

Criação de um Banco de Dados de Trajetórias de Aeronaves

Ronald Annoni Junior e Carlos Henrique Quartucci Forster

Instituto Tecnológico de Aeronáutica

Praça Marechal Eduardo Gomes 50, Vila das Acácias - 12228-900 - São José dos Campos - SP

Resumo—As trajetórias das aeronaves que trafegam em uma área de controle terminal, normalmente realizando procedimentos de saída ou de chegada aos aeroportos localizados em seu interior, tem grande interesse como objeto de pesquisa e para o desenvolvimento de aplicações voltadas para a melhoria da eficiência e da segurança do tráfego aéreo. No entanto, até a realização do trabalho relatado neste artigo, não havia um banco de dados de trajetórias de aeronaves em uma área de controle terminal brasileira que pudesse ser utilizada como massa de dados de treinamento e testes nas pesquisas. Este artigo descreve o processo de criação de um banco de dados de trajetórias de aeronaves na Área de Controle Terminal de São Paulo (TMA-SP) e apresenta alguns resultados de visualização desses dados.

Palavras-chaves—sistemas de informação

I. INTRODUÇÃO

Os serviços de tráfego aéreo possuem, entre outras finalidades, a de fornecer orientações apropriadas para que as aeronaves sejam conduzidas de forma segura e eficiente, respeitando os procedimentos e mantendo as separações mínimas previstas. Para que esses serviços sejam adequadamente prestados, o espaço aéreo é estruturado de forma que cada fase de voo, com suas características próprias, seja atendida por um órgão operacional de controle específico. A fase de voo em que as aeronaves realizam os procedimentos de saída ou de chegada aos aeroportos ocorre na chamada Área de Controle Terminal (TMA), que pode conter um ou mais aeródromos e está sob a jurisdição do órgão de controle operacional denominado Controle de Aproximação (APP).

Em Áreas de Controle Terminal que contenham aeroportos de grande movimento pode se exigir que as aeronaves, sob regras de voo por instrumento (IFR), sigam procedimentos padronizados de saída e de chegada que são conhecidos, respectivamente, por SID (*Standard Instrument Departure*) e STAR (*Standard Terminal Arrival Route*). Estes procedimentos contêm uma série de pontos (*waypoints*) que as aeronaves devem sobrevoar, sequencialmente, a determinadas altitude e velocidade, e tem por finalidade facilitar o controle, evitar o conflito entre tráfegos diferentes, evitar obstáculos naturais e, em alguns casos, reduzir o impacto ambiental (ruído) evitando-se o sobrevoos de determinadas áreas.

Uma das consequências previsíveis dessa padronização é a concentração bem destacada do tráfego em algumas

trajetórias coincidentes com os procedimentos padronizados, sujeitas obviamente a uma dispersão natural.

O tráfego que não segue os procedimentos de uma SID ou de uma STAR acaba também, ao longo do tempo, estabelecendo algum tipo de padrão observável por uma soma de razões, entre elas estão o fato de que as trajetórias normalmente começam ou terminam em aeroportos ou heliportos/helipontos, a observância de regras de voo visual (VFR) e de outras regras gerais de tráfego aéreo e ainda por razões de economia e de segurança. Nestes casos, no entanto, a variabilidade das trajetórias é muito maior.

Outras razões para se observar variabilidade nas trajetórias ou discrepâncias em relação aos padrões incluem o fato de os controladores de tráfego aéreo poderem autorizar rotas diretas para pouso, ou adiantar os procedimentos eliminando um ou mais *waypoints*. É possível, também, o controlador atrasar os procedimentos fazendo as aeronaves circularem por trajetórias mais longas devido a um espaço aéreo congestionado ou a alguma condição meteorológica desfavorável. Finalmente, há a possibilidade indesejável de que aeronaves, por algum erro, sigam trajetórias em desacordo com as regras, procedimentos ou instruções recebidas, elevando o risco para si e para as operações de forma geral.

O estudo das trajetórias das aeronaves em uma TMA possui diversas aplicações e pode gerar inúmeros resultados de interesse para o planejamento e desenvolvimento do tráfego aéreo. Como exemplo podemos citar estudos para otimização do fluxo do tráfego aéreo [1] com o objetivo de aumentar a eficiência das operações, estudos para redução de impactos ambientais, principalmente o ruído aeronáutico [2], e pesquisas na área de monitoramento de conformidade que tem como finalidade verificar a aderência dos voos aos planos preestabelecidos, às regras de tráfego e aos procedimentos padronizados para que se possa detectar desvios e evitar que aeronaves sigam trajetórias que venham a colocar em risco a segurança das operações aéreas [3].

