

Vizualizácia časovo závislých dát



Martin Škorupa

2009

Vizualizácia časovo závislých dát

DIPLOMOVÁ PRÁCA

Martin Škorupa

**UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
KATEDRA ALGEBRY, GEOMETRIE A DIDAKTIKY MATEMATIKY**

Študijný odbor 9.1.1 Matematika

Vedúci diplomovej práce
Mgr. Matej Novotný

Bratislava 2009

Čestne prehlasujem, že som túto prácu vypracoval samostatne s použitím uvedenej literatúry a zdrojov pod odborným vedením Mgr. Mateja Novotného

V Bratislave 27. apríla 2009

.....
Martin Škorupa

Predovšetkým by som sa chcel poďakovať môjmu školiťovi Mgr. Matejovi Novotnému za jeho cenné rady a nápady pri tomto projekte. Ďalej Martinovi Cifovi Štefčekovi za pomoc pri implementácii. A v neposlednom rade svojim rodičom a starým rodičom, ktorí mi umožnili študovať a počas štúdia ma podporovali.

Martin Škorupa

Abstrakt

ŠKORUPA, Martin: Vizualizácia časovo závislých dát. [Diplomová práca] / Martin Škorupa. Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky; Katedra algebry, geometrie a didaktiky matematiky. Školiteľ: Mgr. Matej Novotný. Bratislava: FMFI UK, 2009.

Vizualizácia sa úspešne využíva pri analýze časovo závislých dát. Nie všetky výskumné odvetvia však majú skúsenosti a predstavu čo všetko môžu pomocou vizualizácie dosiahnuť. Predkladaná práca podáva riešenie v podobe nástroja, ktorým môžeme intuitívne riešiť analytické otázky a súčasne poskytuje jednoduché riešenie na prezentovanie poznatkov získaných z tejto analýzy. Bola motivovaná potrebami zoológov o jednoduchú analýzu ich nameraných dát a snahu prezentovať ich v čo najpriateľnejšej podobe.

Práca rozoberá vizualizačné metódy časovo závislých dát, ktoré sú dostatočne intuitívne pre bežného človeka a súčasne kategorizuje časovo závislé dáta.

Kľúčové slová: Vizualizácia, Časové dáta, Pracovná plocha

Obsah

Abstrakt.....	1
Obsah.....	2
Zoznam ilustrácií.....	4
Úvod.....	5
Motivácia.....	6
Cieľ.....	7
1 Základné pojmy.....	8
2 Časovo závislé dáta.....	10
2.1 Vlastnosti.....	10
2.2 Taxonómia.....	10
2.2.1 Diskrétné časové body - Časový interval.....	11
2.2.2 Lineárny čas - Cyklický čas.....	11
2.2.3 Sériový čas - Súvislý čas.....	12
2.2.4 Postupný čas – Stromový čas.....	12
3 Vizualizácia.....	13
3.1 Vlastnosti.....	13
3.2 Taxonómia podľa vizualizačných vlastností.....	15
3.2.1 2D – 3D.....	15
3.2.2 Statická – Dynamická.....	15
3.2.3 Jednorozmerná – Viacrozmerná.....	16
3.2.4 Riadená dátami – Riadená udalosťami.....	16
3.3 Prehľad niektorých vizualizačných metód.....	17
3.3.1 Konvenčné vizualizačné metódy.....	17
3.3.2 Vizualizačné metódy pre viacrozmerné a početné dáta.....	19
4 Návrh aplikácie.....	21
4.1 Dáta.....	21
4.2 Prostredie.....	21
4.2.1 Správa dát.....	22
4.2.2 Pracovná plocha.....	22
4.3 Ďalšia funkcionality.....	23
5 Realizácia.....	24

5.1 Prostredie.....	24
5.1.1 File.....	24
5.1.2 Data (F3).....	24
5.1.3 Visualization (F4).....	25
5.1.4 Animation (F5).....	26
5.1.5 Help (F1).....	26
5.2 Pohyb po pracovnej ploche.....	26
6 Prehľad niekoľkých vizualizácií.....	28
6.1 Prezentácia.....	28
6.2 Analýza.....	29
6.3 Prehľad.....	30
7 Záver.....	31
Zoznam použitej literatúry.....	32
Obsah CD.....	34

Zoznam ilustrácií

1. Obr.: Diskrétné časové body.....	11
2. Obr.: Časový interval.....	11
3. Obr.: Lineárny čas.....	11
4. Obr.: Cyklický čas.....	11
5. Obr.: Sériový čas.....	12
6. Obr.: Súvislý čas.....	12
7. Obr.: Postupný čas.....	12
8. Obr.: Stromový čas.....	12
9. Obr.: 2D viz.....	15
10. Obr.: 3D viz.....	15
11. Obr.: Statická viz.....	16
12. Obr.: Dynamická viz.....	16
13. Obr.: Jednorozmerná viz.....	16
14. Obr.: Viacrozmerná viz.....	16
15. Obr.: Viz. riadená dátami.....	17
16. Obr.: Viz. riadená udalosťami.....	17
17. Obr.: Postup Napoleonských vojsk.....	17
18. Obr.: Bodový, Čiarový, Stĺpcový a Cyklický graf.....	18
19. Obr.: Zložený a Kumulovaný graf.....	18
20. Obr.: ThemeRiver.....	18
21. Obr.: Jednofarebné ofarbenie.....	19
22. Obr.: Dvojfarebné ofarbenie.....	19
23. Obr.: New Data.....	24
24. Obr.: Data.....	25
25. Obr.: Editovací panel.....	25
26. Obr.: Animation.....	26
27. Obr.: Popup.....	26
28. Obr.: Prezentácia.....	28
29. Obr.: Analýza.....	29
30. Obr.: Prehľad.....	30

Úvod

V súčasnosti sú informácie jednou z najdrahších komodít na trhu. S rozvojom informačných technológií vznikajú každodenne rozsiahle databázy takmer v každom odvetví. Mnohé z nich majú obrovskú hodnotu napríklad pri analýze trhu. Dáta z CT, alebo MRI môžu zachraňovať ľudské životy. Analýza počasia a prírody môže predchádzať alebo varovať pred prírodnými katastrofami.

Dáta, zvyčajne v číselnej podobe, nemajú sami o sebe takú informačnú hodnotu. Je potrebné ich hlbšie analyzovať, dať do súvisu s inými dátami a správne interpretovať aby nám dali požadovanú informáciu.

Vizuálne vnímanie je najzákladnejším zmyslom na získavanie informácií. Človek pomocou neho získava najviac informácií o okolitom svete. Nie je preto prekvapujúce, že práve vizualizácia je popri matematickej štatistike najpoužívanější prostriedok na dolovanie informácií z dát. Človek je od malička cvičený v rozpoznávaní ikon, symbolov a vzorov, preto vizualizácia umožňuje dostatočnú abstrakciu na zobrazovanie aj obrovského počtu dát, reprezentáciu vzťahov medzi nimi a ich vlastností.

Vizualizácia dát má dve základné funkcie. Prvou funkciou vizualizácie je analýza dát, alebo „data mining“. Pri ňom sa snažíme zistiť aké informácie nám vlastne dáta dávajú. Štandardný prístup na analýzu dát je získavanie odpovedí na otázky typu: „Majú dáta túto vlastnosť?“. Na získanie odpovede na túto otázku, musíme zvoliť správnu metódu vizualizácie, prípadne využiť ďalšie parametre vizualizácie, ako napríklad zmenu farby, zúženie výberu dát, zlúčenie dát, ... Po zobrazení vizualizácie v nej hľadáme vzory a štruktúry ktoré predstavujú hľadanú vlastnosť. V prípade, že máme správnu metódu vizualizácie, tak prítomnosť alebo neprítomnosť týchto vzorov hovorí o tom, či dáta majú alebo nemajú danú vlastnosť. Problémom a cieľom výskumu pre určité typy dát zostáva určenie správnej vizualizácie a príslušných vzorov a tvarov na určenie danej vlastnosti. Problém nastáva v prípade, že otázku otočíme: „Aké vlastnosti majú dáta?“. Vtedy môžeme postupovať len metódou „pokus omyl“ a buď testovať dáta na rôzne vlastnosti alebo vizualizovať dáta rôznymi metódami a zisťovať čo dané vzory znamenajú.

Druhovou funkciou vizualizácie je prezentácia. Ak prezentujem úspešnosť predaja výrobku, je pre mňa jednoduchšie ukázať graf, ktorý prezentuje tento predaj, ako

popisovať, v ktorom mesiaci predaj stúpal a v ktorom klesal. Dost' však záleží aký typ prezentácie si zvolíme. Ak by sme úspešnosť predaja prezentovali iba za sebou idúcimi číslami, predstavujúcimi počet predaných kusov za mesiac, prezentácia by bola omnoho menej čitateľnejšia a vyžadovala by si ďalšiu analýzu na získanie požadovanej informácie. Najefektívnejšie je na prezentovanie danej vlastnosti využiť tú istú metódu vizualizácie, ktorú sme použili pri analýze dát pri zistení či dáta majú danú vlastnosť.

Motivácia

Existuje viacero aplikácií na analýzu a vizualizáciu dát. Takmer všetky sú však špecializované na jednu z funkcií vizualizácie, teda buď na analýzu, alebo na prezentáciu dát. Ak budeme v texte hovoriť o aplikáciách na analýzu, alebo prezentáciu, myslíme tým aplikácie využívajúce vizualizáciu na tento účel.

Aplikácie na analýzu sú poväčšine komerčné a určené na profesionálne použitie. Teda pre bežného používateľa veľmi zložité na použitie, s množstvom funkcií, ktoré nikdy nevyužije. Výstupom z nich býva väčšinou len obrázok vizualizácie, teda na vytvorenie prezentácie je potrebná ďalšia úprava v grafickom editore.

