

Journal of Technology Management for Growing Economies

Volume 1 Number 1

April 2010

HP Labs India's Technology to Make Printed Paper Documents Tamper-Proof

Ramani Srinivasan
Anjaneyulu Kuchibhotla
Godavari Srinivasu
Bhushan Matad
HP Labs India, Bangalore

HP Labs India's Technology to Make Printed Paper Documents Tamper-Proof

Ramani Srinivasan¹
Anjaneyulu Kuchibhotla
Godavari Srinivasu
Bhushan Matad
HP Labs India, Bangalore

Abstract

The objective of this paper is to discuss the cost-effective technologies developed at HP Labs India to make paper documents tamper-proof. We started with a real customer problem: to ensure that printed government land records should be tamper-proof. We first looked at existing techniques to see if they could address the problem but subsequently had to come up with our own technology. To address the problem, there was a need to identify good pilot customers to promote adoption of technology. Another challenge was around the customer's lack of awareness in terms of the impact of fraudulent documents on their business, and the difficulty of selling solutions versus selling products. Technology invention plays a small part in the final success of any customer driven innovation, and so other factors should not be ignored. There were several other specific conclusions related to the innovator, the customer-facing staff and the customers. The key contributions of this work was starting with a real customer problem and dealing with end-to-end issues related to adoption of technology to make paper documents tamper-proof.

Keywords: Paper document fraud, security of paper documents, document authentication.

INTRODUCTION

Fraud related to paper documents is quite prevalent all over the world but is particularly acute in emerging economies. This paper covers a few cost effective technologies developed at HP Labs India to make paper documents tamper-proof and focuses on our experiences in taking these technologies to the field. We believe several of these experiences would be relevant also in the context of other technologies that are being targeted at emerging economies. Our experience substantiates the fact that technology innovation is only a small part of the successful deployment of technologies. There are many other factors that are equally important, if not more so.

Predictions of the demise of paper records, originally made in the early seventies of the last century, were premature, and continue to be so. Documents are created for every important event in a person's life, starting with the birth certificate and ending with the death certificate! Falsification of paper documents occurs at all levels in society in all countries. For instance, the following news items from an Indian newspaper illustrates just one of many instances of paper related fraud in emerging economies:

Journal of Technology
Management for
Growing Economies
Vol. 1 No. 1, April 2010
pp.9-22



©2010 by Chitkara
University. All Rights
Reserved.

Srinivasan, R.
Kuchibhotla, A.
Srinivasu, G.
Matad, B.

"About encroachment on 1,821 hectares of land, (Revenue Minister Mr. Jagadish Shettar) said there are 337 instances of creation of false documents covering 1,059 hectares (2,617 acres)" - The Hindu (Hindu, 2006).

Clearly, these problems are not unique to emerging economies. They occur even in advanced countries known for their extensive use of computers and networks. For instance, the following news item refers to the Dean of a major university in the US falsifying degree certificates:

10

"MIT dean resigns after falsifying degrees" - Reuters (Reuters, 2007).

Increasing numbers of paper documents are being issued to the persons concerned and are being relied upon by third parties, who know that there is likely to be falsification of these documents. For instance, companies hiring employees depend upon printed educational records, employment records, proof of previous income, etc. They need to spend significant time and effort verifying these or take a risk, go ahead and then punish the employee subsequently if any information produced is found to be false. There are many other areas where false documents are produced, for e.g. property records, insurance policies, birth certificates, etc. Such falsification leads to several problems some of which are listed below:

- **Substantial monetary losses:** A Times of India news item (Times of India, 2002) reports a good example: "Hyderabad: The Abids police on Thursday arrested a man who doctored land records and procured bank loans in crores² by mortgaging them". These losses are typically assumed to be a part of the cost of doing business and are covered indirectly by other customers having to pay more.
- **The impact goes far beyond monetary loss:** The follow-up enquiry and punitive action, including legal steps, cost a lot but are typically hidden costs of doing business. These indirect costs include loss of management and staff time as well as disruption of normal work. It may also result in ungrounded suspicion of one's own staff.
- **Diversion of benefits:** For instance, in India, falsification of caste certificates which are used to enforce positive discrimination leads to diversion of those benefits to others who are not entitled to them.
- **Potential security risks:** Falsification of records such as a birth certificate or educational certificates can lead to security risks, for instance enabling people to get a passport when they do not qualify for one. One false

document is often adequate to use to procure other 'authentic' documents issued by the appropriate authorities.