Um dos grandes empecilhos para o aprofundamento das pesquisas nestas áreas reside justamente na dificuldade de obtenção de um banco de dados com um volume significativo de trajetórias reais de aeronaves em uma área terminal. Até onde nos foi possível levantar por meio da literatura técnica e científica, buscas na Internet, e consultas a pessoal técnico de empresas e de órgãos oficiais ligados ao sistema de controle do espaço aéreo brasileiro, não havia conhecimento de um tal banco com dados do Brasil, apesar do sistema de cobertura radar do nosso Controle de Tráfego Aéreo possuir, armazenados em fita

Ronald Annoni Junior, annoni@ita.br, Tel +55-12-39476890; Carlos Henrique Quartucci Forster, forster@ita.br, Tel +55-12-39475981. Este trabalho contou com o apoio do Controle de Aproximação de São Paulo (APP-SP), subordinado ao Serviço Regional de Proteção ao Voo de São Paulo (SRPV-SP).

magnética, registros das detecções realizadas.

Este artigo relata o nosso desafio de prover um banco de dados de trajetórias reais de tráfego aéreo na Área de Controle Terminal de São Paulo (TMA-SP), a partir dos registros de detecções armazenados em meio magnético no Controle de Aproximação de São Paulo (APP-SP).

II. ORIGEM DOS DADOS

Descrevendo de forma bastante simplificada, o sistema computadorizado de controle do tráfego aéreo utilizado no APP-SP recebe as informações dos radares, produz uma síntese das detecções e distribui as informações já tratadas para os consoles dos controladores e para os servidores de gravação dos dados em fita magnética. A partir deste cenário foram cogitadas algumas opções para extrairmos os dados das detecções dos radares.

A primeira opção era interceptar os dados chegando diretamente dos radares ligados ao APP-SP. Neste caso havia uma relativa complexidade e risco na implementação da conexão física e do protocolo de comunicação direta com os radares em produção que preferimos evitar. Além disso, não teríamos o resultado da síntese das detecções produzidas pelo sistema do APP-SP e nem dados históricos. Poderíamos, no máximo, com os dados correntes produzir a nossa própria síntese e gerar o nosso próprio histórico, o que não era adequado.

Diante dessas dificuldades e considerando o fato de que o sistema possui um módulo de gravação, resolvemos extrair os dados diretamente dos arquivos das fitas magnéticas. Esta seria uma boa opção pois as fitas contém a situação real completa e já tratada dos dados. Inclusive uma das funcionalidades do sistema é exatamente a de se reproduzir uma visualização de determinado período a partir dos arquivos gravados em fita. No entanto, esta opção também se mostrou infrutífera em razão do formato proprietário dos arquivos.

Como alternativa, optamos finalmente por utilizar uma funcionalidade do sistema que permite gerar, a partir dos arquivos gravados, um relatório em texto com as informações de radar. Um fragmento desse relatório pode ser visto na Fig. 1. Este relatório lista, em ordem cronológica, todas as sínteses produzidas pelo sistema e logo abaixo de cada linha de síntese, exibe um grupo de linhas com as informações provenientes de cada radar individual que gerou aquela síntese. Uma breve descrição do significado de cada campo desse relatório pode ser vista na Tabela I.

O processo de leitura das fitas e a geração dos relatórios efetuado pelo sistema em uso no APP-SP é bastante demorado. Cada fita, que contém aproximadamente 24 horas de tráfego, leva também aproximadamente 24 horas para ser lida e gerar o relatório. Foram lidas 82 fitas e gerados 82 relatórios cobrindo o período de novembro/2010, dezembro/2010 e janeiro/2011, com exceção de alguns dias cujas fitas haviam sido recolhidas pelo órgão de prevenção de acidentes aeronáuticos para análise. Cada relatório lista as detecções de um dia de tráfego aéreo sobre uma região quadrada de 256 milhas náuticas de lado com centro no radar de Congonhas, localizado nas coordenadas 23° 37' 31.4" S and 46° 39' 12.1" W, e possui, em média, 150.000 páginas / 7.500.000 linhas de texto (700MB de arquivo .txt) cada.