Aplikácie na prezentáciu podporujú vkladanie rôznych grafov, obrázkov a textov, ale zvyčajne majú málo možností na konfiguráciu danej vizualizácie čím strácajú schopnosť väčšej analýzy dát.

Vizualizácia sa úspešne využíva pri analýze časovo závislých dát aj na ich prezentáciu. Nie všetky výskumné odvetvia však majú dostatočné skúsenosti a predstavu, čo všetko môžu pomocou vizualizácie dosiahnuť, keďže to nie je podstatou ich výskumu. S týmto javom sme sa stretli na Katedre zoológie Prírodovedeckej fakulty Univerzity Komenského v Bratislave, s ktorou sme pri tejto práci spolupracovali. Cieľom výskumu na tejto katedre je výskum živočíchov. Pri tejto práci dlhoročne zbierajú informácie o živočíchoch, analyzujú ich a vyvodzujú závery, ktoré prezentujú ako výsledok svojej práce.

Vzhľadom nato, že pracovníci tejto fakulty nie sú IT odborníci a analýza dát je síce nad rámec základných funkcií prezentačných aplikácií, ale analytická sila profesionálnych analytických aplikácií vysoko prevyšuje ich potreby. Využívajú na analýzu dát len matematickú štatistiku, ktorá síce dáva potrebné výsledky, ale je

príliš abstraktná a pomalá. Na prezentáciu jej výstupu prezentačné aplikácie málokedy stačia.

Cieľ

Cieľom diplomovej práce je naštudovať problematiku vizualizácie časovo závislých dát. Klasifikovať a porovnať rôzne typy vizualizácií a popísať ich využitie pri analýze rôznych typov dát.

V praktickej časti implementujeme niektoré z popísaných vizualizačných metód. Vytvoríme aplikáciu, ktorá spája vlastnosti prezentačných aj analytických aplikácií. Bude umožňovať hlbšiu analýzu dát prostredníctvom dodatočných vizualizačných parametrov. Zároveň poskytne jednoduchý nástroj na prezentáciu výsledkov. Prostredníctvom prostredia pracovnej plochy, využívajúcej metódu „drag and drop“, umožníme jednoduchú možnosť úpravy prezentácie a zároveň budeme mať možnosť porovnávať jednotlivé vizualizácie, čo nám poskytne ďalší analytický prvok. Bude dbať na jednoduché ovládanie, nevyžadujúce veľké IT nároky.

1 Základné pojmy

V tejto sekcii zadefinujeme niektoré pojmy používané v práci. Definície sú v niektorých prípadoch menej rigorózne ako zaužívané formálne definície. Je to v záujme jednoduchosti a pochopiteľnosti práce.

Časovo závislé dáta

Dáta, ktoré v sebe nesú aj informáciu o čase existencie meraného elementu. Teda funkciu meniacu sa v čase.

$$d = f(t)$$

Pre diskkrétne časové hodnoty je to množina prvkov:

$$D = \{(t_1, d_1), (t_2, d_2), (t_3, d_3), \dots, (t_n, d_n)\}$$

kde

$$d_i = f(t_i)$$

Dátový element

Hodnota d_i .

Jednorozmerné a viacrozmerné dáta

Vo všeobecnosti môžu d_i a d_j predstavovať hodnoty dvoch rôznych vlastností. Pre ujasnenie budeme za jednorozmerné dáta považovať množinu D , kde všetky d_i budú popisovať jednu vlastnosť. Za n -rozmerné (viacrozmerné) dáta budeme považovať množinu D , ak $D = D_1 \cup D_2 \cup D_3 \cup \dots \cup D_n$, kde D_i sú jednorozmerné dáta a D_i, D_j popisujú inú vlastnosť.

Jeden rozmer viacrozmerných dát

Pre viacrozmerné dáta $D = D_1 \cup D_2 \cup D_3 \cup \dots \cup D_n$ je jeden rozmer množina D_i .

Teraz si definujeme typy časových dát, podľa časovej osi, na ktorej sa nachádzajú. Príklady k daným typom sú uvedené v odseku 2.2 Taxonómia.

Diskrétné časové body

Časová os sa skladá z diskrétnych časových bodov, (t_i, d_i) hovorí o tom, že dátový element d_i existoval v čase t_i na okamih nulovej dĺžky.

Časový interval

Časová os sa skladá z časových intervalov, (t_i, d_i) hovorí o tom, že dátový element d_i existoval v určitom intervale, napríklad deň, mesiac, rok, ...

Lineárny čas

Časovú os predstavuje začiatkový bod a polpriamka idúca do budúcnosti. Teda predpokladá, že množina $T = \{t_1, t_2, t_3, \dots, t_n\}$ je usporiadaná: $t_i \leq t_{i+1}$

Cyklický čas

Dáva do súvisu jednotlivé periódy a spája príslušné fázy periódy. Časovú os predstavuje kružnica, alebo špirála. Nemá veľký zmysel hovoriť o usporiadaní na globálnej úrovni, lebo ráno je aj pred aj po večere.

Sériový čas

O t_i a t_j vieme rozhodnúť, len ktorý je pred ktorým. Teda nemusíme vedieť presné hodnoty t_i , ale elementy vieme usporiadať podľa ich existencie.

Súvislý čas

Vieme presné hodnoty t_i a teda vieme zistiť časový rozdiel výskytu dvoch elementov.

Postupný čas

Časová os je polpriamka. Dátové elementy sa vyskytujú zaradom, d_{i+1} nasleduje za d_i .

Stromový čas

Časová os je vetviaci sa strom. Čas sa môže vetviť, teda za d_i nasleduje d_{i+1} aj d'_{i+1} aj d''_{i+1} ... Táto časová os podporuje robenie rozhodnutí.

Null element

Null element je hodnota d_i , ktorá sa nepovažuje za nameranú hodnotu. Vyplňuje časový interval prázdnu hodnotou, aby časový rozdiel medzi nameranými hodnotami bol konštantný.

Nulová hodnota

Je hodnota ktorá je pri vizualizácii zvýraznená a od ktorej sa počítajú vzdialenosti pri mriežke.

2 Časovo závislé dáta

Z časovými dátami sa stretávame v bežnom živote, či už každé ráno pri pozeraní predpovede počasia, alebo mesačne pri kontrolovaní výplatnej pásky. Analýza týchto dát nemusí byť vždy jednoduchá. Vo vede, inžinierstve či biznise je to jeden z najvážnejších problémov, ktorého úspešné riešenie môže ušetriť veľa času alebo finančných prostriedkov. S príchodom počítačov, lepších meracích zariadení a všeobecným hladom po informáciách sa počet časových dát výrazne zvyšuje, čo ešte viac zvyrazňuje snahu o efektívnu analýzu týchto dát.

2.1 Vlastnosti

Analýza dát znamená hľadanie ich vlastností. Časovo závislé dáta majú oproti štandardným dátam prvok času, ktorý nás vedie k nasledujúcim otázkam [4]:

- Existoval dátový element v konkrétnom čase?
- Kedy dátový element existoval v čase? Je tam cyklická závislosť?
- Ako dlho trval dátový element?
- Ako často sa dátový element vyskytuje?
- Ako rýchlo sa dátový element mení, alebo aký je rozdiel medzi jedným a druhým dátovým elementom v čase?
- V akom poradí sa dátové elementy objavujú?
- Vyskytujú sa dátové elementy spoločne?

Pri viacrozmerných dátach nás môžu navyše zaujímať vzťahy medzi jednotlivými rozmermi:

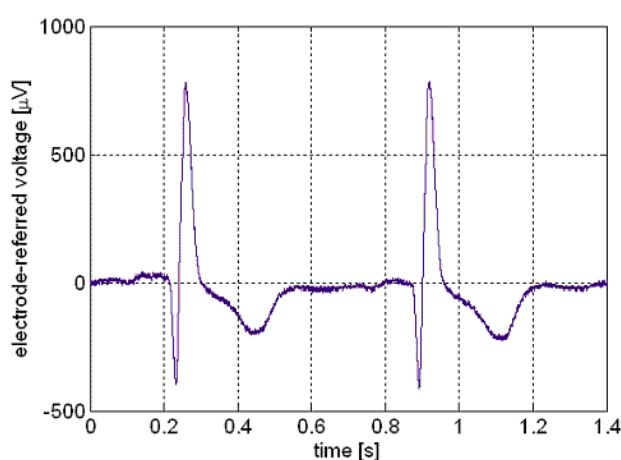
- Je rozmer A závislý od rozmeru B?
- Aká je závislosť medzi rozmerom A a B?

2.2 Taxonómia

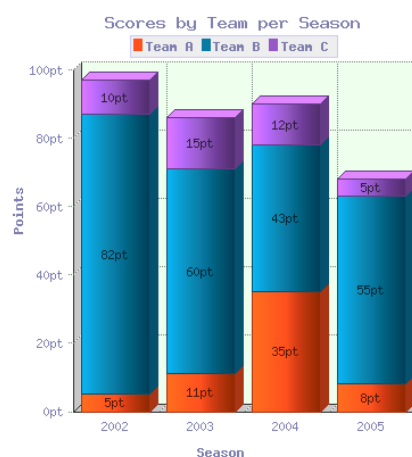
Podľa časovej osi, v ktorej sa dáta vyskytujú vieme časovo závislé dáta presnejšie špecifikovať. Podľa Franka (1998) [5] sa delia nasledovne:

2.2.1 Diskrétné časové body - Časový interval

Diskrétné časové body sa vyskytujú napríklad pri dátach, ktoré vznikajú pomocou meracích prístrojov, ktoré merajú okamžitý stav. Napríklad EKG, teplomer, mikrofón, osciloskop, ... Tie síce zaznamenávajú spojité signály, ale pri jeho digitalizácii vznikajú diskrétné body. Môžeme si to predstaviť ako keď chceme krivku popísať bodmi, ktoré na nej ležia. Časové intervaly vznikajú napríklad, ak dáta s diskretnými časovými bodmi, rozdelíme na podmnožiny reprezentujúce určitý časový interval (minúta, deň, rok, ...) a na tieto podmnožiny aplikujeme matematickú operáciu typu suma, minimum, maximum, priemer, ...