- **Delays due to verifying documents:** Because of the risk of document falsification, considerable delays are introduced in processing documents; since their authenticity need to be checked for every transaction. So even if just 1% of the documents are likely to be false, 100% need to be verified.

It is therefore clear that the effects of document falsification are severe, often underestimated and in many cases even unquantified. Our attempt was to overcome these problems in a cost-effective manner by coming up with cost-effective technology.

THE PROBLEMS WITH PAPER

Unlike electronic documents, paper documents have not evolved. Electronic documents can be digitally signed, making it impossible for anyone to make alterations without being detected at the receiving end; this involves the use of cryptographic techniques (Schneier, 2007). One copy of an electronic document maintained on a database can be declared to be the original, available to all authorized. Copies can be obtained very quickly and at low cost and sent out electronically. They can be encrypted, ensuring privacy. They can be stored and located quickly and accurately. Archives can be replicated, eliminating the risk of damage by fire, theft, etc.

On the other hand, new technology has only made it easier to forge and manipulate paper documents. Desktop publishing offers a variety of techniques that can help replicate a document. High quality color scanning makes it possible to copy parts of an image on paper and incorporate it into another document. High quality images of signatures and seals can be cut and paste with relative ease. Falsified documents can be printed using good quality color printing. Equipment and skills necessary for doing all this is widely available. For instance, counterfeiting currency notes using these techniques has often occurred and it has taken considerable effort on the part of governments to prevent such misuse of printing technology.

In the field of education, forged certificates are quite common. Many large universities are looking for solutions to prevent forgery of mark transcripts and degree certificates. Simple solutions like putting marks obtained by candidates on a website are not satisfactory since they compromise the privacy of the individuals involved.

A basic problem is the absence of interfaces between the databases of enterprises. For instance, consider fraud involving vehicle loans from banks. Falsified bank statements in one bank's name are sometimes given to another

Srinivasan, R.
Kuchibhotla, A.
Srinivasu, G.
Matad, B.

bank to get a loan to finance the purchase of a vehicle. There is no easy way for one bank to verify the bank statements of another bank. Another example is that the income tax department cannot verify bank statements. Creating interfaces between unrelated entities like that would create major privacy problems. This is also true of interfaces of government units with other entities. This is one reason why paper documents continue to be used in significant numbers.

12

One therefore sees that inspite of the problems with paper it will continue to persist for a long time to come. Our work aims to make paper documents tamper-proof and to overcome some of their limitations.

APPROACH/METHODOLOGY ADOPTED FOR THE WORK

We started our work with a specific customer problem: to ensure that printed government land records should be tamper-proof. We first looked at existing techniques to see if they could address the problem but subsequently had to come up with our own technology since none of the existing solutions met our requirement. Similar customer problems also exist in the insurance (for e.g. policy documents) and education (for e.g. certificates) sectors. We looked at the problem of enterprises and customers who use paper records. We also looked at the rationale for working to make paper records machine readable and more secure, comparing these techniques with totally computerized and networked solutions, particularly in the context of countries like India.

Most existing technologies require either special ink, paper, printers or capture devices and therefore require a change in the ecosystem. We wanted technology which could fit into the existing ecosystem. Some techniques for making the document more secure require that the paper document be machine readable. This opens up relatively seamless interfacing of the paper world with the electronic world. If this can be achieved, one government unit or enterprise would not have to get relevant data re-entered, and instead could take it from paper documents issued by other entities. We developed prototypes to make paper documents machines readable and secure and gave demonstrations to potential users. This was followed by contacts with business groups within HP, and by the effort to build a hardware prototype. This hardware prototype was later demonstrated to customers.

