Temporal GIS and Statistical Modelling of Personal Lifelines

Marius Thériault[1], Christophe Claramunt[2], Anne-Marie Séguin[3] and Paul Villeneuve[1]

Centre de recherche en aménagement et développement, Université Laval, Québec, CANADA G1K 7P4, tel. (418) 656-7558, fax (418) 656-2018, {Marius.Theriault, Paul.Villeneuve}@crad.ulaval.ca[1]
Institut de recherche de l'École navale, BP 600, 29240 Brest Naval, FRANCE, Claramunt@ecole-navale.fr[2]
INRS Urbanisation, Culture et Société, Université du Québec, 3465 rue Durocher, Montréal, CANADA H2X 2C6 ; tel. (514) 499-4047, Anne-Marie_Seguin@inrs-urb.uquebec.ca[3]

Abstract

Spatio-temporal data models for describing complex lifelines and trajectories of persons, as well as events that affect their evolution, are prerequisites for the statistical analysis of their relationships. Such analyses are useful to develop a better understanding of urban dynamics and social transformations. This paper develops a spatio-temporal database model for handling personal trajectories along a time line (many complementary lifelines) allowing for the statistical analysis of any pre-defined event. It combines survival analysis, Cox regression and temporal GIS. The combination of these aspects support an assessment of the likelihood of any event to occur in the life of persons at risk, after a given time delay and under some specific conditions. Our model was implemented and tested using a geo-relational approach that also supports spatial and temporal reasoning at complementary levels of abstraction. It allows the cross-analysis of several multi-dimensional lifelines to form individual trajectories. The application example is based on an historical survey of personal biographies (spatially located) of 418 professional workers living in the Quebec Metropolitan Area in 1995-96.

Keywords: temporal GIS, multi-dimensional lifelines, event history analysis, individual trajectories

1 Urban Dynamics, Temporal GIS and Event History Analysis

Modelling the evolution of urban systems implies the analysis of human activities at different levels of abstraction along their temporal and spatial dimensions. Understanding persons' behaviour in a dynamic perspective and the constraints of their environment can provide clues about processes responsible for the evolution of cities. Such analysis, however, clearly implies integration of several multi-dimensional episodes describing the behaviour (biography) of individuals. Within urban studies, space-time representations of individual trajectories have been widely studied using time-geography modelling and analysis concepts following Hägerstrand's seminal work (1967). Other investigations have explored time-geography principles to analyse the multi-dimensional structure of spatial behaviour of people involved in specific activities (Taylor and Parkes, 1975; Goodchild and Janelle, 1984; Miller, 1991; Odland and Shumway, 1993; Whiters, 1997; Janelle *et al*., 1998; Johnson, 2001).

During the last decade, important progress was made in integrating the time dimension within GIS (Langran, 1989; Peuquet, 1994; Frank, 1994; Peuquet and Quian, 1996; Böehlen *et al*., 1999; Hornsby and Egenhofer, 2000). The statistical analysis of spatio-temporal processes, however, is still lagging far behind the development of GIS technology. Some of the temporal GIS models identified so far are based on an explicit description of spatial evolution with respect to change in status for identifiable entities (Claramunt and Theriault, 1996; Hornsby and Egenhofer, 1998 and 2000). These approaches aim at the modelling of transition patterns among events in an effort to identify recurring spatio-temporal processes. They provide an explicit description of changes in the geographic phenomena modelled. Analysis of basic transition primitives and their combination to record geographical change can reveal patterns that could further our understanding of evolution processes. Such an approach, requires exploratory analysis of successive events. To the best of our knowledge, database models are not currently linked to existing statistical analysis of successive events, survival analysis, and modelling of choice processes and therefore not ideally suited to theses types of assessments. In fact, because of the complexity of their data structures, they only allow computation of some basic descriptive statistics on transitions among successive states (E.g. mean values and frequencies).

On the other hand, the literature is replete with existing statistical methods for modelling discrete choices and for performing temporal analysis of state transitions (logistic regression, log-linear models, survival analysis using Kaplan-Meier techniques, Cox regression). These statistical approaches are linked to probability theory. Making assumptions on the error distribution of observations, they provide hypothesis tests based on the comparison of actual events to their theoretical distribution, in space and time. Thus, among a set of events, they could identify those which are unlikely to appear at random in the actual spatio-temporal configuration. Adding the power of inferential statistics to GIS applications would clearly enhance their usefulness for urban studies. It will allow scientists to assess the effect of data sampling procedures on the accuracy of simulation results. As well, it will provide strong guidelines to distinguish patterns of events which are significant from those which are not (Pötter and Blossfeld, 2001).

Event history modelling is a specific type of longitudinal statistical analysis that focuses on the survival rate of a given status and considers various attributes of the observed individuals. It is used to estimate the probability of an event to occur, considering the time elapsed after some condition is met (Blossfeld and Rohwer, 1995; Blossfeld, 1996; Le Bourdais and Marcil-Gratton, 1998). For example, event history analysis (Cox, 1972) can compute the likelihood of a woman to give birth to a child, after she gets married (or she starts a cohabiting union), taking into consideration her age and her income. This likelihood is the complement of her survival function for not having a child under the same conditions.