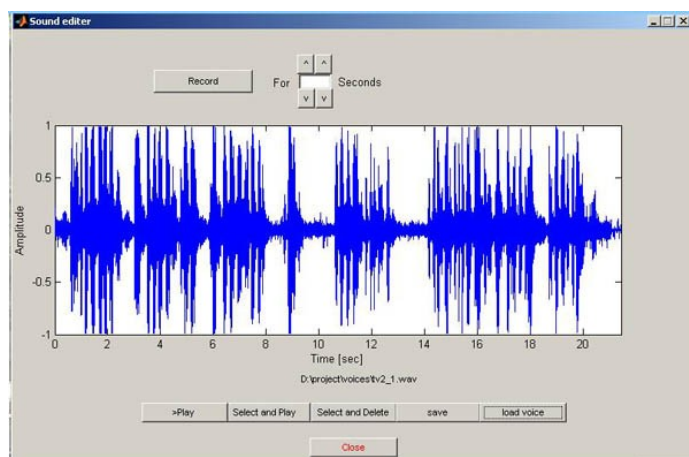
1. Obr.: Diskrétné časové body



2. Obr.: Časový interval

2.2.2 Lineárny čas - Cyklický čas

Lineárny čas je štandardný čas, kedy dáta začínajú v určitý moment a pokračujú do budúcnosti. Využíva sa pri väčšine dát. Cyklický čas využíva sa ak chceme poukázať na periodickosť dát ako napríklad denná teplota.



3. Obr.: Lineárny čas



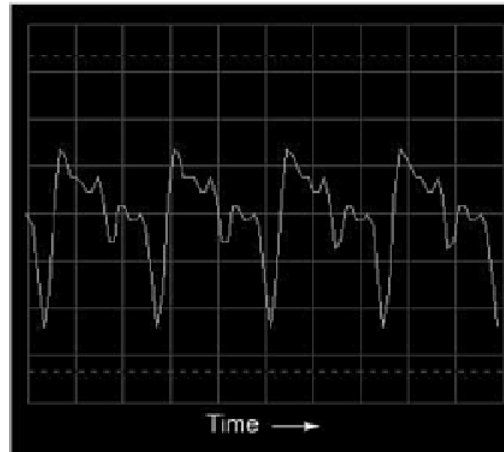
4. Obr.: Cyklický čas

2.2.3 Sériový čas - Súvislý čas

Sériový čas sa vyskytuje pri dátach, ktoré sme merali za sebou, ale nezaujímal nás presný čas merania. Napríklad pri meraní výskytu zveri, nám stačí vedieť že v daný deň sme napočítali 12 kusov, ale nemusíme vedieť či to bolo ráno alebo večer. Súvislý čas zvyčajne dostávame pri podrobnom meraní, keď meriame v časových intervaloch a zaujíma nás presný čas namerania dát.

SENATE Calendar		SÉNAT Calendrier	
Subject to change		Sujet à modification	
JAN.	JANV.	FEB.	MARS
S M T W T F S	S M T W T F S	S M T W T F S	S M T W T F S
D L M M J V S	D L M M J V S	D L M M J V S	D L M M J V S
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12
13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19
20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26
27 28 29 30 31	27 28 29 30 31	27 28 29 30 31	27 28 29 30 31
APRIL	AVRIL	MAY	JUNE
S M T W T F S	S M T W T F S	S M T W T F S	S M T W T F S
D L M M J V S	D L M M J V S	D L M M J V S	D L M M J V S
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12
13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19
20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26
27 28 29 30 31	27 28 29 30 31	27 28 29 30 31	27 28 29 30 31
JULY	JUL.	AUG.	AOÛT
S M T W T F S	S M T W T F S	S M T W T F S	S M T W T F S
D L M M J V S	D L M M J V S	D L M M J V S	D L M M J V S
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12
13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19
20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26
27 28 29 30 31	27 28 29 30 31	27 28 29 30 31	27 28 29 30 31
OCT.	OCT.	NOV.	DÉC.
S M T W T F S	S M T W T F S	S M T W T F S	S M T W T F S
D L M M J V S	D L M M J V S	D L M M J V S	D L M M J V S
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12	6 7 8 9 10 11 12
13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19	13 14 15 16 17 18 19
20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26	20 21 22 23 24 25 26
27 28 29 30 31	27 28 29 30 31	27 28 29 30 31	27 28 29 30 31

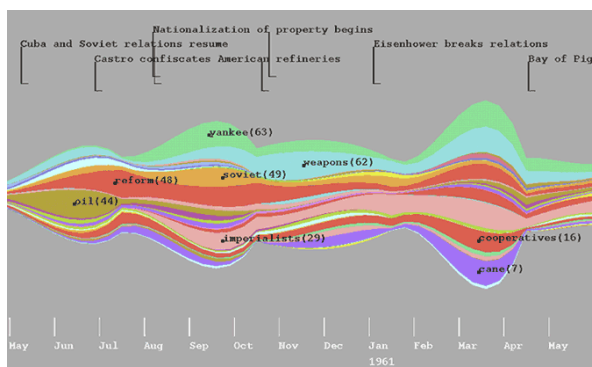
5. Obr.: Sériový čas



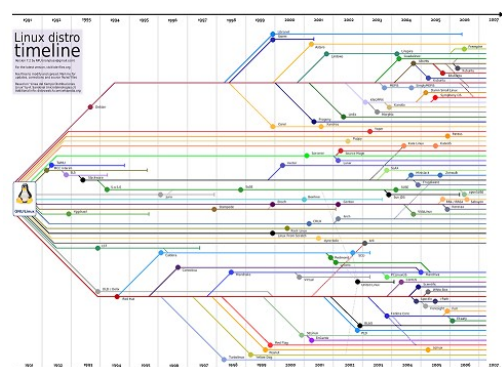
6. Obr.: Súvislý čas

2.2.4 Postupný čas – Stromový čas

Postupný čas je štandardný čas keď meriame určitú vlastnosť v čase. Stromový čas sa využíva napríklad pri scenároch, kde nám umožňuje robiť rozhodnutia a teda vytvára rôzne alternatívy. Prípadne pri meraní určitej vlastnosti, ktorá sa v danom momente rozdelila na viac vlastností a pokračujeme v monitorovaní každej z nich.



7. Obr.: Postupný čas



8. Obr.: Stromový čas

3 Vizualizácia

Vizualizácia je prvá forma, ktorú človek vymyslel na uchovávanie informácií, o čom svedčia nástenné maľby, ktoré sú staré už vyše 30.000 rokov. Tieto zaznamenávali hlavne veľké zvieratá a priebeh ich lovu. Kartografia je po nástenných maľbách druhá najstaršia metóda uchovávania informácií, ktorú používali už okolo roku 7500 p.n.l. a používa sa dodnes. Do vynájdenia písma sa ďalej používali hlavne piktogramy, ktoré slúžili na zaznamenávanie počtu dobytku a zásob [1]. S vynájdením písma a rozvojom spoločnosti sa objavovali nové typy vizualizácie, hlavne rôzne odrody grafov a diagramov. Digitálna technika ešte viac umožnila vznik a využitie vizualizácií. V súčasnosti sa získavajú hlavne viacrozmerné a početné dáta. To spôsobuje, že nové vizualizácie sa vyvíjajú hlavne na analýzu týchto typov dát.

3.1 Vlastnosti

Výber vizualizácie a k nej určenie vzorov na zistenie či dáta majú danú vlastnosť je hlavný problém pri analýze dát. Pri výbere vhodnej vizualizácie sú podstatné tri veci: typ zobrazovaných dát a pozorovacia schopnosť človeka pre ktorého je vizualizácia určená a vlastnosť, ktorú sa snažíme nájsť.

Rôzne typy a množstvo dát si vyžadujú rôzne vizualizácie. Na dobrú analýzu nemôžeme rovnako vizualizovať predpoveď počasia na najbližšie 3 dni a údaje o počasi za posledných 10 rokov. Pre najbližšie 3 dni si môžeme dovoliť nakresliť slniečko, obláčik, alebo mrak a pri nich teplotnú hodnotu. Pri dátach za posledných 10 rokov by bolo výhodnejšie zobrazit' stĺpcový graf, kde by jeho výška predstavovala teplotnú hodnotu a farba stĺpca oblačnosť.

Pozorovacia schopnosť človeka ovplyvňuje schopnosť rozpoznávať príznaky a vzory, ktoré predstavujú vizualizovanú vlastnosť. Ak treba porovnať dve hodnoty, tak dospelému človeku ich stačí napísať. Ak by sme ich ukázali malému dieťaťu nevedelo by to. Preň by bolo výhodnejšie zobrazit' ich napríklad ako dva štvorce, kde ich veľkosť by predstavovala hodnotu čísla. Rovnako však môžeme dospelému človeku ukázať vizualizáciu pomocou paralelných súradníc, ak ju predtým ešte nevidel, a jej analýza mu zaberie omnoho viac času ako keby sme vizualizovali niektorou štandardnou metódou.

Cieľom vizualizácie je zistenie, či majú dáta určitú vlastnosť a teda nájdienia vzoru prislúchajúceho k danej vlastnosti. Vlastnosť teda ovplyvňuje, ktorá vizualizácia je vhodnejšia. Máme informácie o veľkosti prichádzajúcich a odchádzajúcich dát, ktoré sme prečerpali na mobilnom internete. Ak chceme zistiť či sme mali viac prijatých, alebo viac odoslaných dát je nám najvýhodnejší koláčový diagram, ktorý na jednej strane zobrazuje sumu pochádzajúcich dát a na druhej sumu odchádzajúcich dát. Ak chceme zistiť celkový objem dát, môžeme použiť čiarový graf, kde hodnota v danom čase bude zodpovedať dovedy prenesenému objemu dát. Koláčový diagram nám však túto vlastnosť neposkytne.