We found an early adopter, the International Institute of Information Technology (IIITB). This institution ran a pilot during the years 2007 and 2008, affixing two dimensional barcodes on mark sheets using our

technology. Anyone who had one of these mark sheets in hand could scan it and send the image for verification to a server at IITB. The server would read the barcode and send a copy of the authenticated marks for the user to compare the printed transcript against. The pilot results were very satisfactory, but the impact was limited because of the small number of documents involved - less than 150 per year.

A sample copy of the IITB grade card along with the 2D barcode is given in the Figure 1.

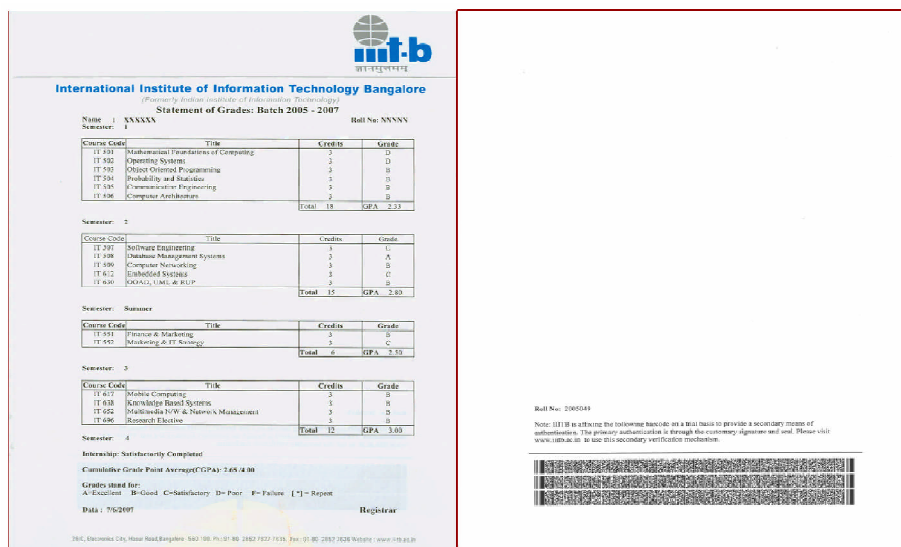


Figure 1: IITB Transcript with a Barcode to Support Verification

THE TECHNOLOGY

During our work we came up with a variety of techniques that evolved over a period of time. Most of the techniques were related to authentication of paper documents but some also looked at other related aspects such as checking if a paper document is original or not. Based on this work we have also been able to generate significant IP (US Patent and Trademark Office, 2009). We summarize some of the key techniques in this section.

We started with the observation that paper documents need to evolve and incorporate security techniques such as digital signatures. One way to do this is to start with the document before it is printed - this is easy, because the bulk of important documents are computer generated. The document can be digitally signed while it is still in the electronic form. Two dimensional barcodes exist that can be used to print a machine readable version of the

Srinivasan, R.
Kuchibhotla, A.
Srinivasu, G.
Matad, B.

document on paper, along with the human readable text. The standard we adopted was PDF417 (International Organization for Standardization, 2006). The electronic document encoded as a barcode may contain a copy of the text in machine readable form; even better, the barcode can carry information marked in a structured form, for instance in the form of XML documents, so that this data can be readily processed by a computer and/or stored in a database. The 2D barcodes can also incorporate information other than the printed text. For instance, the same sheet of paper may have human readable material in the form of conventional printed text along with a barcode containing the same information in a structured form for importing into any database anywhere.

14

We subsequently developed several variants of the technology involving, hashing and digitally signing the document, and imprinting it as a machine readable information on paper using a 2 D barcode. It is worth noting here that a digital signature is quite different from a digital image of a person's handwritten signature. The "digital signature" is a string of bits computed using the content of the document and the personal key of the signatory. Thus, the digital signature of a person will vary from document to document. If even a single letter, numeral or punctuation mark on a document is altered, the digital signature will no longer match the document.