Tabela I
DESCRIÇÃO DOS DADOS DO RELATÓRIO MULTIRRADAR

Código	Descrição	Valores
ID	Número da Pista Síntese	
NP	Número da Pista Local	
QUAD	Quadrícula da Pista	
RAD	Nome do Radar	
HORA	Instante da Mensagem	
XPOS	Posição X do alvo (NM)	
YPOS	Posição Y do alvo (NM)	
TD	Origem da Detecção	NC=Não Comunic. EX=Extrapolada C=Secundária 2D=Primário 2D 2C=Associado 2D 3D=Primário 3D 3C=Associado 3D
QF	Fator de Qualidade da Pista	0=Pior 7=Melhor
HEA	Direção (GRAUS)	
VELC	Velocidade (NÓS)	
SSR3	Código SSR Modo 3A	
A	Validade do Código SSR	0=Válido 1=Duidoso 3=Inválido
1/2	Código SSR 1A ou 2A	
C	Tipo do Código SSR	1=Código SSR 1A 2=Código SSR 2A
MODEC	Altitude Modo C (x 100 PÉS)	
TNV	Tendência de Nível de Voo	= Não Aplicável ? = Não Comunic. - = Descendente + = Ascendente = = Estável
V	Validade de Tendencial de Nível	N=Idade (0-3) ?=Não Válido
3D	Altitude 3D (x 100 PÉS)	
PLN	Plano de Voo Correlacionado	
CFL	CFL	
INDICAT	Indicativo	
LABELOP	Quarta linha da etiqueta	
TP	Tipo da Pista	R=Real T=Teste
BGEC	IDs do Primeiro Radar de Detecção, Primeiro Centro de Detecção, Radar de Detecção Corrente e Centro de Detecção Corrente	
X	Número do exercício de treinamento	0=Real 1=Teste
P	Indicação de Emergência (7500, 7600, 7700)	
E	Indicação de Emergência Militar	
S	Identif. Especial (SPI)	
W	Identif. Colisão AR-SOLO	
C	Identif. Colisão AR-AR	
R	Identif. Conformidade com a Rota	
M	Indic. Múltipla Assoc.	P=Pista Principal S=Pista Suplente
H	Indicador de Plano em Espera (HOLDING)	
OBS	Observações	E=Msg. ERRO CRI=Criação SUP=Supressão DEF=Defusão

Relatorio Multi Radar Pagina 166055 Novembro 20 00:00:39 2010

ID	QUAD	HORA	XPOS	YPOS	TD	QF	HEA	VELC	SSR3	A	1/2	C	MODEC	TNV	V	3D	PLN	CFL	INDICAT	LABELOP	TP	BGEC	X	P	E	S	W	C	R	M	H	OBS		
NP	RAD	PRI	HORA	XPOS	YPOS	TD	QF	PRO	VELC	SSR3			MODEC			3D							E								OBS			
44	755		74.70	57.34	20.56	2C	7	90	381	3352	0	0000	2	171	?	0	----	73	0	ONE6258	RJ	R	I	A	F	A	0	0	0	-	-	-	-	
77	RAGR	13	71.86	57.03	20.59	2C	7	90	381	3352			171			----	0						0											
292	RAMO	11	72.10	57.09	20.53	2C	7	90	380	3352			171			----	0						0											
243	RASJ	15	71.46	56.94	20.53	2C	7	90	381	3352			171			----	0						0											
354	800		74.70	71.91	125.06	C	7	36	507	4265	0	0000	2	370	=	0	----	0	0	-----	GL01266	R	J	A	F	A	0	0	0	-	-	-	-	-
67	RAGR	11	71.38	71.41	125.06	C	7	35	508	4265			370			----	0						0											
937	RAKP	9	69.86	71.91	123.94	C	7	36	506	4265			370			----	0						0											
234	RASJ	15	74.36	71.72	125.28	C	7	35	509	4265			370			----	0						0											

Figura 1. Fragmento de um Relatório Multirradar

De posse dessa enorme fonte em texto, passamos então a desenvolver um processo para a extração dos dados e o seu armazenamento utilizando um sistema gerenciador de banco de dados (SGBD).