Survival analysis uses Kaplan-Meier (product-limit) technique to estimate the length of time elapsed before occurrence of an event after a change in personal status enables it to occur. Using observed duration of elapsed time, event, survival analysis computes the base line hazard function of an event to occur. However, at survey time, the event would not have occurred for some enabled individuals. These are then called censored. Because censored cases were at risk, they must be considered when estimating the proportion of individuals experiencing the event. Therefore, one cannot estimate hazard using only those objects experiencing a given event. The stochastic part of the phenomenon must also be considered explicitly to estimate the odds ratio (number of individuals experiencing the event divided by those who do not). It must be accumulated over time to model hazard evolution. When the event occurs for a person, he is generally no longer at risk. Thus, the number of persons at risk changes over time while hazard decreases or increases, depending on the nature of the event (E.g., death and marriage). Moreover, the chance that an event occurs depends partly on the attributes of each individual and the context in which he/she lives. For example, risk of death increases with age of the person, someone living in dangerous conditions is also more at risk. The marginal effect of the local context on hazard assessment (decrease in survival) is modelled using time regression based on exponential mathematical models: $h(t) = e^{(\beta_0 + \alpha \ln t)} e^{(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n)}$. In this equation, hazard is modelled using two terms. The first part estimates the base hazard using time elapsed after enabling occurred (natural logarithm of t). The second one is specified using any number of independent variables (X) providing values specific to each observation. Coefficients are computed for observed data using maximum likelihood techniques. This defines the basis of Cox regression techniques (Cox, 1972) available in most statistical software, like SPSS.

Unfortunately, SPSS, like all other statistical software, can manipulate only flat files. It allocates columns to variables and rows to observed cases. For most applications in time regression (and survival analysis), the task of restructuring information to build this flat file is time-consuming and prone to errors. Moreover, computation of many attributes depends on the choice of reference events (enabling conditions and modelled event). While it is certainly feasible to integrate spatial context in survival analysis (even time-varying context), to the best of our knowledge, there is no existing example of such an application. This requires integration of data models that defines events, their ordering and semantics within a GIS environment. In addition, a data manipulation language that allows for retrieval of transition sequences among events is required. We believe that such an improvement would open new frontiers for the application of GIS in urban studies.

The objective of this paper is to present a GIS modelling strategy to efficiently generate event history tables compatible with most statistical software. The procedure is illustrated and tested using an historical survey of 418 professionals living in the Quebec Metropolitan Area during 1995-96. The purpose of this survey was to collect detailed information about all significant events that occurred during their adult life (residential, familial and professional trajectories), with specific references to space and time. The following sections will discuss some lifeline modelling principles needed for this project (Section 2) and the linkage to temporal GIS (Section 3). Section 4 presents a geo-relational extension of our model. Section 5 discusses the application of the survey and discusses some preliminary results. Finally, the conclusion assesses the overall efficiency of the procedure and outlines further work.

2 Lifeline Modelling Principles

Modelling individual trajectories within GIS might offer a better data support for the development of urban land-use and transportation models. An important achievement of recent urban modelling is the integration of the decision-making behaviours of urban actors. These include activity location and travel decisions which are intricately linked with household structures and professional profiles of persons (Hunt and Simmonds, 1992). Information about these contextual attributes is needed for temporal regression analysis (Ben-Akiva and Lerman, 1985) or computer-intensive micro-simulation. Temporal regression is also required for modelling complex systems, which integrate decision rules of many interacting individuals.

During the last decades, individual trajectories generated patterns of events of increasing complexity. This is linked to many factors such as an escalation of divorces, extension of contractual short-term employment and increasing geographical mobility. This trend is highly related to the economic restructuring occurring in most countries since the mid seventies (Rose and Villeneuve, 1993; Séguin, 1994). Within cities, individual trajectories aggregate to yield demographic, professional and residential patterns that can be observed using census data. However, the processes by which personal biographies aggregate to form macro patterns cannot be derived from censuses. The latter give only the barest snapshot reports on complex situations (Thériault *et al*., 1999).