V nasledujúcom odstavci budeme popisovať aké vlastnosti by mala mať dobrá vizualizácia prezentovaná na PC. [2]

- Priestrannosť

Použiť čo najmenšiu možnú plochu pri zachovaní informácie, alebo minimalizovať priestor pri maximálnej informácii.

- Dynamickosť

Schopnosť vidieť priebežný vývoj. Plynulý prechod vizualizácie pri zmene parametrov na udržanie focus-u. Tu sa využíva animácia.

- Multi-parameter

Umožniť súčasné porovnávanie vizualizácie pri rôznych nastaveniach.

- Vzor

Uľahčiť rozpoznávanie vzorov a zoskupení v dátach. Zvýraznenie štruktúr, ktoré predstavujú zobrazovanú vlastnosť.

- Vnímateľnosť

Jednoduché vnímanie dosiahnuté pomocou vizuálnych podnetov.

- Sémantickosť

Pochopiť význam vizuálnej reprezentácie. Identifikovanie, na ktoré vlastnosti sa vizualizácia zameriava.

- Všeobecnosť

Schopnosť aplikovať rovnaké vzory na rôzne vizuálne nastavenia.

- Výstup

Umožniť nahratie a spätné zobrazenie predtým použitých vybraných vizuálnych efektov.

- Personalizácia

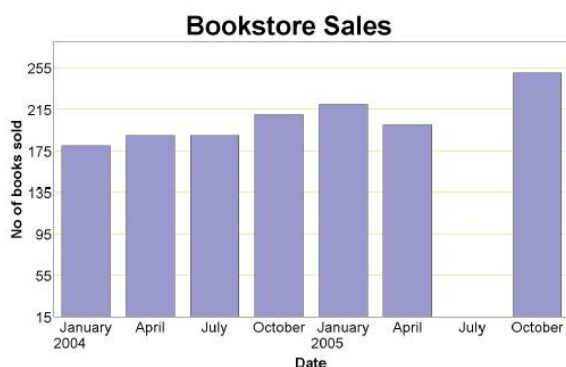
Vnímanie je subjektívne a je potrebné prispôbiť vizualizáciu osobným potrebám.

3.2 Taxonómia podľa vizualizačných vlastností

Každá vizualizácia má svoje vlastné špecifiká. Môžeme ich však rozdeliť do skupín a to podľa nasledujúcich kritérií: [6]

3.2.1 2D – 3D

2D je klasická reprezentácia v rovine pomocou obrázkov. 3D reprezentácia nie je reálne v 3D priestore, keďže zatiaľ nám to technika neumožňuje. Pri 3D sa teda využíva perspektíva a premietanie, prípadne interaktívna 3D reprezentácia podobná virtuálnej realite. Keďže sa 3D priestor premietá do 2D, zvyčajne sa stráca informácia. Na plné využitie tejto vizualizácie je teda zvyčajne potrebná interakcia. Často sa z 2D stáva 3D reprezentácia pridaním z-ovej osi, ktorá nenesie žiadne informácie. Tento krok sa robí len pre prezentačnú vizualizáciu, aby vizualizácia vyzerala krajšie, nemá žiadny význam pri analýze dát.



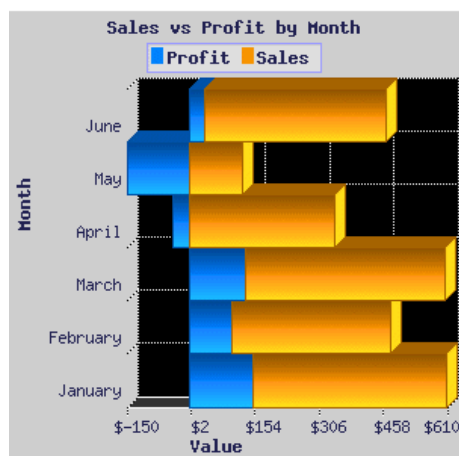
9. Obr.: 2D viz.



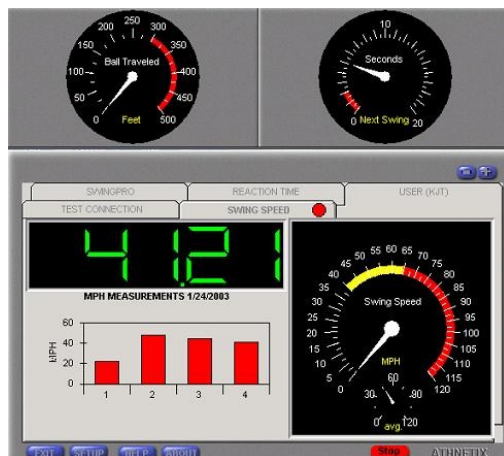
10. Obr.: 3D viz.

3.2.2 Statická – Dynamická

Statická vizualizácia pozostáva len z obyčajného obrázku. Dynamická poskytuje aj dynamický aspekt. Väčšinou to býva interakcia s vizualizáciou, buď prostredníctvom zmeny pohľadu, alebo nastavenia, prípadne animácia.



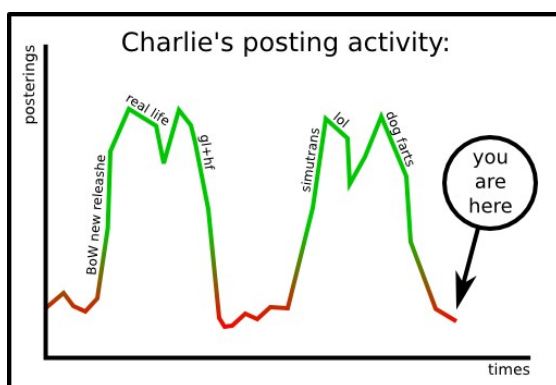
11. Obr.: Statická viz.



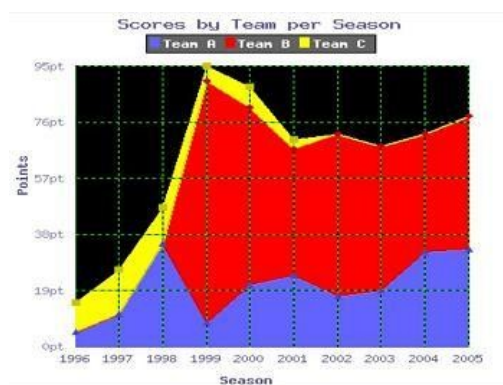
12. Obr.: Dynamická viz.

3.2.3 Jednorozmerná – Viacrozmerná

Jednorozmerné zobrazujú len jeden dátový element pre jeden čas. Viacrozmerné zobrazujú viac čo býva pri veľkom počte dosť obtiažne. Tieto metódy sa stále vyvíjajú.



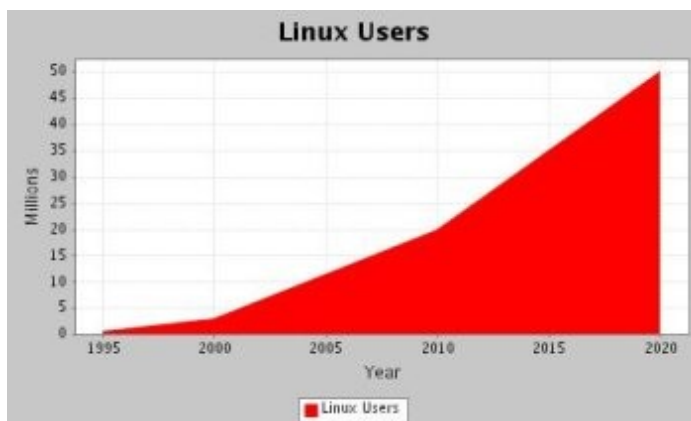
13. Obr.: Jednorozmerná viz.



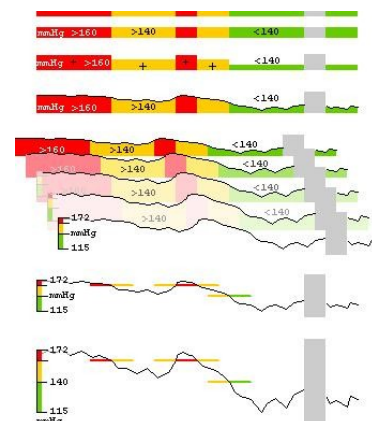
14. Obr.: Viacrozmerná viz.

3.2.4 Riadená dátami – Riadená udalosťami

Pri dátami riadenej vizualizácii sa sústreďíme na vizualizáciu dát a na vlastnosti súvisiace s hodnotami dát. Pri udalosťami riadenej vizualizácii sa zase zameriavame na čas, kedy nastala určitá zmena dát. Zvýrazňuje body záujmu.