III. EXTRAÇÃO E ARMAZENAMENTO DOS DADOS

Primeiramente selecionamos o sistema gerenciador de banco de dados PostgreSQL [4] como ferramenta para administrarmos o nosso repositório de trajetórias. O PostgreSQL foi escolhido por ter, reconhecidamente, confiabilidade, estabilidade e desempenho comparáveis aos melhores SGBDs do mercado para lidar com grandes volumes de dados, sem custo de licença. Contribuiu também para essa escolha o fato de se poder utilizar o software PostGIS [5] para adicionar ao servidor PostgreSQL suporte a objetos geográficos, permitindo que ele seja empregado como um banco de dados espacial backend para um sistema de informações geográficas. Essa extensão possibilita o armazenamento, no banco, de objetos do tipo POINT, LINestring, POLYGON, MULTIPOINT, MULTILINESTRING, MULTIPOLYGON e GEOMETRYCOLLECTION com suporte à representação em coordenadas geográficas. Todos os cálculos (áreas, distâncias, comprimentos, intersecções, etc.) com os tipos GEOGRAPHY são efetuados sobre um elipsóide definido pelo padrão WGS84.

Após a instalação e a configuração do SGBD passamos para o desenvolvimento do modelo conceitual do banco de dados de trajetórias. A linha mestra para o desenvolvimento do modelo inicial para essa primeira extração dos dados dos relatórios era de que ele deveria ser o mais simples possível e, ao mesmo tempo, ser capaz de armazenar todos os tipos de dados contidos nos relatórios em texto. A idéia era não desprezar nenhum tipo de informação e também permitir que qualquer informação armazenada no banco pudesse ser facilmente localizada de volta nos relatórios, para conferência, mesmo que isso custasse alguma desnormalização do modelo. Seguindo essas diretrizes o modelo conceitual foi desenvolvido utilizando a ajuda da ferramenta DB Designer Fork [6] para a construção gráfica do diagrama entidade-relacionamento e geração automática dos scripts SQL correspondentes. A Fig. 2 ilustra o modelo para o armazenamento inicial dos dados.

Uma vez aplicados os scripts gerados pela ferramenta DB Designer Fork para a criação física do banco de dados, passamos então a desenvolver uma solução para efetuar a leitura dos arquivos de relatório, identificação dos dados e gravação nas tabelas apropriadas. Um exame

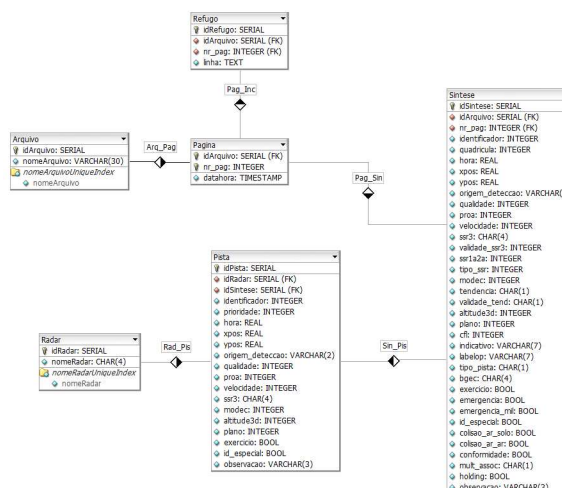


Figura 2. Modelo para o armazenamento inicial dos dados

dos relatórios identificou três tipos de linhas com dados úteis para serem extraídos:

- Linha de cabeçalho de página
- Linha de pista síntese
- Linha de pista radar

Para cada um desses tipos de linha foi desenvolvida uma expressão regular que permitisse identificá-la e facilitasse a separação de cada dado individualmente. Além desses três tipos de linha foram identificados outros tipos que não possuíam dados para serem retirados:

- Linhas em branco
- Linhas de cabeçalho de dados
- Linhas de mensagens do sistema

Para estes casos foi desenvolvida uma expressão regular para identificar se a linha pertencia a esse grupo. A idéia era de que se uma linha não fosse identificada por nenhuma das expressões regulares ela seria um caso não previsto e então ela deveria ser armazenada em uma tabela de refugio (ver Fig. 2) para ser analisada manualmente. Esta foi uma precaução tomada pois não era viável inspecionar visualmente todas as linhas de todos os relatórios antes do processamento.