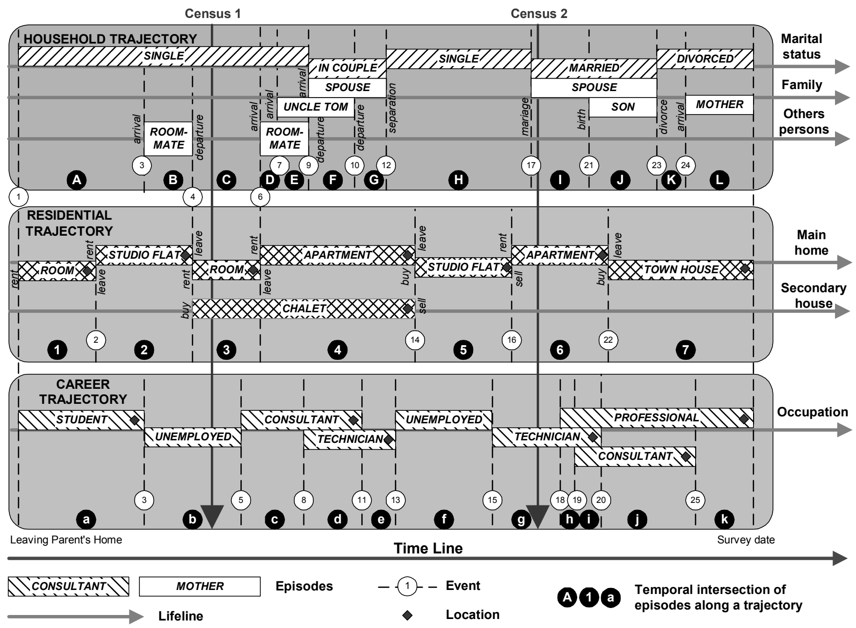

Fig. 1. Schematic example of an individual history

In our database schema, an individual history (Fig. 1) is formed by a set of complementary lifelines defining three trajectories (household, residential and professional career). An individual's history is altered when an event occurs modifying one of his lifelines, or in other words, the episodes of one of his lifelines. Let us remark that such an event may alter - simultaneously or afterwards - several episodes of many lifelines, leading to interdependent episodes. For example, finding employment could be a preliminary condition for renting an apartment. Other events can be anticipated and lead to prior adjustment: a household's move may anticipate a new child birth. Therefore, trajectories show interlocked evolution based on personal behavioural strategies. They are formed of partially or totally ordered episodes. Moreover, each of these trajectories is somewhat different in structure. A household composition changes with arrival or departure of any member. The residential trajectory depicts the succession of living places. The career pathway may describe successive mixes of potentially simultaneous occupational activities (Thériault *et al.*, 1999).

An individual's lifeline is defined as a set of time-stamped episodes along one of his life dimensions (E.g. marital status, having children). These lifelines are described using episodes and events. The global status of an individual for an instant of time can be derived from an aggregation of his different lifelines. A lifeline has a form that is either linear or multi-linear. For example, some individuals may hold several jobs or residential locations simultaneously. Lifeline episodes constitute logical sequences related to life cycles. Studying these evolution (or decision) patterns is certainly more relevant for urban planning than knowing the exact timing of events for each individual.

Fig. 1 illustrates a complex mix of real world phenomena (persons, trajectories, lifelines, episodes and events). All these notions, including the individuals and trajectories themselves, are integrated in our model. The first basic modelling concept is the episode, which is used to describe individuals and their historical properties. An episode denotes an individual's homogeneous value along one modelling dimension, that is, a lifeline. An individual's history is composed of many lifelines that include partially ordered time-stamped (or not) episodes, using temporal periods. An episode shows some status that endures over a given period of time and corresponds to the individuals lifeline. Following Snodgrass (1999), we define a period as an anchored interval of time delimited by two time-stamps (begins and ends). A time interval also indicates the length of the period (its duration). In the sense of Allen (1984), an episode - a property in Allen's terminology - is divisible. Therefore, if an episode holds over a period of time *i*, then all its properties hold during any subinterval of *i*.

A second important modelling concept relates to events. An event models one-to-many changes of episodes from one-to-many individuals' lifelines. Each episode is bounded by two events (begins and ends), potentially building multidimensional networks of ordered successive episodes or events. An event is time-stamped by either an instant or a period. An event using a period of time is indivisible. For example, if a marriage occurred in June 1999 (time granularity of a month), we cannot derive that this marriage happened on the 15th of June 1999.

For prototyping purposes, our modelling concepts are specified using a relational database approach. However, these modelling concepts are general enough to be applied to other database models such as object-oriented models. Two relations constitute the main components of this database: events and episodes (Fig. 2). They are described as follows:

An **episode** is defined as a tuple Episode (<u>EpisodeId</u>, EpisodeName, LifeLineDimension, SetOfIndividualId, SetOfRelatedIndividualId, EpisodeTime, QuestTime, EpisodeSpatial). EpisodeId uniquely identifies each episode. EpisodeName represents and names the class of episodes represented along a lifeline (E.g. single or married). LifeLineDimension indicates the dimension modelled by this lifeline (E.g. individual’s life, marital status, and occupation). SetOfIndividualId keeps the list of individuals to whom this episode belongs (E.g. the person answering the questionnaire, and his/her relatives actively involved in this episode). SetOfRelatedIndividualId lists the other persons linked to this episode (E.g. members of an individual’s household playing only a passive role). EpisodeTime models the temporal period over which this episode is valid. EpisodeSpatial indicates the spatial location of this episode. QuestTime keeps the date when the questionnaire was completed (important for surveys including persons answering at different times).