15. Obr.: Viz. riadená dátami



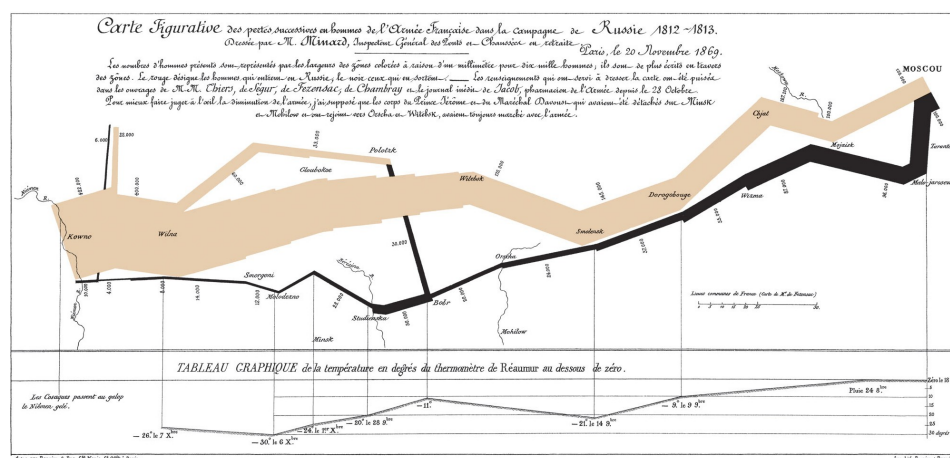
16. Obr.: Viz. riadená udalosťami

3.3 Prehľad niektorých vizualizačných metód

3.3.1 Konvenčné vizualizačné metódy

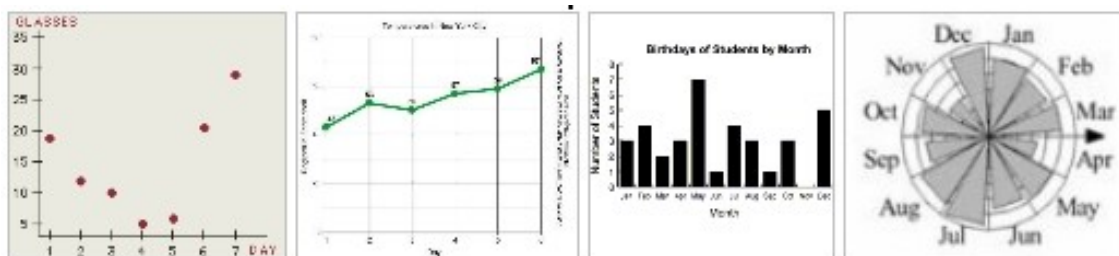
Sú to bežne používané vizualizácie, jednoduché na porozumenie a hľadanie vzorov. Jedná sa hlavne o statické vizualizácie v 2D priestore. Väčšina z nich je však omedzená na jednorozmerné dáta či menšie množstvo zobrazovaných dát. Umožňujú hlavne súhrn kvantitatívnych uzáverov. V závislosti od škálovania umožňujú jednoduché zisťovanie dátových hodnôt a časových informácií.

Jeden z najlepších príkladov vizualizácie časovo závislých dát je Mirandov graf zobrazujúci postup Napoleonských vojsk v kampani proti Rusku z rokov 1812-1813. [8] Zobrazuje, ako sa menila v závislosti od času geografická poloha a početnosť vojsk aj teplota.



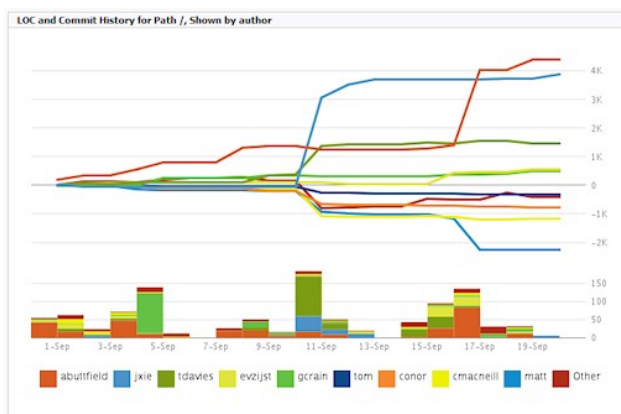
17. Obr.: Postup Napoleonských vojsk

Človek je od mladosti učení predstavovať si čas, buď ako priamku, alebo ako hodiny, teda kruh. Teda časová os zvykne mať tvar priamky, alebo kruhu/špirály, podľa toho či sa jedná o *Lineárny*, alebo *Cyklický čas*. Keďže dáta majú väčšinou číselný charakter a hodnota čísla je predstavovaná jeho vzdialenosťou od začiatku číselnej osi, tak vzdialenosť vykresleného bodu od časovej osi predstavuje hodnotu elementu. Podľa typu časovej osi tak dostávame *Bodový graf* pre *Časové body/Súvislý čas*, *Čiarový graf* pre *Časové body/Súvislý čas*, *Stĺpcový graf* pre *Intervalový čas/Sériový čas* a *Cyklický graf* pre *Cyklický čas*. Všetky tieto vizualizácie sú určené pre jednorozmerné dáta. Sú veľmi jednoduché na pochopenie, vzhľadom na škálovanie ponúkajú dostatočne presný odhad na určenie času existencie a hodnotu elementu.

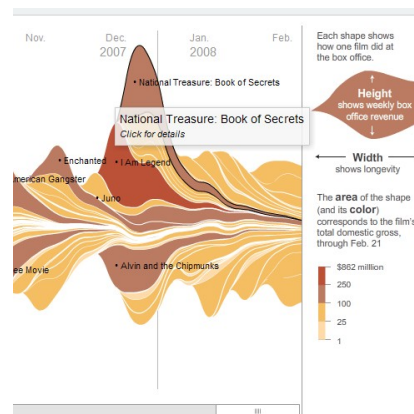


18. Obr.: Bodový, Čiarový, Stĺpcový a Cyklický graf

Ak prekryjeme dva a viac grafov s rovnakou časovou osou dostaneme jeden *Zložený graf*. Podobným princípom prekrytia *Stĺpcových grafov* dostávame *Graf zmien*. Tieto grafy sú použiteľné len pri zjednotení malého počtu grafov, pri väčšom počte sa v nich stráca prehľad. Majú vlastnosti ako grafy, z ktorých vznikli a naviac získavajú schopnosť porovnávať jednotlivé dátové sady. V prípade, že hodnoty elementov budeme sčítovať, dostaneme *Kumulovaný graf* zo *Stĺpcových grafov* alebo *ThemeRiver* [13] z *Čiarového grafu*. Získavame pri nich informáciu o sume elementov v určitom čase, ale zhoršuje sa schopnosť vyčítať hodnotu jednotlivých elementov. Ne strácajú však prehľadnosť, pri väčšom počte zlúčených grafov, až tak výrazne.

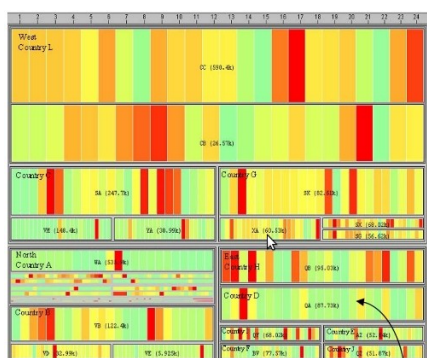


19. Obr.: Zložený a Kumulovaný graf

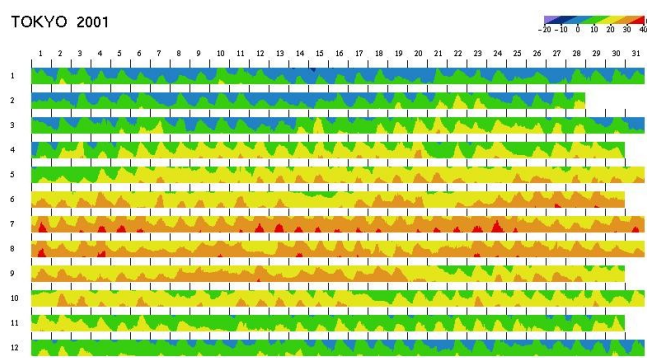


20. Obr.: ThemeRiver

Okrem vzdialenosti od časovej osi sa môžu hodnoty elementov mapovať aj na iný parameter ako napríklad farbu. Podľa toho či pre jednu hodnotu použijeme jednu farbu, alebo dve farby dostávame *Jednofarebné*, alebo *Dvojfarebné ofarbovanie* [11], [12]. Tieto vizualizácie majú oproti vyššie spomenutým vizualizáciám výhodu, že nepotrebujú hodnotovú os. Stačí im úzky pásik. Strácajú však schopnosť presného určenia hodnoty elementu. Na jej určenie je potrebná legenda hodnotovej škály, no aj z nej sa presná hodnota ťažko dočítava. *Dvojfarebné ofarbovanie* sa túto neschopnosť snaží riešiť použitím len niekoľkých farieb na hodnotovej škále, čím uľahčí identifikáciu hodnoty, i keď presná identifikácia je stále zložitá. Dvojfarebné ofarbenie napomáha aj lepšej rozoznatelnosti zmeny hodnoty za sebou idúcimi elementami, keďže na malom hodnotovom intervale z hodnotovej škály, má *Dvojfarebné ofarbovanie* podobu stĺpcového grafu.



21. Obr.: Jednofarebné ofarbenie



22. Obr.: Dvojfarebné ofarbenie

3.3.2 Vizualizačné metódy pre viacrozmerné a početné dáta

Rozvoj digitálnej techniky umožňuje vznik čoraz väčších databáz. Vďaka tomu je vizualizácia viacrozmerných a početných dát stále predmetom výskumu. Naša implementácia neplánuje vizualizáciu takýchto dát, preto uvedieme len niektoré názvy vizualizácii s odkazmi na články, ktoré ich predstavujú a nebudeme ich podrobnejšie rozoberať.

- Spiral graph [14]
- Calendar View [15]
- Lexis Pencils [16]
- SpiraClock [18]
- Wormplots [17]
- Parallel Coordinates [10]

- TimeWheel [19]
- MultiComb [20]
- Time Dependent Histogram [21]

4 Návrh aplikácie

Pri návrhu aplikácie sme boli motivovaný potrebami Katedry zoológie Prírodovedeckej fakulty Univerzity Komenského v Bratislave. Potrebovali nástroj, ktorý by jednoduchou formou umožnil analyzovať časovo závislé dáta a hľadať vzťahy medzi nimi a prezentovať nájdené závery.

4.1 Dáta

Dáta, ktorými disponuje Katedra zoológie sú nameraná vo voľnej prírode pracovníkmi a študentmi. Jedná sa o ručné merania rôznych vlastností živočíchov. Teda interval medzi jednotlivými meraniami nie je rovnaký. Ak by sme požadovali pre dáta konštantný interval medzi meraniami, niektoré dáta by chýbali. Namerané dáta sa budú porovnávať s inými údajmi, napríklad o počasí, ktoré sú kompletne. Kvôli kompatibilite sme zaviedli „null element“, ako hodnotu dát v čase, kde chýba nameraná hodnota. To nám umožní jednoducho porovnávať kompletne a nekompletné dáta a hľadať medzi nimi závislosti.