We took one of the many solutions we had developed to the field. It involved using a scanned barcode containing a digital signature to point uniquely to an electronic document on a server without endangering privacy. The server will not give any information available for scanning and verification unless the owner makes the document available for online verification, thereby ensuring privacy. We also developed techniques for offline verification. This allows a document to be verified without access to any network.

The next step in our work was a centralized Document Authentication System (DAS) (see Figure 2) which serves many entities such as government departments or enterprises. The DAS can be run by an independent agency and can host a document authentication service on behalf of multiple document issuing agencies. Here, the Issuing Agencies do not need to share their original database/document information with the agency hosting DAS.

It would merely process a scanned 2D barcode and forward the extracted digital signature to the concerned database for verification. It is worth answering a question here: Why does one need a digital signature to point to a database record; will not a serial number suffice? The answer involves

technical and operational considerations. Assume that an employee looks up a remote database over the network and authenticates a document being processed using only its serial number. Proper audit requires a record to show that this authentication was actually carried out and that the information obtained was effectively used. Audit procedure has to rule out insider fraud where an employee falsifies the authentication record, which may not be traced for months. A digital signature on a paper document leaves a dependable and easy-to-verify trail which can be used for audit. This trail is independent of what was done over the network. Secondly, digital signatures can also be verified offline when a network connection is unavailable.

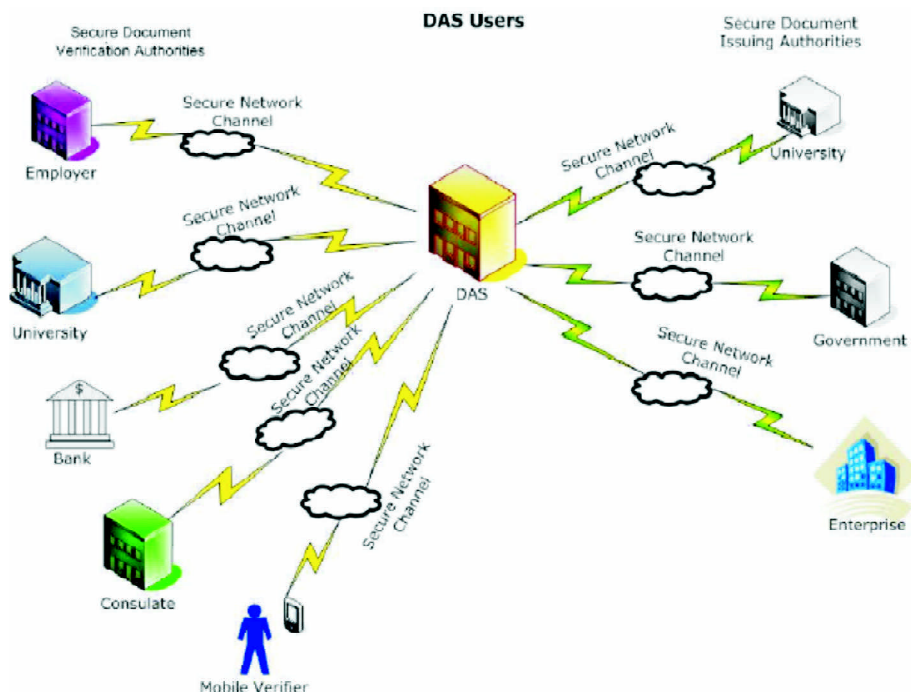


Figure 2: Document Authentication System Supporting Multiple Issuers and Verifiers.

There are important privacy issues related to verification of documents over the web. Privacy is endangered if anyone can get a copy of a document merely by knowing a serial number. To protect against this risk, many websites use a well-known technique. For instance, the website of a stock exchange requires the user to enter a variety of details such as Expiry date, Strike price, Trade number, Order number and Quantity of trades involving futures and options. The system merely confirms or repudiates the details

Srinivasan, R.
Kuchibhotla, A.
Srinivasu, G.
Matad, B.

given. This ensures privacy, but imposes an effort penalty on the user. A digital signature in a barcode on any protected document achieves the same degree of privacy protection with less effort. It can be verified offline as well using a PC equipped with a scanner.