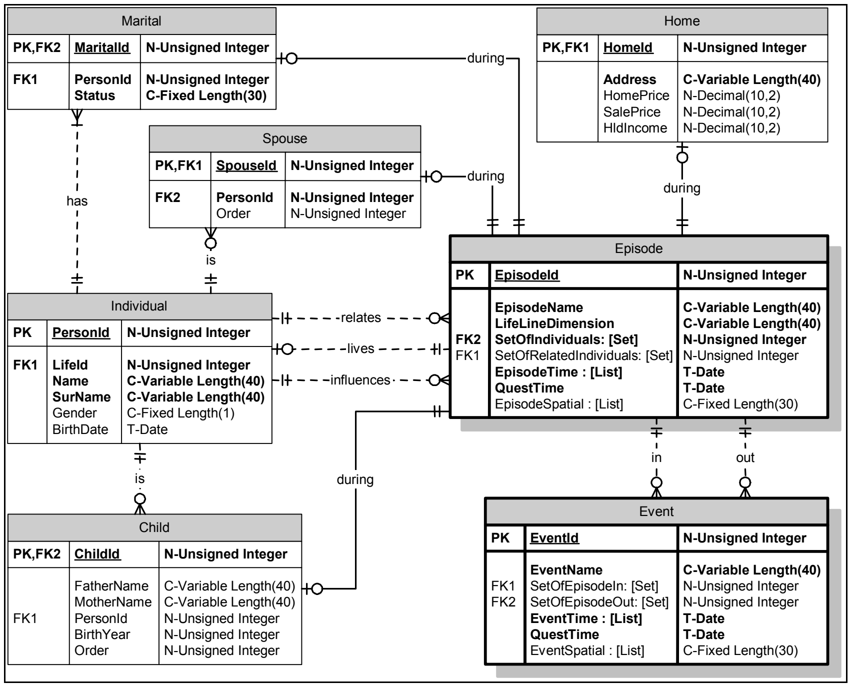

Fig. 2. Relational schema for describing household and residential lifelines

An **event** is defined as a tuple *Event* (*EventId*, *EventName*, *SetOfEpisodeInId, SetOfEpisodeOutId, EventTime, QuestTime, EventSpatial*). *EventId* uniquely identifies each event. *EventName* represents the class of events represented (E.g. marriage, birth). *SetOfEpisodeInId* gives the set of episodes which are terminated by this event. *SetOfEpisodeOutId* gives the set of episodes which are initiated by this event. As previously mentioned, an event can be linked to several episodes. Conversely an episode might be either terminated or generated by the conjunction of several events. These links provide means of selecting the previous and the following episodes, using their explicit sequence ordering. This could work even when time stamps are missing or when these episodes are non-immediate successors or predecessors. *EventTime* models the time when this event happens (using either temporal instants or periods). *QuestTime* indicates the date when the information was collected. *EventSpatial* keeps the spatial location of this event.

Temporal and spatial attributes use multi-level data representations: a calendar of type (*Year: Month*) for temporal attributes *EventTime, EpisodeTime* and *QuestTime*; and, a spatial hierarchy of type (*MunCode: Municipality: NeighbourhoodType: PostalCode: Longitude: Latitude*) for spatial attributes *EventSpatial* and *EpisodeSpatial*. Such representations support reasoning at various levels of abstraction as introduced in Claramunt and Jiang (2000). As necessary, this primary structure can be extended using additional relations to handle specific attributes (thematic, temporal and spatial) that complement the semantics of the application (E.g. Child, Spouse, Home as illustrated in Fig. 2).

3 Temporal GIS Modelling

Lifeline histories are modelled at the basic individual level using events and episodes. Thanks to a graph structure linking successive episodes and events, this model also provides means to handle chronological time and historical sequences (ordered events) in the same structure. For example, *SetofEpisodeIn* points to every episode ended by the current event, while *SetofEpisodeOut* lists all those following it (Fig. 2). Additionally, because individuals move within the city, the same database schema can be used to analyse spatio-temporal patterns.

In Fig. 2, the relational schema uses sets to model groups of individuals related through an episode (E.g. a newborn child and his parents). Lists are used to bound time periods (begin- and end-date) and to record multi-level spatial location. In an effort to map this relational schema to the software environment used for prototyping, (combination of MapInfo and MS Access) some minor adaptations were required (Fig. 3). Sets were replaced by intermediate tables (*EpisodeIn* and *EpisodeOut*), thus-allow for many-to-many relationships among episodes and events. Lists of individuals were mapped to tables (*ActingIndividuals* and *RelatedIndividuals*) and linked to episodes. *EpisodeTime* and *EventTime* were disaggregated in two fields giving their beginning and ending dates. Spatial lists (holding spatial locations) are defined using a *Spatial* table that maintains relationships with both episodes and events. Finally, the resulting mapping is monitored by an administration table (*MapInfo_MapCatalog*) that provides instructions for generating map symbols when specified tables and views are opened within MapInfo, using ODBC services.

Events and episodes are stored into two tables using the fields *EpisodeName*, *LifeLineDimension* and *EventName* to distinguish their type, and to identify the lifeline to which the episode belongs. All geographical locations are recorded in a unique table, named *Spatial*. Therefore, several events and episodes can be related to a single location tuple as required. Time management is operated in two ways by four interrelated tables: *Episode*, *Event*, *EpisodeIn*, *EpisodeOut*. Date fields, in both *Event* and *Episode* tables, keep track of chronological time, using T-Date types. Links to *EpisodeIn* and *EpisodeOut* define multi-dimensional ordering of related episodes and events. All spatio-temporal features managed by these five general-purpose tables, greatly eases the implementation of spatio-temporal views and the formulation of spatio-temporal queries.