Dáta sa budú načítavať z textového súboru. Textový súbor bude mať nasledujúci formát: každý riadok bude predstavovať jedny dáta. Riadok bude začínať dátumom namerania prvého elementu v tvare YYYY\MM\DD HH:MM:SS, ďalej bude pokračovať postupne hodnotami jednotlivých elementov oddelených medzerou, alebo tabulátorom. V ďalších riadkoch môžu byť zase uložené dáta s rovnakým formátom. V aplikácii sa bude dať nastaviť dĺžka periódy merania elementov.

Textový súbor poskytuje vhodnú možnosť na predspracovanie dát. Jeho ľahkou editáciou vieme rozdeliť jedny dáta na viacero menších dát, napríklad podľa periódy, čo nám neskôr umožní vizualizovať dáta s *Lineárnym časom* ako dáta s *Cyklickým časom*.

4.2 Prostredie

Užívateľské prostredie má byť v prvom rade jednoduché na ovládanie, keďže aplikácie je určená pre bežných ľudí, ktorý nemusia mať veľké skúsenosti s IT technikou.

4.2.1 Správa dát

Aplikácia bude podporovať porovnávanie dát, bude teda umožňovať súčasné načítanie viacerých databáz.

Rôznym dátam bude môcť byť priradený rôzny typ vizualizácie, čím dostaneme viac možností na skúmanie vzťahov medzi dátami. Jedným dátam sa bude dať priradiť viacero vizualizácií, prípadne viac krát tá istá, čo umožní porovnávanie vplyvu rôznych nastavení na vizualizáciu. Obr.: Analýza

Jednotlivým dátam sa bude dať nastaviť hodnota „null elementu“, čas a intervaly v akých boli dáta namerané a pre lepšiu orientáciu pri práci s nimi im budeme môcť priradiť aj názov.

4.2.2 Pracovná plocha

Pracovná plocha skĺbi analýzu a prezentáciu dát. Implementujeme tri vizualizačné metódy: *Čiarový graf*, *Jednofarebné ofarbovanie* a *Dvojfarebné ofarbovanie*. Pre každú vizualizáciu poskytneme vlastný panel pre nastavenie parametrov vizualizácie, ktorý bude mať jednotnú formu pre zjednodušenie ovládania.

Implementované bude:

- Zužovanie výberu časového úseku pre zobrazenie. Tým získame schopnosť focus-u na určitý časový úsek záujmu. Obr.: Prehľad
- Možnosť zapnutia a vypnutia zobrazenia názvu vizualizácie, popisku časovej osi a popisku hodnotovej osi. Tieto nastavenia slúžia na lepšiu prezentačnú schopnosť. Obr.: Prehľad
- Zmena ofarbenia vizualizácie. Táto možnosť slúži aj pre prezentáciu, ale hlavne pri analýze, kde niektoré typy ofarbenia v sebe nesú ďalšiu informáciu o dátach. Pri *Dvojfarebnom ofarbovaní* bude aj možnosť nastaviť z koľkých farieb sa má skladať hodnotová škála. Obr.: Analýza
- *Čiarový graf* a *Dvojfarebné ofarbovanie* bude možné otočiť okolo x-ovej osi. Spolu s priložením dvoch vizualizácií k sebe dostaneme ďalšiu možnosť porovnávania dvoch dátových atribútov. Obr.: Prehľad
- Nastavenia „null elementu“. Ponuka troch možností, ako zobrazovať „null element“ a nastavenie jeho farby. Obr.: Prehľad

- Nastavenia mriežky. Hustota mriežky v časovej a v hodnotovej osi.
- Nastavenie polohy, veľkosti a odsadenia vizualizácie.
- Nastavenie rozsahu hodnotovej škály. Poslúži na zjednocovanie nastavení vizualizácií a spolu s určitými ofarbeniami môže slúžiť na hľadanie prahovej hodnoty. Obr.: Prehľad

Použitím metódy „drag and drop“ pri pohybe a škálovaní vizualizácií dostaneme vynikajúci nástroj na porovnávanie vizualizácií a dát a to prostredníctvom prekrývania a priloženia vizualizácií vedľa seba. Pohyb vizualizácií bude umožňovať zakotvenie o druhú vizualizáciu, teda pri posúvaní prvej vizualizácie sa bude posúvať aj druhá a všetky ďalšie naviazané. Na zjednodušenie určenia polohy pri posúvaní sa bude dať zapnúť posúvanie po vrcholoch mriežky. Pre potrebu zväčšenia pracovnej plochy ju bude možné priblížiť alebo oddialiť.

4.3 Ďalšia funkcionlita

Implementované budú štandardné prvky aplikácií ako uloženie a načítanie vizualizácie. Ďalej uloženie obrazovky pracovnej plochy, animáciu rastu časovej osi, pridávanie textu na prezentáciu na pracovnej ploche a zobrazenie hodnotových škál pre vizualizácie *Jednofarebné ofarbovanie* a *Dvojfarebné ofarbenie*. Dodatočné nastavenia, ktoré sa nevyužívajú na analýzu dát a pre prezentáciu nie sú veľmi podstatné sa budú načítavať z externého súboru, aby bola možnosť ich zmeny, ale aby zbytočne nekomplikovali aplikáciu.

5 Realizácia

Vizualizačný program bol naprogramovaný v jazyku C++ v prostredí Borland C++ Builder 6. Na zobrazovanie využíva OpenGL.

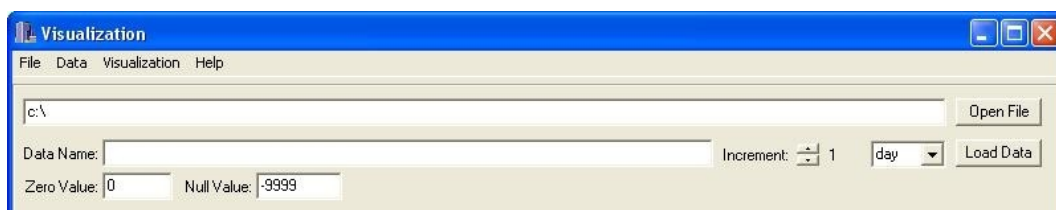
5.1 Prostredie

Program obsahuje menu, ktoré popisuje celú jeho štruktúru.

5.1.1 File

5.1.1.1 New Data (F2)

Okno, ktoré umožňuje načítanie dát z textového súboru. Textový súbor musí mať korektný formát popísaný v sekcii 4.1Dáta. Ďalej sa sa tu nastavuje hodnota „null element“, „nulová hodnota“ a veľkosť periódy merania elementov pre načítané dáta.



23. Obr.: New Data

5.1.1.2 Save Visualization (Ctrl+S)

Uloží načítané dáta, zvolené vizualizácie a všetky nastavenia vizualizácie práve tak ako sa momentálne nachádzajú.

5.1.1.3 Load Visualization (Ctrl+O)

Zruší všetky momentálne načítané dáta a vizualizácie a načíta uloženú vizualizáciu zo súboru.

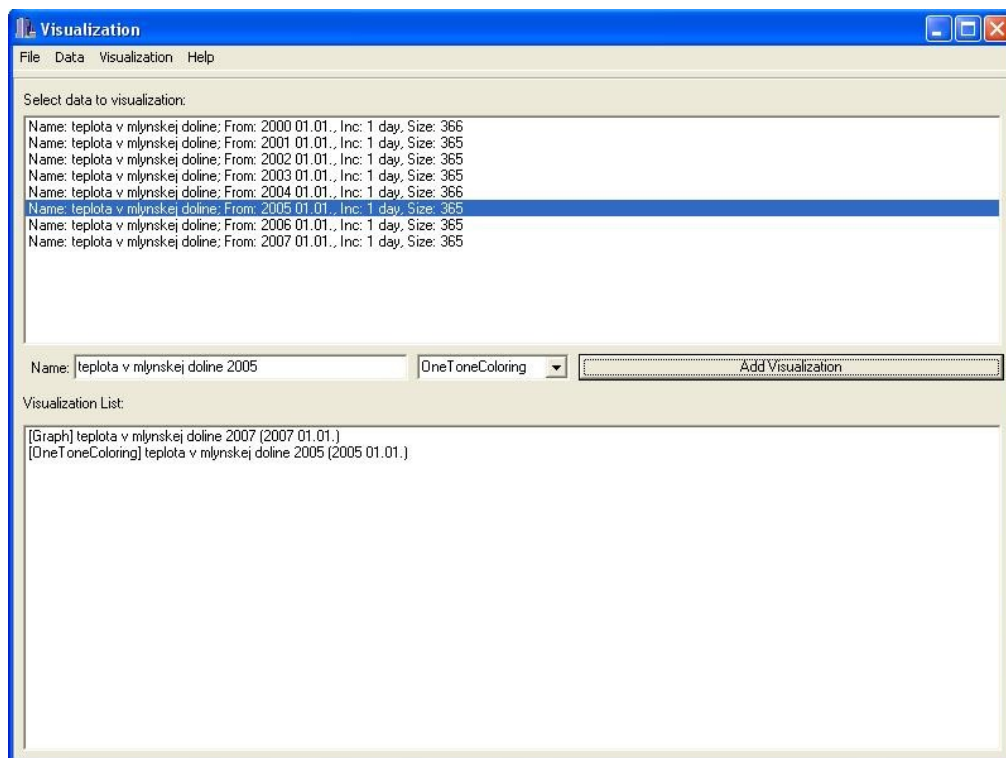
5.1.1.4 Save Image (Ctrl+I)

Uloží aktuálny obraz pracovnej plochy do obrázka. Na výber sú dva formáty, JPEG a PNG.