User feedback indicated other user problems which led to further research. Can we authenticate user documents in Indian scripts? Can we do this in a way that will stand the test of time - practices such as the use of ASCII or Unicode cannot stand the challenge of change over a few decades. Our team developed image signing techniques which took the text image from a paper document and authenticated it using the technology of digital signatures. This works in a script and font-independent manner. It also enables easy identification by the machine of what alterations have been made on the text image. Our technique prints out a color document derived by verifying a black and white document. Colored parts indicate altered text or components of diagrams (see Figure 3). One of the techniques we developed shows the unaltered text or diagram-element in place of the altered content.

It is difficult to authenticate images as compared to digital text. Digital text is error protected in storage and transmission. In contrast, images gather "noise" - that is, imperfections - as they are printed, stored, handled and scanned. Ensuring error tolerance of an adequate level without compromising the authentication technique posed a number of interesting technical challenges, which our team solved. An additional benefit we gained was that benign imperfections resulting from the document getting folded, the paper aging, and from folding, stretching or shrinking due to humidity changes could be distinguished from intentional alterations. The color indication of changes from the original document enables even an untrained user to recognize the significance of every alteration, and to guess the intent behind it. For instance, intentional change of a numeral or addition of text stands out as compared to imperfections due to fading or shrinking/stretching of text (Sankarasubramaniam et al, 2009).

All the previous techniques were related to authentication of a printed document. We also carried out some work on detection of originality of a paper document. One technique our team worked on involved detecting a photocopy as distinct from the original. With an adequately high density of data in the barcode (over 1000 bytes/square inch), it becomes difficult to make photocopies without introducing detectable levels of noise. Our copy detection technique used this information to determine if a document is the original or a photocopy.

We also developed a very strong technique for testifying the originality of a document, by associating a hardware token with it which is applicable

for high-value documents. The hardware token, say in the form of a smart card, demonstrates its link with the original document which carries the public key associated with that token. The token also stores and makes available a printable image of the original document. In conjunction with the above-mentioned techniques, this technique pairs the token unalterably with the document. The uniqueness is in the hardware token now, as against the paper document. The owner of the token can even present the token and a photocopy of the document as proof of his ownership of the original. An important feature of our solution is that the use of the token for verification does not use a network.



Figure 3: The Output of Verification Showing Possible Alterations in Red

Srinivasan, R.
Kuchibhotla, A.
Srinivasu, G.
Matad, B.

PROBLEMS IN DEVELOPING AND DEPLOYING SOLUTIONS

In this section we share some of our experiences in developing and deploying such solutions. We focus on experiences that may be relevant in the context of other technologies that are focused at emerging economies being developed by others.

The Quest for Low-Hanging Fruits

18

There is a tendency to focus on computerization and networking, ignoring the problems related to paper documents. Statements like "the important 75% of our revenue comes from urban branches; they will soon be covered by full computerization" indicate a problem. What about the other branches? What about third party entities that need to verify documents without having access to the issuer's database? Though they realize this is an important problem, the tendency is to address the seemingly more important and easier problem first.

The Inertia to Change

We did face the well-known problem of persuading customers to accept a change in their process - the problem of crossing the chasm (Moore, 1991). Customers have accepted computerization and use of computer networking as inevitable. But innovations on paper documents do not have the same visibility and hence the willingness to change processes is poor.

An inherent conservatism leading to resistance to innovation, perhaps more in the emerging economies, was another problem we faced. Potential users we dealt with were very well aware that paper signatures are often forged and go undetected. But there was hesitation to replace it with digital signatures even in valuable documents. One of the potential users made a remark which illustrates the point very well: "If there is a fraud involving a forgery in the traditional system, they will blame the culprit, but if there is a fraud exploiting some weakness of a new system I introduce, they will hang me!" In other words, the fear of possible failure of new solutions is not at all reduced by the known weaknesses of existing solutions and leads to a riskaverse behaviour.