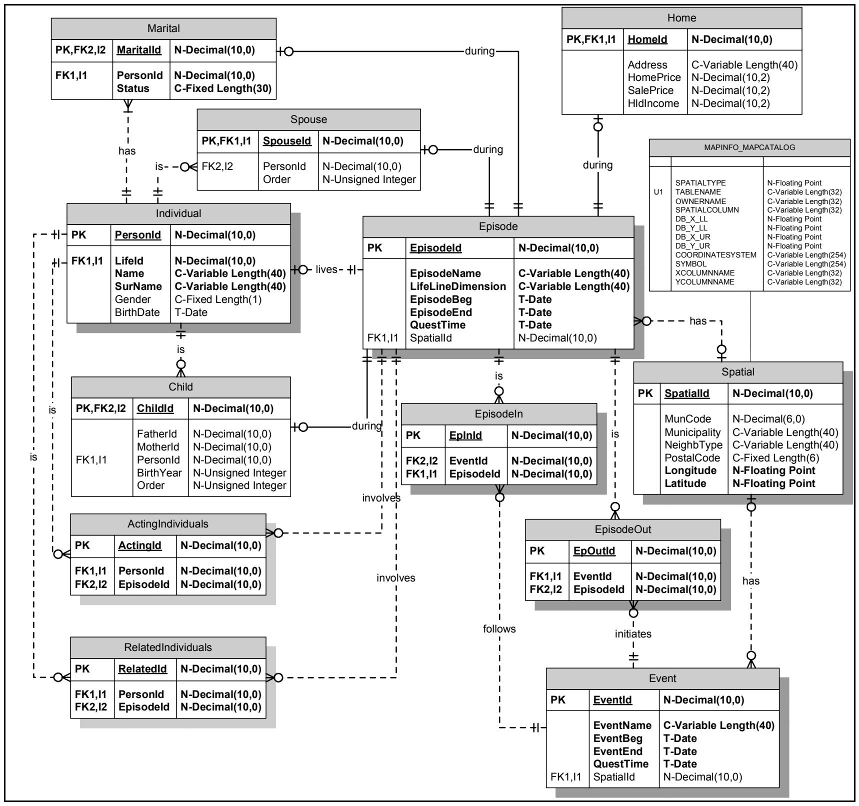

Fig. 3. Access-MapInfo expanded schema for household and residential lifelines

4 Spatio-Temporal Views

Manipulating several tables using a given query is not a straightforward and efficient task. Complex queries can be predefined using conventional relational views embedded in the database schema. Fig. 4 presents some examples of views integrating various components of respondents' events and episodes with data about the respondent itself. Using views, one can model some specific schemas showing any type of episodes or events, providing an efficient link with the secondary tables of Fig. 3 (E.g. Home, Marital). Moreover, these views can integrate any spatial and temporal information in simple structures. Depending on the DBMS capabilities, these views can later be joined to build more complex views, thereby defining a comprehensive system to solve queries.

Table 1 details a list of tables and views involved in defining views for our application, indicating join relationships used. *Respondent_Episodes* and *Respondent_Events* views (Fig. 4) give, for each respondent of our survey, the list of his/her reported episodes and events. Using database triggers, these views automatically add the respondent's *Age* (*EpisodeBeg*-*BirthDate*) and the *Duration* of the episode (in months) to each tuple. They also indicate if, at survey time, the

Table 1. Relational links defining spatio-temporal views of individual lifelines

Resulting View	From Tables / Views	Relational / Sorting Conditions
RespEpis	[ActingIndividuals] As [c1] [Episode] As [c2] [Individual] As [c3] [Spatial] As [c4]	[c1].[EpisodeId] = [c2].[EpisodeId] ^ [c2].[SpatialId] = [c4].[SpatialId] ^ [c1].[PersonId] = [c3].[PersonId]
IndivEvent	[RespEpis] As [c1] [Event] As [c2] [EpisodeOut] AS [c3] [Spatial] As [c4]	[c1].[EpisodeId] = [c3].[EpisodeId] ^ [c2].[EventId] = [c3].[EventId] ^ [c2].[SpatialId] = [c4].[SpatialId]
Couples	[Spouse] As [c1] [Episode] As [c2] [RespEpis] As [c3]	[c1].[SpouseId] = [c2].[EpisodeId] ^ [c2].[EpisodeId] = [c3].[EpisodeId] ^ [c3].[Respondent]
Respondent_Episodes	[RespEpis] As [c1] [Episode] As [c2] [Individual] As [c3]	[c1].[EpisodeId] = [c2].[EpisodeId] ^ [c1].[Respondent] ^ [c1].[PersonId] = [c3].[PersonId] Order By [c1].[PersonId], [c1].[EpisodeBeg]
Respondent_Events	[IndivEvent] As [c1] [Event] As [c2] [Individual] As [c3]	[c1].[EventId] = [c2].[EventId] ^ [c1].[PersonId] = [c3].[PersonId] ^ [c1].[Respondent] Order By [c1].[PersonId], [c1].[EventBeg]
Episodes_Before_Event	[RespEpis] As [c1] [Respondent_Events] As [c2]	[c1].[PersonId] = [c2].[PersonId] ^ [c1].[EpisodeEnd] <= [c2].[EventBeg] Order By [c1].[PersonId], [c2].[EventBeg], [c1].[EpisodeBeg]
Episodes_After_Event	[RespEpis] As [c1] [Respondent_Events] As [c2]	[c1].[PersonId] = [c2].[PersonId] ^ [c1].[EpisodeBeg] >= [c2].[EventEnd] Order By [c1].[PersonId], [c2].[EventEnd], [c1].[EpisodeBeg]