5.1.2 Data (F3)

V tomto okne sa k dátam pridávajú vizualizácie. V **Select data to visualization** sa nachádzajú načítané dáta zo súborov. Vo **Visualization List** sa nachádzajú práve zobrazované vizualizácie. Prostredníctvom ComboBox-u sa zvolí typ vizualizácie.

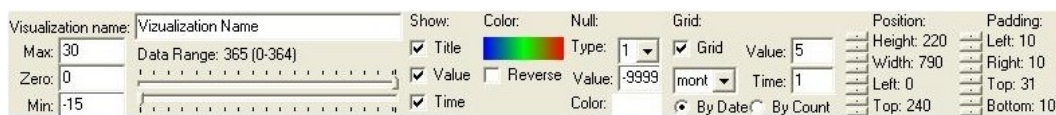
Kliknutím na **Add Visualization** sa dáta označené v **Select data to visualization** začnú zobrazovať pomocou zvolenej vizualizácie a pridajú sa do **Visualization List**. Vizualizácia sa dá vytvoriť aj dvojklikom na dáta v **Select data to visualization**. Z **Visualization List** sa dá vizualizácia odstrániť dvojklikom.



24. Obr.: Data

5.1.3 Visualization (F4)

Toto okno je okno pracovnej plochy. V ňom prebieha samotná vizualizácia a analýza dát. Na editáciu vizualizácie je potrebné najprv ju označiť a to kliknutím na ňu. Po označení sa objaví editovací panel, ktorý umožňuje zmenu nastavenia vizualizácie.



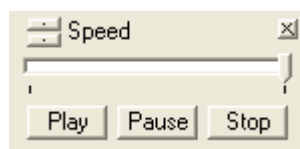
25. Obr.: Editovací panel

- Visualization name - názov vizualizácie.
- Max, Zero, Min - nastavenie maximálnej, minimálnej a nulovej hodnoty.
- Data Range - zúženie výberu dát od - do.
- Show - zobrazovanie názvu, času a hodnôt.
- Null - zobrazovanie „null element“.
- Color - farba vizualizácie.

- Reverse - otočenie vizualizácie.
- Grid - zobrazovanie mriežky, šírka mriežky pre hodnoty a pre čas/počet.
- Position - pozícia a veľkosť vizualizácie.
- Padding - vzdialenosť vykreslenia od vonkajšieho okraja vizualizácie.

5.1.4 Animation (F5)

Tento panel slúži na spustenie animácie. Dá sa tu aj nastaviť rýchlosť animácie a pretáčať animáciu.



26. Obr.: Animation

5.1.5 Help (F1)

Zobrazí html súbor s návodom na obsluhu programu.

5.2 Pohyb po pracovnej ploche

Pracovná plocha využíva princípy „drag and drop“, teda posun a škálovanie vizualizácií sa robí jednoduchým kliknutím myši a ťahaním. Pre jednoduchšie zarovnávanie vizualizácií môžeme držaním CTRL zabezpečiť aby sa vizualizácia pohybovala len po mriežke o vzdialenosti 20px. Ak chceme pracovať s vizualizáciou, potrebujeme ju najprv označiť a to kliknutím myši na ňu. Ak máme označenú vizualizáciu a kurzorom chodíme ponad vizualizované dáta zobrazuje sa hodnota a čas výskytu elementu, nad ktorým je kurzor.

Po pravom kliknutí na plochu sa zobrazí popup:



27. Obr.: Popup

- add Text - pridá Text do vizualizácie.
- add OneTone - pridá hodnotovú škálu pre *Jednofarebné ofarbovanie*.

- add TwoTone - pridá hodnotovú škálu pre *Dvojfarebné ofarbovanie*.
- delete Vis - ak bolo kliknuté na označenú vizualizáciu, tak ju zmaže.

6 Prehľad niekoľkých vizualizácií

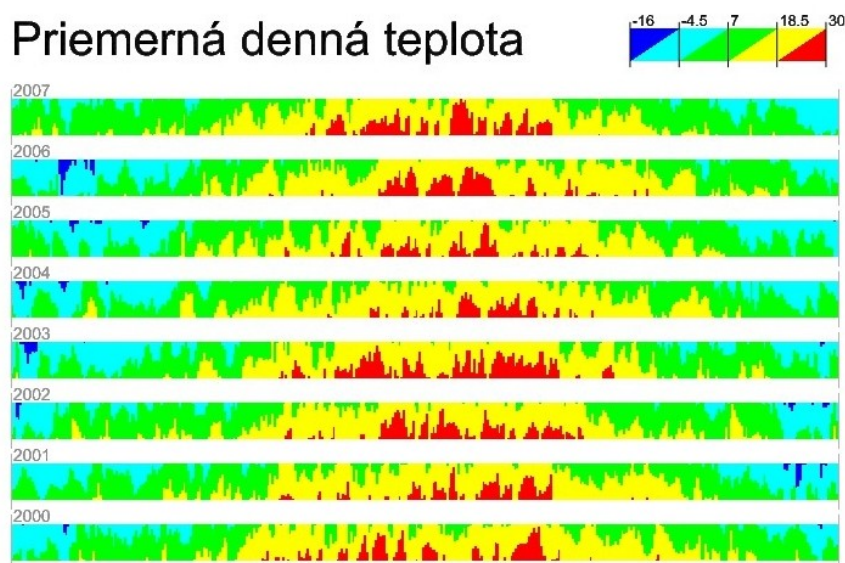
V tejto kapitole si ukážeme niekoľko možností vizualizácie, ktoré nám táto aplikácia poskytuje.

Na vizualizáciu sme použili dáta, ktoré nám poskytli SHMU a Katedra zoológie Prírodovedeckej fakulty Univerzity Komenského v Bratislave. Od SHMU máme k dispozícii priemerné denné teploty vzduchu a priemerné denné hodnoty relatívnej vlhkosti vzduchu v Mlynskej doline namerané v rokoch 2000 až 2007. Od PriFUK máme k dispozícii údaje o hniezdení vtákov v Lamači za roky 2000 až 2005.

6.1 Prezentácia

Na tomto obrázku je prezentácia teplôt nameraných za 8 rokov prostredníctvom *Dvojfarebného ofarbovania*. Okrem samotných dát sa tam nachádza aj text, ktorý popisuje prezentáciu a hodnotová škála pre *Dvojfarebné ofarbovanie*, ktorá umožňuje vyčítať hodnoty dát. Text a všetky vizualizácie sú na seba zakotvené, teda prezentáciu môžeme jednoducho presúvať. Vizualizácie majú zapnuté zobrazenie mriežky a časovej hodnoty mriežky. Veľkosť mriežky je nastavená na jeden rok, teda sa zobrazia len krajné dve čiary.

Na tejto prezentácii môžeme vidieť, že vieme efektívne vizualizovať aj dáta s *Cyklickým časom*.



28. Obr.: Prezentácia

6.2 Analýza

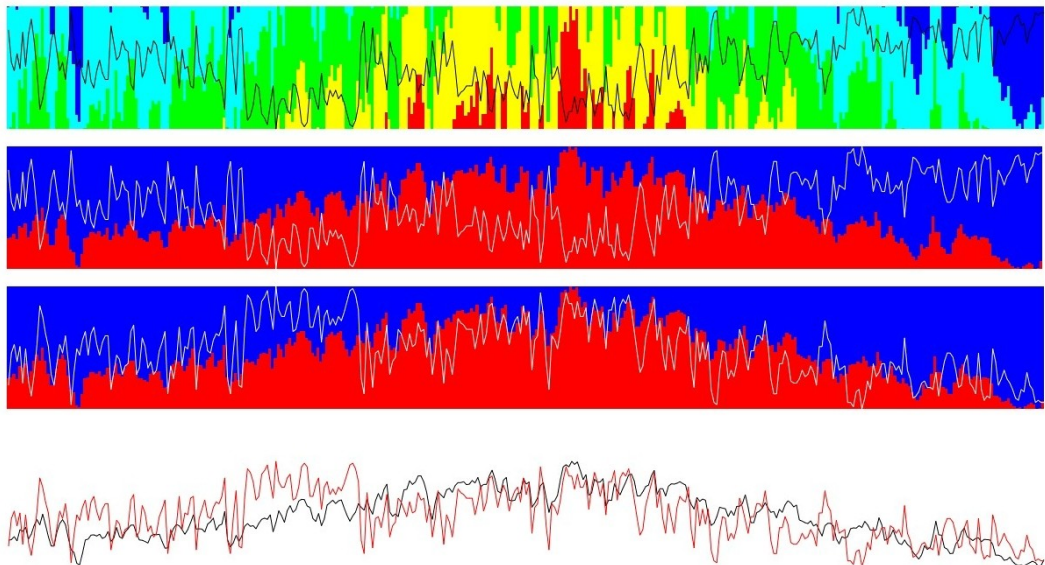
Na tomto obrázku sa snažíme zistiť vzťah medzi teplotou vzduchu a relatívnou vlhkosťou vzduchu za rok 2001. V každej časti sú presne tie isté dáta.

V prvej časti teplotu reprezentuje vizualizácia pomocou *Dvojfarebného ofarbovania* s 5 farbami v hodnotovej škále a vlhkosť reprezentuje *Čiarový graf* ofarbený na čierne. Z tejto vizualizácie sa nedá zistiť či je vzťah medzi týmito dátami.

V druhej časti teplotu reprezentujeme už len pomocou 2 farieb a farbu vlhkosti sme zmenili na bielu, aby sa dala lepšie rozpoznať na červeno-modrom pozadí. Skúsenejší užívateľ si môže všimnúť tam, kde teplota klesá, tak vlhkosť stúpa. Keďže človek najľahšie zaregistruje závislosť dvoch javov ak sa správajú rovnako, otočíme v tretej časti graf vlhkosti aby korešpondoval s grafom teploty.