Um aplicativo em Java foi implementado para realizar a leitura dos arquivos de relatório, linha a linha, identificando cada uma por meio das expressões regulares desenvolvidas. Caso a linha fosse identificada como uma das linhas de dados, ela era passada para uma classe de *parsing* cuja função era a de retornar os dados individuais para a inserção no banco. Já se a linha fosse identificada como uma das linhas sem dados, ela era simplesmente

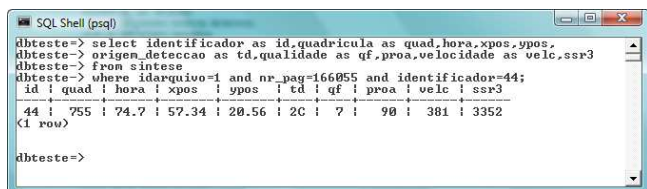


Figura 3. Consulta ao banco de dados para verificação de coerência com os relatórios

ignorada. Por fim, se a linha não fosse classificada em nenhuma das hipóteses anteriores, ou se houvesse alguma discrepância durante o *parsing* (tipos incongruentes, por exemplo), a linha seria inserida na tabela de refugo para ser posteriormente analisada de forma manual.

Até a data de redação deste artigo, 32 dos 82 arquivos de relatório haviam sido processados dessa forma, levando em média 2 horas de processamento para cada arquivo. Nenhuma linha dos relatórios foi refugada o que indica que os mesmos têm uma estrutura “bem comportada”, regular. Algumas consultas feitas no banco de dados para verificar se os dados estavam sendo armazenados corretamente, de forma coerente com os relatórios, foram positivas como se pode ver na Fig. 3 confrontada com a primeira linha do relatório da Fig. 1. Faltava ainda verificar se os dados tinham coerência espaço-temporal, ou seja, era preciso verificar se o conjunto de pontos sequenciais de uma determinada pista (de radar) representava uma trajetória e se essas trajetórias estavam corretamente posicionadas sobre o terreno. Para esse fim foi necessário desenvolver aplicações de visualização como será descrito a seguir.

IV. VISUALIZAÇÃO DAS TRAJETÓRIAS

A primeira aplicação de visualização desenvolvida fez uso da API Java 2D para simplesmente plotar os pontos de detecção retornados pelas consultas ao banco de dados. Na Fig. 4 estão plotadas 24 horas de detecções sobre toda a área de cobertura (256NM x 256NM) dos dados, resultado da seguinte consulta SQL:

```

SELECT xpos, ypos
FROM sintese
WHERE idarquivo=1;
    
```

Devido à escala e à quantidade de informação a imagem da Fig. 4 está bem poluída, mas já é possível perceber a presença de trajetórias.

Aproveitando-se da facilidade de consulta que um sistema gerenciador de banco de dados proporciona, reduzimos o escopo de nossa consulta para retornar apenas as detecções das aeronaves utilizando código transponder entre 3000 e 4000, e distando até 40 milhas náuticas do radar de Congonhas. O código SQL abaixo mostra como esta consulta mais complexa pode ser feita ainda de forma simples. O seu resultado pode ser visto na Fig. 5:

```

SELECT xpos,ypos
FROM sintese
WHERE idarquivo=1 AND
      ssr3>'3000' AND ssr3<'4000' AND
      sqrt(xpos^2+ypos^2)<=40;
    
```

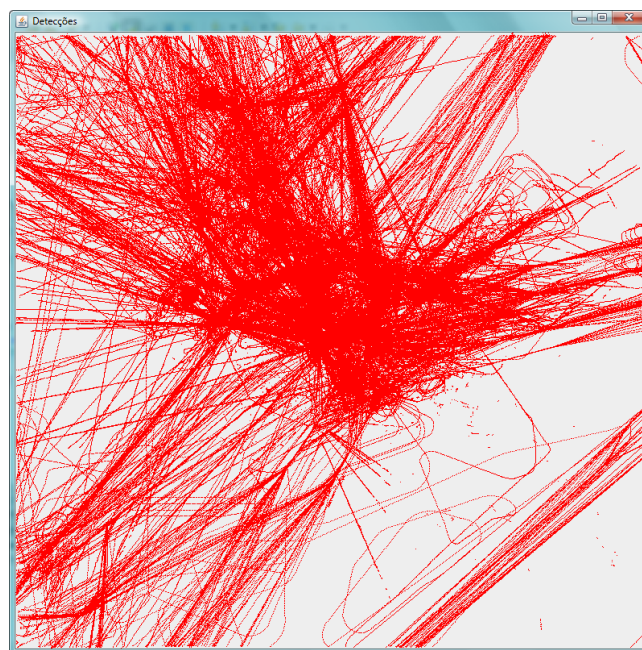


Figura 4. 24 horas de detecções de radar na TMA-SP (256NM x 256NM)

