference to one group over another. In accordance with the Court's opinion, many schools share the belief that diversity brings educational benefits to the entire student body, and therefore should strive to achieve it in their admissions procedures.

Although the goal of achieving diversity is clear, the process in which to achieve it is significantly unclear. First and foremost, institutions want to establish a holistic application process. The meaning of holistically accepting an application is the process by which institutions review student applications using both subjective and objective admissions criteria. Typically, a holistic evaluation uses race as one of many attributes as part of included in its decision, yet all attributes play a role and no single attribute is the determining attribute (Gilbert, 2006). Current practices for selecting applicants to a particular institution are based on a holistic review of the applicant's attributes, which include everything from a student's high school GPA, major, ethnicity/race, to standardized test scores. However, traditional holistic methodologies are vulnerable to legal challenges because of their inherent subjectivity (Gilbert, 2009). The element lacking from current practices in the admissions process is a standardized and objective conventional procedure which each institution can follow. When the Supreme Court ruled in favor of affirmative action, they left a very broad and ambiguous idea of how to go about following the court's ruling while still being fair to all

applicants.

Generally, one of two processes is employed when selecting applicants for admittance to a particular institution.

*Manual Review*

The first, and most commonly used method, is the use of a manual review. In this process, a committee essentially reads each application, meets, and then recommends applicants for admittance. The chief discrepancy in this method is the unclear and varied processes carried out at different institutions. It is possible that not all of an applicant's attributes are taken into consideration equally, especially since human input is the sole means for selection. Furthermore, this process is extremely costly with respect to time spent on making decisions. This process can take weeks if not months to complete.

*Descriptor PLUS*

A second method, which is becoming rapidly popular among institutions, is the use of Descriptor PLUS Software. Descriptor PLUS is a software program developed by The College Board in an effort to conform to the Supreme Court's ruling. The software is a geo-demographic tagging service that identifies and groups students into neighborhood and high school clusters according to neighborhood and high school respectively (Board, 2009). Although this method is very intuitive, the problem still remains that human analysis plays a major role in the selection of applicants. Moreover, this process has high time costs because the applications within those clusters must be manually evaluated and the committee must reach a common consensus on the recommendations for admittance.

*Applications Quest*

A third alternative has been proposed is the applications Quest. The Applications Quest software holistically selects applicants in a diverse manner which can be used in place of, or interchangeably with, the aforementioned methods. It is a software program that groups an institution's applicants into numerous clusters, each containing students with similar backgrounds and qualifications. The students are grouped using a broad range of common attributes rather than a single criterion, such as race, ethnicity or geo-demographic attributes to name a few. The admissions team determines which attributes will be used in the clustering process and Applications Quest creates clusters using holistic comparisons between all of the applications. The Next, the admissions officer will tell Applications Quest to recommend a specific number of applicants for admissions. The tool will then create the specified number of clusters and recommend one applicant from each cluster. The applicant that is holistically the most unique will be recommended by the tool. Applications may then be selected from each cluster to ensure a diverse group of accepted applicants. Unlike the holistic approaches previously described, Applications Quest uses a model that is objective and reproducible and therefore free of bias. This approach adheres to the Supreme Court's ruling since it does not allow for preference to be given on the basis of race, ethnicity, gender or national origin (Gilbert, 2006) (Gilbert, 2009).

Another challenge surrounding college admissions has been the steady rise in competition for slots. Simply put, the supply of spaces has not kept pace with demand. A recent study at Harvard University and University of Michigan Ann Arbor reported declining acceptance rates at 40 highly ranked private and public institutions, which saw the proportion of applicants they admit fall by an average of about 25 percent between 1986 and 2003 (Bound, Hershbein, & Long, 2009). The authors also devised a way of estimating how the same applicant would have fared over time. They found that, in general, the likelihood of being admitted to a four-year institution fell nearly 9 percent from 1972 to 2004. The largest declines were among "low ability" students; those in the two lowest quintiles saw their odds decline by 43 percent and 23 percent, respectively (Hoover, 2009). This is a capacity issue. There are more qualified applicants than available admissions slots; therefore, by definition, a qualified applicant must be denied admission. This opens the door for legal challenges, whereby a determined and resourceful qualified applicant that is denied admission can legally challenge the admissions committee's

decisions. As such, it is critical that admissions decisions are justifiable, reproducible and transparent.