Finally overly conservative purchase practices involve asking such questions as "Who are using this technology?" This rules out innovation, by definition!

Who benefits from innovation and who pays for it?

When an innovation is introduced to make paper documents more secure, the cost is borne by the issuer of the documents, who adopts the innovation.

The benefits, on the other hand, are to the third party users of the document. It is difficult to pass on the cost to the person who is issued the document, particularly in government applications and this reduces the incentive to introduce new technology.

Selling Solutions vs Selling Products

Development of solutions and related technology is a relatively long term effort. Development of an IP portfolio in the area takes years and field testing requires sustained cooperation of the R & D group and the business group involved.

19

Solutions also typically constitute a small fraction of the total business of computer companies (though that is changing now). They require greater employee time to sell, commission and service than readymade products. They play an important role in solving the problems of a customer by offering a complete ready-to-use solution. Very often, this need is met by involving a third party vendor who has a readymade solution. However, this makes it very difficult to satisfactorily utilize one's own understanding of user problems and innovative capability. So the company would rather continue to sell in established areas rather than venture into new ones that involve more selling time.

Costing and pricing solutions also poses its own problems. The cost of product development in the computer field is usually amortized over millions of customers. In contrast, solutions are sold in smaller numbers and it is difficult to set the development cost that the first few first customers should pay. A reliable estimate of how many customers will eventually benefit from a technology/solution is difficult to arrive at.

On the other hand customers also face problems in buying solutions, particularly in the public sector. Fairly rigid purchase practices require release of tenders that can be responded to by a number of vendors. This often rules out innovative solutions, which only an innovative company would be able to offer.

CONCLUSIONS

We are proud to say that the technology we created has had an impact. We now have a major insurance company in India adopting this technology for use on their policies and several others interested. The experiences we mentioned have greatly helped us understand the challenges in taking technology to the field in emerging economies and should greatly help us in the future too.

Srinivasan, R.
Kuchibhotla, A.
Srinivasu, G.
Matad, B.

We have also got substantial peer recognition. The technology was one of the finalists for the Asian Wall Street Journal Award in 2007 and the NASSCOM Innovation Award in 2007.

We will finally end this paper by summarizing some of the key insights we have derived from these experiences to benefit the readers of this paper. As we mentioned in the introduction, technology innovation plays a small part in the final success of any customer driven innovation, and factors other than technology should not be ignored. Some points worth remembering are:

20

For the innovator:

- Stay in close touch with customer-facing employees from the beginning to ensure that customer issues are given adequate importance.
- Fine tune the innovations to meet customer needs. Welcome the customer pointing out a weakness in your invention or innovation. This provides a challenge to take to the next step of invention or innovation.
- Customers do not adopt innovative ideas unless they see a significant need or cost/benefit ratio. Perseverance is necessary. In our case, we developed a variety of techniques to address customer needs. A single technique would not have been adequate.
- Continuous efforts are necessary to share information and ideas with everyone involved in the team. We have spent many hours brainstorming and sharing multiple perspectives and that has been extremely stimulating for us during this work.
- Look for opportunities to create intellectual property (IP); an innovation without IP creation denies the company the opportunity to amortize the cost of R & D by selling related products and services to a large number of customers without cut-throat competition.
- Do not fail to publish related to the innovation - it earns peer respect.

For the customer-facing staff in the business:

- Innovation to meet a big customer's need is not selling; it is a part of marketing.
- However, innovative fine tuning a product or service to meet a customer's need becomes often necessary to make a sale. It differentiates your offering from the others.
- Innovation on a big customer's demand earns valuable mindshare, usually at the highest levels of management.
- Innovation earns visibility in the media, which is also important.

