Information Everywhere.. But not a Drop to Think

Today we are surrounded by all sorts of information – if fact most of use think that we are drowning in the stuff. An IBM report a few years ago estimated that only 4% of machine stored data is *ever* used. The explosion in the amount of machine-stored information is creating a problem: the technologies for analysing data have not advanced at a commensurate rate. In fact that is the problem; it is not information that is causing our problem, it is having raw data. If we could use expert systems to transform the raw data into information then we would have a solution – information is so precious that we cannot have too much of the stuff.

What we need is a means of turning this raw data into Information. Data Mining is the name given to this process of searching large databases for useful information. Supermarkets use this approach to see what loyalty card holders are buying and what combinations of products are bought together. American Express mines its customer data so that appropriate 'bill-stuffer' advertisements are included with customer statements. People who use their AMEX card to buy wine will receive wine promotional material, and so on. There is also "profiling" of purchases for security reasons. If for example you regularly buy petrol, train tickets, hotel reservations etc. the profiler would look upon a purchase of 50,000 litres of beer as unusual and flag this to the security personnel. Equally a sudden rash of purchases for relatively low value electronic items would cause similar concern – these are easily dispersed "for a few quid" at the local pub!

Internet shopping clicks are a source of data for internet vendors and a great deal of work now goes into the analysis of online 'clickstreams'. This can be expanded to an integrated view of all information about a customer, at all stages of their relationship with the company from initial contact onwards. This is called electronic Customer Relationship Management, (*e*-CRM).

However, marketing is not the only use for data mining:

After a couple of bad losses in the first two games of the 1997 season, Tom Sterner, Assistant Coach for the Orlando Magic basketball team turned to IBM to develop a data mining application to analyse the team's performance. The software discovered relationships between the player's behaviour. When Brian Shaw and Darrell Armstrong were in the game, the behaviour of their team-mate Penny Hardaway changed. (Hardaway was the Magic's leading scorer at that time.) Armstrong received more play-time and hence Hardaway was far more effective. Orlando Magic went on to win the next two games and millions of additional dollars in gate money and sponsorship.

Terrorism is, unfortunately, high on everyone's list of things to worry about these days, but technologies related to data mining may help to capture terrorists. Work is being done on image analysis software which can examine two-dimensional photographs and measure facial features, such as the separation of the eyes and the angle of the nose in order to aid identification. In future, this technology may be a integral part of airport security and town centre CCTV surveillance systems, making it difficult for terrorists and fugitives from justice to move around the world. CCTV is now a large industry employing substantial numbers of people to watch TV monitors. Even partial automation of this process would yield substantial benefits in a number of industries. Imagine if a town centre CCTV system drew the attention of police officers to a man walking the street who bears a strong resemblance to unknown man filmed weeks earlier robbing a bank.

An early example of such systems is an automated video system that will be launched in the UK in 2002 which will read car registration numbers and check if the car has a valid tax disc and insurance, or appears on a list of stolen vehicles.

Recently an American security group launched a video system that can tell when people are lying. Apparently the autonomic flow of blood around the eyes changes when people are lying and this can be detected as a change in temperature. This could become the non-invasive but always present security system at airport check-ins. "Have you packed these bags yourself, sir?" And "Are you carrying any weapons that you intend to use to harm others?" etc. could become serious and meaningful security questions. However, the next question of "And is this your wife that is travelling with you, sir?" might cause problems….

For democratic countries there is the difficult problem of balancing personal privacy on the one hand and the needs of the law enforcement and intelligence agencies on the other. I think it is better to try and possibly make mistakes than do nothing in case it goes wrong.