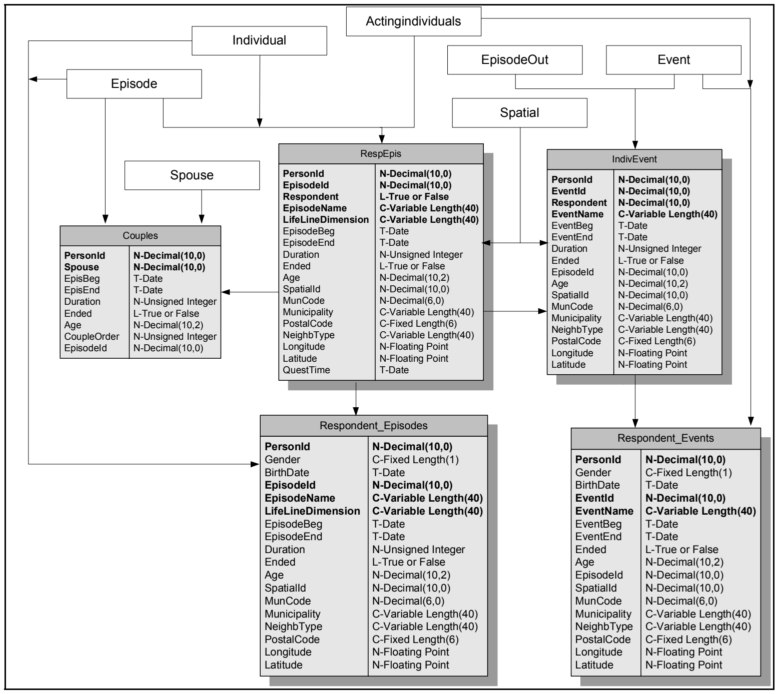

Fig. 4. Spatio-temporal integrated views of respondents' episodes and events

episode was *Ended*, or still happening. Using time-stamps to order each respondent's episodes and events, the next step defines multi-dimensional links among lifelines (Fig. 5; Table 1). Two multi-dimensional views (*Episodes_Before_Event* and *Episodes_After_Event*) show respectively, for each respondent, the list of all episodes that started before any event occurred during his entire life and all those that started after the event. In the view *Episodes_After_Event*, the field *TransTime* indicates elapsed time (in years) between the event and the episode (*[c1].[EpisodeBeg]-[c2].[EventEnd]*). If an episode was not ended at survey time (*[c1].[EpisodeBeg]>=[c1].[QuestTime]*), it is called *Censored*.

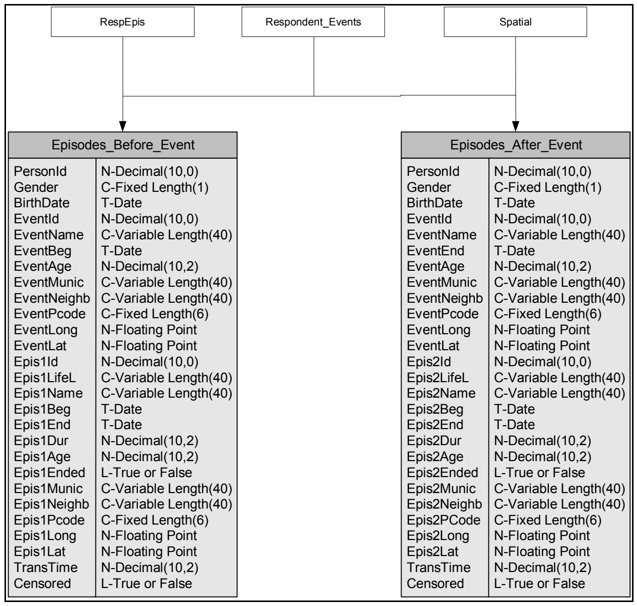

Fig. 5. Multi-dimensional lifeline transitions among episodes through each event

5 Application of the Quebec Metropolitan Area

In our application example, 418 respondents reported 5,541 events and episodes. Episodes are distributed into four lifelines (Family, Individual, Marital and Residential), 10 episode types and 12 event types (Table 2). For each respondent, an episode and an event were assigned to the interview because its timing is needed to distinguish censored events and ongoing periods. Building multi-dimensional sequences of periods preceding and following each event during the life of each individual, generated 22,676 tuples in view *Episodes_Before_Event*, and 44,839 tuples in view *Episodes_After_Event*. This last view provides relevant data and logical structure to generate the flat files needed for time regression in SPSS.

Table 3 presents a relational-like query selecting sequences of episodes and events according to the following criteria: Among people who were tenants when their first child was born, distinguish those who decided to buy a house afterwards from those who did not; then compute how long they postponed their decision and to which distance they moved.

Table 2. Typology of episodes and events

	Episodes		Events	
Lifeline	EpisodeName	Count	EventName	Count
Family	Child	750	Alone	234
Family	Child Departed	102	Birth	418

Individual	Interview	418	Buy	646
Individual	Life	418	Child Adoption	3
Individual	Self	418	Child Arrival	64
Marital	Couple	533	Child Birth	683
Marital	Separated	174	Child Departure	102
Marital	Single	234	Inhabit	169
Residential	Owner	646	Interview	418
Residential	Roomate	169	Leaving Parent	418
Residential	Tenant	1679	Rent	1679
			Separation	174
			Union	533

Table 3. Relational-like query to generate the event history flat file for SPSS

Query command	Notes
Select *, Distance([c1].[Longitude],[c1].[Latitude],[c2].[Epis2Long],[c2].[Epis2Lat],"Km") "DIST" From [Respondent_Episodes] As [c1], [Episodes_After_Event] As [c2] Where [c1].[PersonId]=[c2].[PersonId] And (([c2].[EventName] = "Child Birth" and [c2].[Epis2Name] = "Owner") or ([c2].EventName = "Child Birth" and [c2].[Epis2Name] = "Interview")) and ([c1].[EpisodeName]="Tenant") and ([c1].[NeighbType] in ("City Core","Old Suburbs","New Suburbs","Urban Fringe")) and {[c2].[EventEnd] .TDuring. [c1].[EpisodeBeg;EpisodeEnd]} Group By [c1].[PersonId], [c1].[EpisodeId] Order By [c1].[PersonId], [c1].[EpisodeId], [c2].[Epis2Beg] Object [Line, [c1].[Longitude]:[c1].[Latitude];[c2].[Epis2Long]:[c2].[Epis2Lat]] Into Event_History_Table	Computing moving distance Using Spatio-temporal views Restricting to each person's life Buying home after child's birth Censored: still tenants at survey time Respondent was tenant at that time Respondent was living in Quebec CMA Birth date is during residential episode Retaining only one child birth by person Ordering to retain the first set (child-buy) Generating a line between locations Generating the event history table