For the customer

- Identify innovation required for your needs; this often requires brainstorming with your vendors and their R & D staff.
- A company would usually hold back its best suggestions in a multivendor brainstorming session. Nonetheless everyone gains from a certain amount of multi-vendor brainstorming. This would have to be followed soon by the development of a working relationship with a selected vendor's team for one-on-one discussions.
- Have reasonable expectations of the charge-free effort a company would put into a particular customer's need for innovation. Beyond this level it is profitable to pay, for instance through a small contract for a proof-of-concept.
- Media attention to an innovation benefits the customer as much as it benefits the company.

21

END NOTES

1. Work done while the author was with HP Labs India.
2. One crore is ten million. The currency referred to here is the Indian Rupee.

REFERENCES

- Hindu (2006) 'Officials face action for transferring land records', *The Hindu*, 23 March (online). Available from <URL: <http://www.hinduonnet.com/2006/03/23/stories/2006032312870400.htm>>.
- International Organization for Standardization (2006) *ISO/IEC 15438:2006: Information technology - Automatic identification and data capture techniques — PDF417 bar code symbology specification*. Geneva, International Organization for Standardization (online) Available from <URL:http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43816>
- Moore, G.A. (1991) *Crossing the Chasm*, New York, Harper Perennial.
- Reuters (2007) 'MIT dean resigns after falsifying degrees', Reuters, 26 April. (online) Available from <URL:<http://www.reuters.com/article/domesticNews/idUSN2631144020070426>>
- Sankarasubramaniam, Y., Narayanan, B., Viswanathan, K., and Kuchibhotla, A. (2009) 'Detecting Modifications in Paper Documents: A Coding Approach', *paper presented at Document Recognition and Retrieval XVII Conference (SPIE), 20-21 January 2010*, San Jose, CA, USA.
- Schneier, B. (2007) *Applied Cryptography: Protocols, Algorithms, and Source Code in C*, New Delhi, Wiley-India.
- Times of India (2002) 'Man obtains bank loans with forged documents', *The Times of India*, 29 March (online). Available from <URL:<http://timesofindia.indiatimes.com/news/city/hyderabad/Man-obtainsbank-loans-with-forged-documents/articleshow/5211893.cms>>
- US Patent and Trademark Office (2009) USPA 20090180698 Document Verification Method and System; 20090103803 Machine Readable Documents and Reading Methods, 20090059309 Document and Method of Producing a Document, 20080292136 Data Processing System and Method, 20080101601 Method and Apparatus for Preparing and Verifying Documents, New York, US Patent and Trademark Office.

Srinivasan, R.
Kuchibhotla, A.
Srinivasu, G.
Matad, B.

Dr. Ramani Srinivasan is Consultant in HP Labs India, Adugodi, Bangalore.

Dr. Anjaneyulu Kuchibhotla is Department Director in HP Labs India, Adugodi, Bangalore.

Godavari Srinivasu is Architect in Hewlett Packard-Bestshore, Hosur Road, Bangalore.

Bhushan Matad is Consultant in HP Labs India, Adugodi, Bangalore.

22

ACKNOWLEDGEMENTS

We thank our colleagues who have worked on this project and several of the technologies that have come out of it: Darpan Goel, Badri Narayanan, Anil Pai and Yogesh Sankarasubramaniam. We thank Chintakuntla Srinivas who carried out the necessary engineering for some of the prototypes. Sustained support of Sunil Narang and Shefali Nagpal of HP's Imaging and Printing Group was invaluable. Special thanks are due to Rajeev Chawla of the Indian Administrative Service. As a very experienced administrator of e-Governance projects, he gave us insight into user requirements and encouraged us to persevere in our efforts. Another senior government official who gave us encouragement is M N Vidyashankar. His keen interest and support to innovation have been very helpful. We thank Prof. Sadagopan for agreeing to pilot the technology with IITB.

Chitkara University

Saraswati Kendra, Plot 11-12, Dainik Bhaskar Building

Sector 25-D, Chandigarh-160014, India

Email: journal@chitkarauniversity.edu.in

Website: www.chitkara.edu.in/journal/index.php