So, even at a time when competition for slots is at its highest and slots are at a premium, Applications Quest still allows institutions to achieve a diverse group of admitted applicants following in line with school's academic standards and objectives, without giving preference to any racial group or ethnicity in admissions decisions. This has been accomplished, however, under the assumption that the information provided by applicants has been factual and true. Up to this point, Applications Quest has not considered applicants who may have lied in their admission application.

Lying on applications is nothing new to admission departments. Given the progressively higher competitive state of admission, it is no surprise either. A web search of "lying on college application" returns a large number of forum entries on sites such as *Yahoo! Answers* and *College Confidential* of students essentially wondering how much they could stand to gain and how much trouble they could get in if they lied on their admission application. Moreover, although the rate at which this happens is unknown, cases of a student or cluster of students being caught do regularly appear. Penalty can be as light as the student's application simply being rejected, such as the case of five applicants at the University of Pennsylvania's Wharton Business School (Roller, 2003), or as severe as the student facing criminal charges, such as the case of Akash Maharaj at Yale (Arenson, 2008). Lying on an application has become so widespread that college counselors have even reported receiving phone calls suggesting certain students are lying about accomplishments or extracurricular activities listed on their admissions application (Montgomery, 2003).

Due to the growing popularity of Applications Quest, it is important that ongoing case studies are held to ensure the tool's efficiency and robustness, even under less than ideal and dishonest circumstances. So, this study was comprised with the purpose of testing whether students are able to increase their chances of being admitted to an institution by lying on their application, in an effort to game Applications Quest.

Thus, the remainder of this paper will be outlined as follows. Section 1 will detail the theory and methods utilized in this study. Sections 2 and 3 will discuss the experiment, and analyze its results, respectively. Then, finally concluding remarks are added.

### 1. Theory and Methods

Applications Quest is a software tool used in the aid of selecting applicants in a diverse, unbiased, and holistic manner. Applications Quest's software algorithm is comprised of two fundamental theories that together assist in the analysis and selection of applicants: clustering with Euclidean distance.

Clustering is an essential part of how Applications Quest goes about holistically evaluating applications. In clustering, the goal is that the objects within a group be more similar to one another and different from the objects in other groups (Tan, Steinbach, & Kumar, 2006). Essentially, the dataset is partitioned into subsets (clusters), so that the data in each subset share some common traits; often proximity correlates according to some defined distance measure (Wikipedia, 2009). Clustering algorithms can be put in one of two categories: hierarchical and non-hierarchal.

Hierarchical clustering methods create clusters or groups by merging or dividing. These actions may occur in one of two forms: agglomeration or division. Agglomerative clustering methods form clusters by merging individuals and begin by assuming each instance in the collection population is an individual cluster and then forming clusters by merging individuals. In the course of each processing cycle, two clusters are merged. This process continues until either there is only one cluster remaining that which contains all instances in the population, or some other predefined stopping point has been reached, such as a specified number of clusters. The divisive clustering approach works in the opposite direction. It starts by assuming that all instances belong to one cluster. In each step of the process, a cluster is split into two clusters, until all clusters contain a single instance, or some other predefined stopping point has been reached, such as a specified number of clusters (Gilbert,

2006). Applications Quest uses a divisive clustering approach to group applications.

One of the main advantages of non-hierarchical methods over hierarchical methods, however, is that the former typically result in faster execution times. The most common non-hierarchical method is *k*-means. Before the *k*-means algorithm can be executed, the number of clusters is specified, which is *k*. The algorithm begins by selecting *k* instances from the dataset and assigns them as centroids. A centroid is the most representative instance within a cluster. It is the instance within a cluster that has the shortest distance from all the other instances within the cluster. The centroid instances are typically selected at random or by utilizing some heuristic. Much like the divisive approach, all the remaining non-centroid instances are compared to each centroid. The non-centroid instances are placed in the cluster with the most similar centroid. At the end of each cycle, the centroids are recalculated for each cluster and the instances are redistributed until the centroids do not change. There are several variations of *k*-means, such as bisecting *k*-means, but they all follow a slight variation of this basic approach (Gilbert, 2006).

Once clustering is complete, clusters are compared with the use of a distance or similarity method. These measures can be calculated using a variety of methodsways, but the Euclidean distance is the most commonly used one (Gilbert, 2006).

Euclidean distance is based on Pythagoras' theorem, where instances are represented as points in an *n*-dimensional space. The distance between any two points, *p* and *q*, in an *n*-dimensional space is calculated as the square root of the sum of the squared sides between the two points along each dimension (Equation 1).

$$D(p,q) = \sqrt{\sum_{i=1}^{n} (pi\text{-}qi)^2} \qquad (1)$$

Euclidean distance measures are used by clustering algorithms to determine distance or similarity, yielding a basis for comparison between instances, or objects, with the same attributes/characteristics (Gilbert, 2006).