In Table 3, the relational-like query uses a join between the view *Respondent_Episodes* and the table *Episodes*. The view *Respondent_Episodes* gives the home location of some tenants when the first child was born. This describes the neighbourhood where the 151 respondents who were tenants when their first child was born were living at that time (City core, Old Suburbs, New suburbs, Urban fringe). This gives an indication of their relative location within the Quebec Metropolitan Area. Knowing which ones did not change their housing tenure afterwards (*Censored*), and which ones did, we can test the hypothesis of a relationship linking neighbourhood types where young parents live and their decision to buy a home less than five years later (Table 4). The chi-square test indicates that no significant relationship was found between these two facts in our sample of tenants (probability = 0.566). Therefore, willingness to purchase a house after the first child's birth seems invariant across the city.

Table 4. Cross-tabulation of home location of tenants and their decision to buy a home

Decision to buy a home within 5 years after the first child is born	Type of living neighbourhood when the first child was born								
	City Core		Old suburbs		New suburbs		Urban fringe		
	Count	Expected	Count	Expected	Count	Expected	Count	Expected	Sum
Did buy a home	33	34.8	64	60.5	7	8.3	1	1.4	105
Did not buy a home	17	15.2	23	26.5	5	3.7	1	0.6	46
Sum	50	50	87	87	12	12	2	2	151

$\chi^2 = 2.032$; $df = 3$; $probability = 0.566$

Table 5 shows the result of a Cox regression estimating attributes influencing the decision of tenants to buy a home after the first child is born. According to the overall chi-square test, the relationship is highly significant. Three factors were found to have a significant influence (Wald's statistics) on the willingness to buy a home: decade of birth of the first child (*PeriodB*: odds ratio increasing with time), distance at which people are ready to move in order to access home ownership (*Dist*: odds ratio increases with distance), and duration of their stay in the new home (*Epis2Dur*: odds ratio increases with expected duration of stay).

Table 5. Attributes influencing tenants for buying a new home after the first child is born

	B	SE	Wald	df	Sig.	Odds e^B
EPIS2DUR : Duration of stay at destination home (Years)	0.137	0.019	50.650	1	0.000	1.147
DIST : Distance from child birthplace to new home location (Km)	0.007	0.002	12.545	1	0.000	1.007
PERIODB : Decade of birth of the first child			26.687	3	0.000	
1960-69	-2.115	0.734	8.311	1	0.004	0.121
1970-79	-1.574	0.485	10.530	1	0.001	0.207
1980-89	-0.267	0.426	0.393	1	0.530	0.765
Reference 1990-95	0.000					1.000

$\chi^2 = 61.314$; $df = 5$; $probability < 0.0001$

Table 5 can be used to build a mathematical expression for the likelihood of tenants to buy a home (*Censored=0*) at time *TransTime* after their first child is born (*EVENTNAME="Child Birth"*):

$$h(t) = e^{(\beta_0 + \alpha \ln TRANSTIME)} e^{(0.137 EPIS2DUR + 0.007 DIST - 2.115 Period60 - 1.574 Period70 - 0.267 Period80)}$$

, using respondent's spatio-temporal attributes. This function has two parts:

$e^{(\beta_0 + \alpha \ln TRANSTIME)}$, the base line hazard function related to elapsed time after the first child is born (see Fig. 6),

and $e^{(0.137 EPIS2DUR + 0.007 DIST - 2.115 Period60 - 1.574 Period70 - 0.267 Period80)}$, the marginal effect of independent variables on respondent's decision to buy a home. Moving distances come from measurement of moving paths from the location of previous home to that of the newer (Fig. 6).

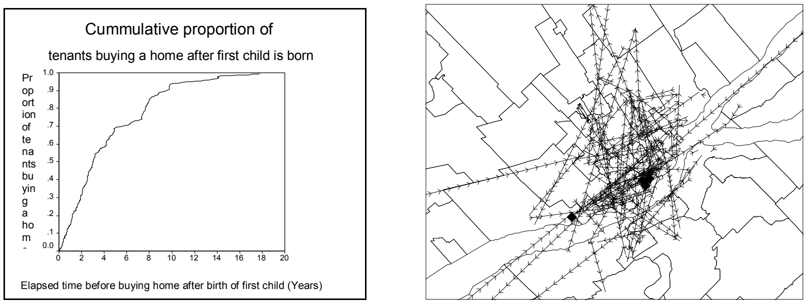

Fig. 6. Likelihood function of buying a home and moving paths of house places

6 Conclusion

This paper introduces a GIS database approach illustrated by an application example that clearly demonstrates the interest of analysing event histories at the individual level using temporal statistical methods. Moreover, integrating events and episodes management within GIS provides efficient means for integrating geographical criteria into the modelling of people's decision. This goes far beyond the mere mapping of individual evolution path over space. Because, they isolate space-time attributes in specialised tables, the geo-relational schemas of this application can be readily used in many other temporal GIS applications needing to relate successive events occurring along lifelines.