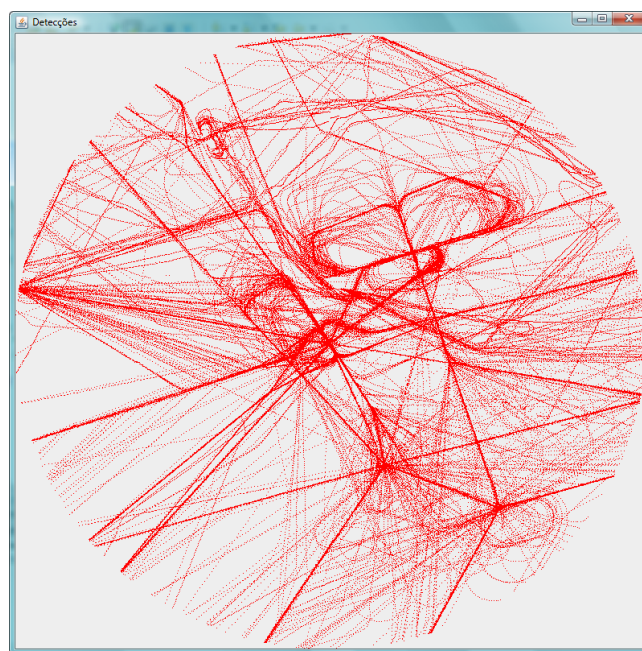


Figura 5. 24 horas de detecções com códigos transponder na faixa 3000-4000 distando até 40NM do radar de Congonhas

Como se pôde notar, na Fig. 5 as trajetórias já ficaram bem mais nítidas e já foi possível até observar alguns padrões com a formação de linhas onde o tráfego é mais frequente. Esta e outras consultas similares foram também executadas com as detecções sendo plotadas em sequência cronológica e ficou demonstrada a relação espaço-temporal das detecções.

Como último estudo de caso, para demonstrarmos a coerência geográfica das trajetórias armazenadas em nosso banco de dados e o potencial de utilização dessas informações, desenvolvemos um aplicativo em Java que gera, a partir de uma consulta ao banco de dados, um arquivo no formato KML [7] para ser visualizado no “geobrowser”