Through the use of a hierarchical clustering algorithm (divisive clustering), in addition of the use of the Euclidean distance, Applications Quest is able to holistically evaluate applications using the applicant's attributes and group (cluster) those applications based on their similarities, forming holistically diverse applicant pools of qualified applications.

## 2. The Research Study

### *Motivation*

Given the potential impact of Applications Quest in many institutions, it is imperative to test its performance and efficiency, even under less than honorable situations. The following section will describe the methodology and approach used to assess the null hypothesis, aforementioned in the abstract, which states that "students will not be able to increase their chances of being admitted by lying on their application in an effort to 'game' the system".

### *The Study*

In order to test the aforementioned hypothesis, a group of 20 applicants, all 19 years of age or older, were asked to read and agree to a consent form detailing the study, then go online and fill out two separate online applications, each containing 25 application attributes and identical to the one used by Auburn University for undergraduate admissions. The consent form informs participants of the study and of their rights as research participants. Next, the participant completes the first application with all factual information. After the first application, participants learn how Applications Quest selects applicants and are further given detailed demographic information about 2,500 other applicants in the applicant pool. Given this information, participants are then presented with a second application, which they are allowed and encouraged to lie on, in an effort to increase their chances of being selected by Applications Quest.

Once the participants have completed both applications, data from the first online application, also known as the control group, combined with the other 2,500 applications, known as the base group, is processed through Applications Quest to see which applications are recommended for admission for 434

slots. After the applications have been processed and analyzed, the control group is removed from the applicant pool, leaving just the original base group.

Next, the data from the second online application, also known as the experiment group, is combined with the base group. These applications are then evaluated through the same process as the control in Applications Quest. The software once again selects applications for 434 slots. After this portion of the study has been executed and analyzed, the results are compared and analyzed in order to see if the chances of an applicant are improved from the control group to the experiment group as a result of the research participants trying to game the system by lying on their application.

### 3. Data & Analysis

Twenty participants took part in this study. Based on the experiment described in Section 2, the findings are as follows.

The first evaluation of the data began by the processing of the applications control group, allowing Applications Quest to use a minimal set of 5 attributes to select applications. This evaluation found that 16 participants of the control group were recommended for admissions by the software. The experiment group was evaluated next. From the same 20 participants, 19 were recommended for admissions by Applications Quest. Moreover, of those 19 applications accepted, 16 were of the same alias that were accepted in the control group processing, leaving only three of the experimental group that actually managed to "game" their way in.

Following these results, a t-test was used to measure statistical significance. It was found that the 't' test yielded a p-value of .041 was found, at an alpha value of .05. Thus, it can be concluded that in this experiment with 5 attributes and 20 applications, there was a correlation between students lying on their applications and being accepted by Applications Quest. However, it also states true that the correlation is so small that it is scientifically trivial.

Upon a second evaluation of the data, Applications Quest was allowed to use a larger set of attributes (7), for which to select applications. From the control group, 17 participants were recommended for admissions by the software. From the experimental group 18 participants were recommended for admissions by the software. In addition, of those 18 participants from the experimental group, 17 were of the same alias accepted in the control group. This leaves one participant from the experimental group who actually managed to "game" the system.

Again, the results were analyzed using a 't'-test to measure statistical significance. From this 't'-test, a p-value of .289 was found, at an alpha of .05. This p-value is greater than the alpha value of .05, therefore the test failed to reject the aforementioned null hypothesis which states that "a student will not be able to increase their chances of being admitted to an institution by lying on their application in an effort to game the system".

### Discussions and Conclusion

In Applications Quest, holistic review and diversity are achieved by creating distinctive groupings or clusters of applications which share a high degree of similarity. Since each grouping/cluster represents a distinctive pattern of similarities across a set of attributes (relative to other clusters), clusters form the basis of diversity in Applications Quest (Gilbert, 2009).

At the time this article was written, no other software applications exist on the market, which achieve what Applications Quest achieves, making this study unique and the first of its kind. The potential significance of this research study is the emergence of an objective, explainable, measurable, and reproducible way of selecting candidates for admissions with significant reductions in time.

Based on the data collected and analyzed in this study, the following conclusions can be reached. First, statistically significant evidence was not found which allows the null hypothesis that "a student will not be able to increase their chances of being admitted to an institution by lying on their application in an effort to game the system" to be rejected. Although the first data evaluation produced somewhat contradicting results, conclusive results, which support the aforementioned statement,

were found by simply allowing Applications Quest to utilize a larger number of application attributes. As such, the findings presented here suggest that using Applications Quest with a higher number of attributes will reduce the likelihood of students gaming the system or admissions officers inserting bias into the admissions decisions.