Z tretieho obrázka je už vidieť, že dané grafy sú príbuzné a teda teplota a vlhkosť vzduchu sú od seba závislé.

Na štvrtom obrázku je ukázané, že namiesto Dvojfarebného ofarbovania sme mohli použiť aj Čiarový graf a mohli sme dospieť k rovnakému výsledku. Je však na užívateľovi, z ktorého obrázku vie ľahšie vyčítať vzor tejto vlastnosti.



29. Obr.: Analýza

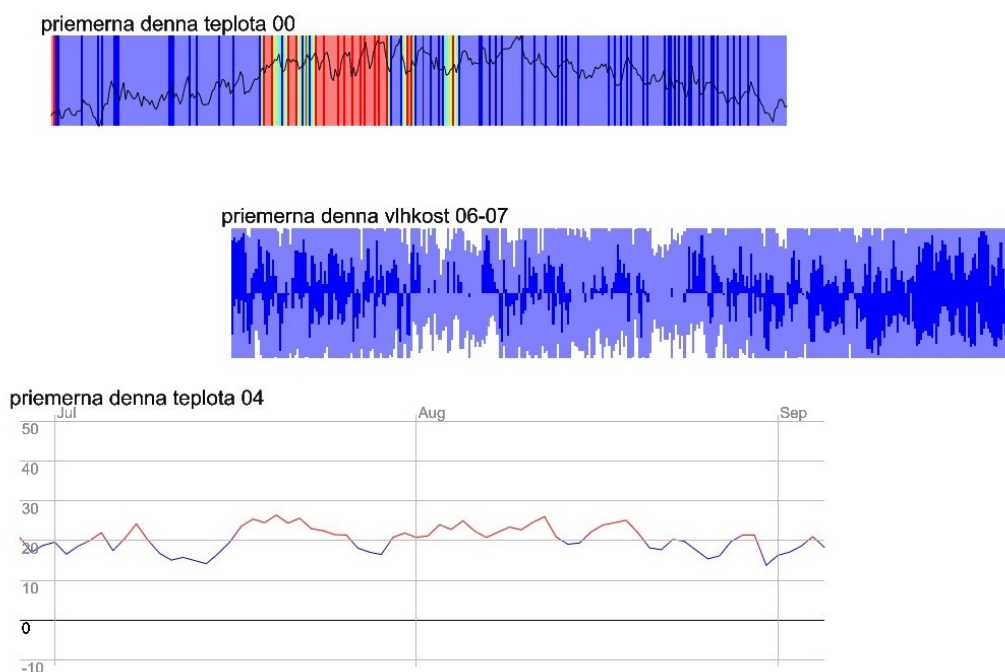
6.3 Prehľad

Na tomto obrázku ukážeme niektoré nastavenia vizualizácií a popíšeme ich význam pri analýze.

Na prvom obrázku je pomocou Čiarového grafu zobrazená denná teplota. *Jednofarebným ofarbením* je znázornené či vták v daný čas hniezdil alebo nie. Keďže informácie o hniezdení nie sú kompletne, null elementy boli dopočítané z krajných hodnôt a vykreslené polo-priesvitne. Z vizualizácie vidíme, že vták hniezdil len v II. štvrtroku za teplého počasia.

Na druhom obrázku porovnávame dennú priemernú vlhkosť vzduchu medzi rokmi 2006 a 2007. Obe dáta sú vykreslené pomocou *Dvojfarebného ofarbenia* rovnakej farby a jedna je otočená okolo x-ovej osi. Vzhľadom na to, že človek ľahko rozpoznáva súvislosť farby, môže sa takáto vizualizácia využiť na hľadanie závislostí medzi dvoma dátami.

Na tretom obrázku máme priemerné denné teploty za rok 2004. Zúžili sme výber dát na Júl a August. Pomocou škálovania osi a vhodného ofarbenia sme zvýraznili hodnoty nad 20°C.



30. Obr.: Prehľad

7 Záver

Primárnym cieľom mojej práce bolo implementovať niekoľko vizualizácií časovo závislých dát a vyvinúť jednoduchý nástroj na analýzu časových dát a prezentovanie výsledkov zistených pri tejto analýze. Tento cieľ som splnil, čo dokazuje priložená aplikácia a obrázky z nej. Vytvorená aplikácia je grafický príjemná a jednoduchá na ovládanie. Podporuje tri metódy vizualizácie časovo závislých dát a pomocou viacerých parametrov vizualizácií umožňuje prispôbiť vizualizácie potrebám používateľa aj z hľadiska analýzy aj z hľadiska prezentácie. Použitím pracovnej plochy využívajúcej metódu „drag and drop“ som pridal nový aspekt pri analýze dát a to porovnávanie jednotlivých vizualizácií, ktoré nastáva pri prekryve ale priložení dvoch vizualizácií vedľa seba. Okrem iného to ponúka jednoduchý nástroj na rozmiestňovanie vizualizácií pre potreby prezentácie výsledkov.

Aplikácia by sa dala rozšíriť o ďalšie metódy vizualizácií, čím by sa stala ešte viac využiteľná aj v bežnej praxi. Pridaním ďalších parametrov vizualizácií by sa ešte viac zväčšila analytická sila aplikácie, nemôže to však byť na úkor jednoduchosti ovládania, keďže práve tento prvok je podstatný pre cieľového používateľa.

Druhý cieľ bol popísať rôzne metódy vizualizácie časovo závislých dát. Porovnal som základné vizualizácie, ktoré sú jednoduché na pochopenie vizuálnej informácie, bez potreby širších znalostí z IT, keďže tie boli prioritné pri implementácii.

Táto práca mi dala veľa znalostí z programovania v C++ s využitím OpenGL a prehľad vo formách vizualizácií časovo závislých dát.

Zoznam použitej literatúry

- [1] http://en.wikipedia.org/wiki/Information_graphics, (27-04-2009)
- [2] N. Gunther, "On the Application of Barycentric Coordinates to the Prompt and Visually Efficient Display of Multiprocessor Performance Data," in Proceedings of Sixth International Conference on Modelling Techniques and Tools for Computer Performance Evaluation, eds. R. Pooley and J. Hillston, 67-80, September, Edinburgh, Scotland, Antony Rowe Ltd., Wiltshire, U.K., 1992
- [3] N. Gunther, http://www.cmg.org/measureit/issues/mit22/m_22_1.html, 2005, (27-04-2009)
- [4] A. M. MacEachren, How Maps Work. The Guilford Press, New York, USA, 1995
- [5] A. U. Frank, Different Types of „Times“ in GIS. Oxford University Press, New York, Oxford, 1998
- [6] W. Müller, H. Schumann, Visualization of Time-dependent Data: Methods & Challenges, Anhalt University of Applied Sciences, 2003
- [7] W. Müller, H. Schumann, Visualization methods for time-dependent data - an overview, Simulation Conference, 2003. Proceedings of the 2003 Winter, 2003
- [8] C. J. Minard, Tableaux graphiques et cartes figuratives. (Collection of graphic works) École , Nationale des Ponts et Chaussées (ENPC). Fol 10975. , 1844-1870
- [9] St. Feiner, C. Beshers, World within Worlds – Metaphors for Exploring n-Dimensional Virtual Worlds, Proc. ACM Symposium on User Interface Software and Technology (UIST '90), 76-83, 1990
- [10] A. Inselberg, Multidimensional Detective. Proc. IEEE Symposium on Information Visualization 1997 (InfoVis '97), 100-107, 1997
- [11] C. Tominski, S. Hadlak, H. Schumann, Two-Tone Pseudo Colored Spiral Display, Java Applet, <http://vcg.informatik.uni-rostock.de/~ct/TTS/TTS.html> (accessed Sept. 2008), 2007
- [12] T. Saito, H.N. Miyamura, M. Yamamoto, H. Saito, Y. Hoshiya, T. Kaseda, Two-Tone Pseudo Coloring: Compact Visualization for One-Dimensional Data, Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on, 2005

- [13] S. Havere, B. Hetzler, L. Nowell, ThemeRiver“ Visualizing Theme Changes over Time. Proc. IEEE Symposium on Information Visualization 2000 (InfoVis '00). IEEE Computer Society, Los Alamitos, USA, 115-123, 2000
- [14] J.V. Carlis, J.A. Konstan, Interactive visualization of serial periodic data, Proc. 11th annual ACM symposium on User interface software and technology, San Francisco, California, USA. 29-38., 1998
- [15] J.J van Wijk, E. van Selow, Cluster and Calendar-based Visualization of Time Series Data, In Proc. IEEE Symposium on Information Visualization (InfoVis '99), IEEE Computer Society, 4-9., 1999
- [16] F. Brian, J. Pritchard, Visualisation of historical events using Lexis pencils, In: Case Studies of Visualization in the Social Sciences [online], Available online via <<http://www.agocg.ac.uk/reports/visual/casestud/contents.htm>> [accessed July 18th , 2003], 1997
- [17] L. Treinish, D. Silver, Worm Plots, IEEE Computer Graphics and Applications, 17: 17-20., 1997
- [18] P. Dragicevic, S. Hout, SpiraClock: A Continuous and Non-Intrusive Display for Upcoming Events , Conference on Human Factors in Computing Systems, Minneapolis, Minnesota, USA, 2002
- [19] C. Tominski, J. Abello, H. Schumann, Axes-Based Visualizations with Radial Layouts, Proc. ACM Symp. Applied Computing, pp. 1242-1247, 2004.
- [20] C. Tominski, J. Abello, H. Schumann, Interactive Poster: Axes-based Visualizations for Time Series Data. Proc. IEEE Symp. on Information Visualization, Seattle, 2003.
- [21] R. Kosara, F. Bendix, H. Hauser, Time histograms for large, time-dependent data. In Proceedings of the 2004 Eurographics/IEEE TVCG Symposium on Visualization, pp. 45–54, 340, 2004

Obsah CD

- Aplikácia
- Zdrojový kódy aplikácie
- Sada dát na vizualizáciu