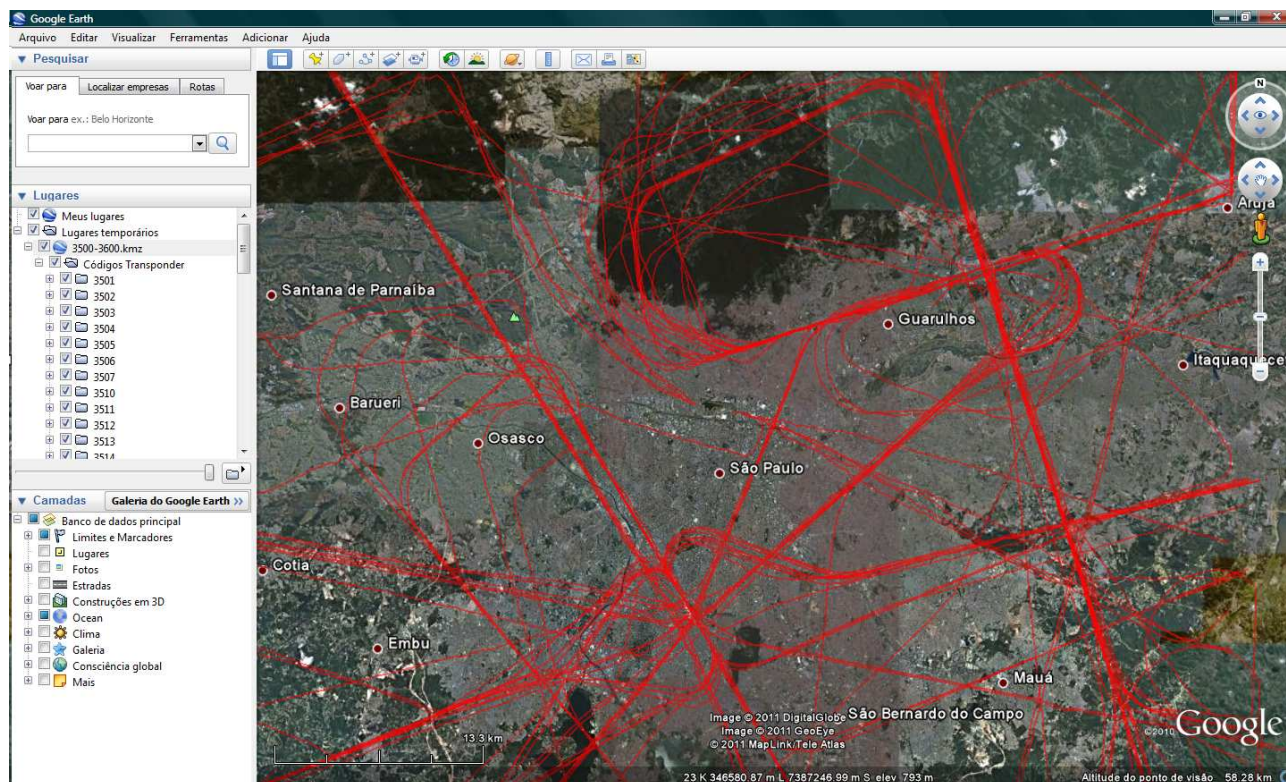


Figura 6. 24 horas de detecções com códigos transponder na faixa 3500-3600

Google Earth [8]. O Google Earth é utilizado para visualização de dados geográficos e espaciais de forma interativa e muito simples. Os dados são visualizados sobre imagens de alta definição do globo terrestre e o usuário pode navegar livremente (*pan/tilt/zoom*) sobre eles, filtrá-los e acrescentar ou remover camadas de informações. Dados de usuários são fornecidos ao Google Earth no formato KML que é baseado no padrão XML e possui o mesmo tipo de estrutura hierárquica, aninhada e *tag-based*.

A primeira sequência de três imagens mostra o resultado de uma consulta que buscou em um dia as aeronaves que utilizaram código transponder na faixa de 3500 a 3600. A Fig. 6 mostra a visão mais aberta da área com o tráfego coincidindo perfeitamente com os aeroportos de Congonhas, Guarulhos e com o Campo de Marte. Na barra à esquerda há a opção de se habilitar ou desabilitar a visualização das trajetórias.

A Fig. 7 mostra um zoom desta mesma imagem sobre o Aeroporto de Guarulhos enquanto a Fig. 8 mostra uma visão em perspectiva 3D desse mesmo local. Esta última mostra claramente o tráfego cruzando transversalmente a pista de Guarulhos e fazendo uma curva à esquerda para o pouso. Todos esses pontos de vista foram possíveis de atingir apenas navegando sobre os mesmos dados formatados no padrão KML.

Como último resultado, a Fig. 9 mostra uma imagem interessante obtida pela seleção apenas dos códigos transponder utilizados pelos helicópteros sobre a cidade de São Paulo. É nítida a concentração sobre as Rotas Especiais de Helicópteros (REH), principalmente sobre o rio Pinheiros, e também na região da Avenida Engenheiro Luis Carlos Berrini, onde há um grande número de helipontos.

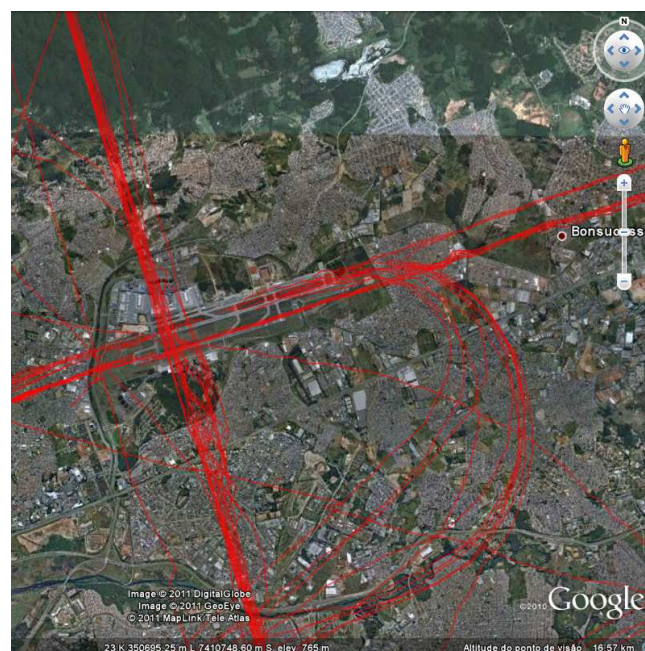


Figura 7. Zoom para o Airport de Guarulhos

V. CONCLUSÕES

Neste artigo foram relatadas as atividades realizadas para a criação de um banco de dados de trajetórias de aeronaves na Área de Controle Terminal de São Paulo (TMA-SP).