References

[1]. Arenson, K. W. (2008, April 10). *Yale Student Is Accused of Lying on Application*. Retrieved October 1, 2009, from New York Times: http://www.nytimes.com/2008/04/10 / education/10yale.html

[2]. Barbara Grutter v. Lee Bollinger, et alJeffrey Lehman, Dennis Shields, Regents of the University of Michigan, and The University of Michigan Law School, 539 U.S. 306 (123 S.Ct. 2325 June 23, 2003).

[3]. Board, T. C. (2009). *Enrollment Planning Services User Guide*. Retrieved October 1, 2009, from The College Board: http://www.collegeboard.com/eps/help/descriptorplus/index.html

[4]. Bound, J., Hershbein, B., & Long, B. T. (2009). *Playing the Admissions Game: Student Reactions to Increasing College Competition.* NBER Working Papers 15272, National Bureau of Economic Research, Inc.

[5]. Gilbert, J. E. (2009). *Applications Quest*. Retrieved October 1, 2009, from Applications Quest: http://www.Applicationsquest.com

[6]. Gilbert, J. E. (2006). Applications Quest: Computing Diversity. *Communications of the ACM , 49* (3), 99 - 104.

[7]. Hoover, E. (2009, August 24). Admissions Competition Heats Up, but Does Pressure Help Students? *The Chronicle of Higher Education - Admissions and Student Aid*

[8]. Jennifer Gratz and Patrick Hamacher v. Lee Bollinger, et alJames J. Duderstadt, and the Board of Regents of the University of Michigan, 539 U.S. 244 (123 S. Ct. 2411 June 23, 2003).

[9]. Montgomery, M. (2003, October 23). *How Competitive is College Admissions? Enough to Tempt Parents to Behave Badly*. Retrieved October 1, 2009, from Great College Advice: http://greatcollegeadvice.com/how-competitive-is-college-admissions-enough-to-tempt-parents-to-behave-badly/

[10]. Regents of the University of California v. Allan Bakke, 438 U.S. 265 (98 S. Ct. 2733 June 28, 1978).

[11]. Roller, J. (2003, July/August). *College rejects applicants for lying*. Retrieved October 1, 2009, from Group: http://findarticles.com/p/articles/ mi_ qa 3835 / is_ 200307/ai_n9252141/

[12]. Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining.* Addison-Wesley.

[13]. Wikipedia. (2009, October 9). *Cluster analysis*. Retrieved October 9, 2009, from Wikipedia, the free encyclopedia: http://en.wikipedia.org/ wiki/Data_ clustering

## ABOUT THE AUTHORS

*Micael S. George Jr. is a senior in a duel degree program, EuroTech, at the University of Connecticut. He is graduating in May 2010 with a B.S.E in Computer Science Engineering and a B.A. in German Studies. Micael has done research with REU groups at Lehigh University, Auburn University, and the University of Connecticut. He has also interned at the New York City Department of Transportation, Bureau of Engineering Review & Support, Division of Bridges. Recently, he studied abroad and interned in Stuttgart, Germany at Stuttgart University and eCCOMES GmbH. After graduation he intends on pursing opportunities in the Field of Technology.*

*Caio Soares is a Ph.D. candidate in the Computer Science and Software Engineering Department of Auburn University, in Auburn, Alabama. His research interests include Machine Learning, Data Mining, and Evolutionary Computation, focusing on real-world and human centered problems. He has earned B.S. degrees in Computer Science and Mathematics from Berry College and an M.S. degree in Computer Science from Auburn University. He is a SREB Fellow and a recipient of the Google Hispanic Scholarship.*

*Juan E. Gilbert is a Professor and Chair of the Human Centered Computing Division in the School of Computing at Clemson University where he leads the Human-Centered Computing (HCC) Lab. Dr. Gilbert has research projects in spoken language systems, advanced learning technologies, usability and accessibility, Ethnocomputing (Culturally Relevant Computing) and databases/data mining. He is a National Associate of the National Research Council of the National Academies, an ACM Distinguished Speaker, a member of the IEEE Computer Society Distinguished Visitors Program and a Senior Member of the IEEE Computer Society. He has a B.S. degree in Systems Analysis from Miami University, M.S. and Ph.D. degrees from the University of Cincinnati.*