Further work is needed to extend the concept of view towards spatial and temporal operators. Integrity constraints have to be enforced to check the convergence of chronological and ordered temporal information handled in the database. A user-friendly query interface has to be developed for better manipulation of data. Finally, adding dynamics would improve mapping of results.

Acknowledgements

This project was funded by the Quebec Province's FCAR, the Canadian SSHRC, the Canadian NSERC, the Québec Province's Ministère de l'Éducation and the Canadian Network of Centre of Excellence in Geomatics (GEOIDE).

References

Allen JF (1984) Toward a general theory of actions and time. Artificial Intelligence 23:123-154

Ben-Akiva M, Lerman SR (1985) Discrete choice analysis: Theory and application to travel demand. MIT Press, Cambridge, Mass

Blossfeld HP (1996) Macro-sociology, rational choice theory, and time: A theoretical perspective on the empirical analysis of social processes. European Sociological Review 12(2):181-206
Blossfeld HP, Rohwer G (1995) Techniques of event history modelling. In: New approaches to causal analysis. Lawrence Erlbaum Associates, Mahwah, New Jersey
Böhlen M, Jensen C, Scholl M (1999) Spatio-temporal database management. In: Lecture Notes in Computer Sciences #1678, Springer-Verlag, Berlin
Claramunt C, Jiang B (2000) Hierarchical reasoning in space and time. In: Forer P, Yeh AGO and He J (eds) Proceedings of the 9th International Symposium on Spatial Data Handling. Beijing, China, 3a, pp 41-51
Claramunt C, Thériault M (1996) Toward semantics for modelling spatio-temporal processes within GIS. Proceedings of the 7th International Symposium on Spatial Data Handling. Delft, Holland, 47-63
Cox DR (1972) Regression models and life tables. Journal of the Royal Statistical Society B-34:187-220
Egenhofer M (1991) Reasoning about binary topological relations. In: O. Günther and H.-J. Schek (eds) Advances in Spatial Databases. Springer-Verlag, Berlin, pp 143-160
Frank AU (1994) Qualitative temporal reasoning in GIS – Ordered time scales. Proceedings of the 6th International Symposium on Spatial Data Handling. Zurich, Switzerland, pp 410-430
Goodchild MF, Janelle DG (1984) The city around the clock: Space-time patterns of urban ecological structure. Environment Planning A 16:807-820
Hägerstrand T (1967) Innovation Diffusion as a Spatial Process. University of Chicago Press, Chicago, Illinois
Hornsby K, Egenhofer M (1998) Identity-based change operations for composite objects. Proceedings of the 8th International Symposium on Spatial Data Handling. Vancouver, Canada, pp 202-213
Hornsby K, Egenhofer M (2000) Identity-based change: A foundation for spatio-temporal knowledge representation. International Journal of Geographical Information Systems 14(3):207-224
Hunt JD, Simmonds DC (1992) Theory and application of an integrated land-use and transport modelling framework. Environment and Planning B: Planning and Design 20:221-244
Janelle DG, Klinkenberg B, Goodchild MF (1998) The temporal ordering of urban space and daily activity patterns for population role groups. Geographical Systems 5(1-2):117-137
Johnson MP (2001) A spatial decision support system prototype for housing mobility program planning. Geographical Systems 8(1):49-67
Langran, G. (1989) A review of temporal database research and its use in GIS applications. International Journal of Geographical Information Systems 3(3):215-232
Le Bourdais C, Marcil-Gratton N (1998) The impact of family disruption in childhood on demographic outcomes in young adulthood. In M. Corak (ed) Labour Markets, Social Institutions and the Future of Canada's Children. Statistics Canada, Ottawa, pp 91-105
Miller HJ (1991) Modelling accessibility using space-time prism concepts within geographical information systems. International Journal of Geographical Information Systems 5(3):287-301
Odland J, Shumway JM (1993) Interdependencies in the timing of migration and mobility events. Papers in Regional Science 72(3):221-237

Peuquet DJ (1994) It's about time: A conceptual framework for the representation of temporal dynamics in geographical information systems. Annals of the Association of American Geographers 84(3):441-461

Peuquet DJ, Quian L (1996) An integrated database design for temporal GIS. Proceedings of the 7th International Symposium on Spatial Data Handling. Delft, Holland, pp 1-11

Pötter U, Blossfeld HP (2001) Causal inference from series of events. European Sociological Review 17(1):21-32

Rose D, Villeneuve P (1993) Work, labour markets and households in transition. In Bourne L and Ley D (eds) The Social Geography of Canadian Cities. McGill-Queens Press, Montreal, pp 153-174

Séguin AM (1994) Transformations sociales et carences de l'habitation. In: Dumont F, Langlois S, Martin Y (eds) Traité des problèmes sociaux. Institut québécois de recherche sur la culture, Québec, pp 65-79

Snodgrass RT (1999) Developing Time-Oriented Database Applications in SQL. Morgan Kaufmann, New York

Taylor PJ, Parkes DN (1975) A Kantian view of the city: A factorial-ecology experiment in space and time. Environment and Planning A 7:671-688

Thériault M, Séguin AM, Aubé Y, Villeneuve P (1999) A spatio-temporal data model for analysing personal biographies. Proceedings of the 10th International Workshop on Database and Expert Systems Applications. IEEE Computer Society, Florence, Italy, pp 410-418

Whiters SD (1997) Methodological considerations in the analysis of residential mobility: A test of duration, state dependence, and associated events. Geographical Analysis 29(4):354-374