Demonstramos a existência de interesse técnico e científico no estudo desses dados como, por exemplo, na otimização do fluxo do tráfego aéreo, no estudo de impactos ambientais e também para monitoramento de conformi-

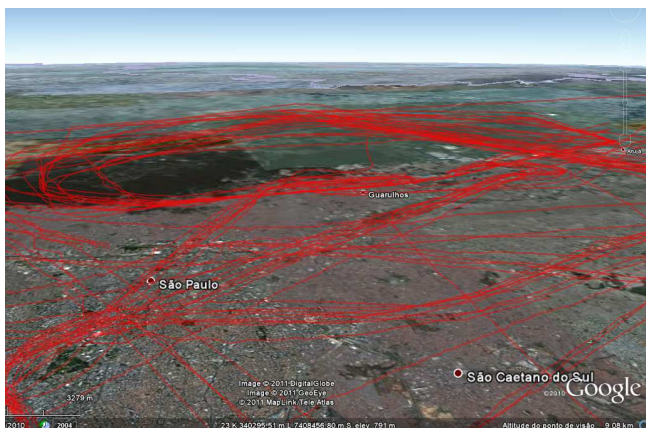


Figura 8. Visão em perspectiva 3D do tráfego

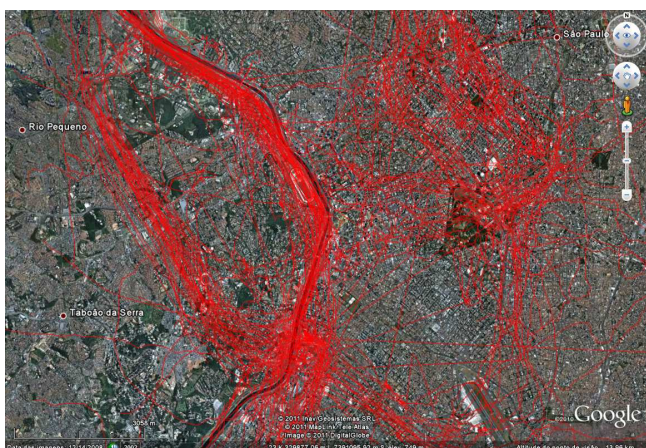


Figura 9. Tráfego de helicópteros sobre a cidade de São Paulo

dade das operações aéreas.

Apesar de haver esta demanda, estes dados até hoje não estavam disponíveis numa forma adequada para o seu uso. Empregando métodos, técnicas e ferramentas de modelagem de dados e de programação, extraímos os dados brutos de sua fonte primária e os colocamos dispostos em um ambiente de fácil consulta e acesso.

Para demonstrarmos essa facilidade de se obter resultados rápidos de questões complexas e o potencial de utilização deste novo repositório, apresentamos estudos de caso de visualização do tráfego aéreo.

Com isso espera-se agora imprimir um maior ritmo para as pesquisas nesta área, abrindo novas perspectivas e fomentando novas idéias.

AGRADECIMENTOS

Ao Controle de Aproximação de São Paulo (APP-SP), subordinado ao Serviço Regional de Proteção ao Voo de São Paulo (SRPV-SP), pela produção e pelo fornecimento dos relatórios multirradar utilizados neste trabalho.

REFERÊNCIAS

[1] D. P. Dias, C. Müller, and P. P. Porto, “Procedimentos de melhoria do fluxo de tráfego aéreo para operações de pouso no aeroporto internacional de São Paulo,” in *Anais do IX Simpósio de Transporte Aéreo (IX SITRAER)*, Manaus, Brasil, Outubro 2010, pp. 92–103.

[2] R. A. A. Wijnen and H. G. Visser, “Optimal departure trajectories with respect to sleep disturbance,” *Aerospace Science and Technology*, vol. 7, no. 1, pp. 81–91, 2003.

[3] T. G. Reynolds and R. J. Hansman, “Conformance monitoring approaches in current and future air traffic control environments,” in *21st Digital Avionics Systems Conference*, vol. 2. Piscataway, NJ: IEEE Publications, October 2002, pp. 1–12.

[4] PostgreSQL Database Management System. PostgreSQL Global Development Group. <http://www.postgresql.org>

[5] PostGIS. Refrations Research. <http://postgis.refrations.net>

[6] DB Designer Fork. <http://sourceforge.net/projects/dbdesigner-fork>

[7] KML 2.2 Reference. Google Inc. <http://code.google.com/apis/kml/>

[8] Google Earth. Google Inc. <http://earth.google.com>