

Toward approaches to big data analysis for terroristic behavior identification: child soldiers in illegal armed groups during the conflict in Donbas region (East Ukraine)

Yuriy V. Kostyuchenko

Scientific Centre for Aerospace Research of the Earth, National Academy of Sciences of Ukraine

Maxim Yuschenko

Scientific Centre for Aerospace Research of the Earth, National Academy of Sciences of Ukraine

ABSTRACT

Paper aimed to consider of approaches to big data (social network content) utilization for understanding of social behavior in the conflict zones, and analysis of dynamics of illegal armed groups. Analysis directed to identify of underage militants. The probabilistic and stochastic methods of analysis and classification of number, composition and dynamics of illegal armed groups in active conflict areas are proposed. Data of armed conflict – antiterrorist operation in Donbas (Eastern Ukraine in the period 2014-2015) is used for analysis. The numerical distribution of age, gender composition, origin, social status and nationality of child militants among illegal armed groups has been calculated. Conclusions on the applicability of described method in criminological practice, as well as about the possibilities of interpretation of obtaining results in the context of study of terrorism are proposed.

Keywords: child soldiers; illegal armed groups, social networks, data analysis, statistics, crime identification

INTRODUCTION

In criminological practice and research of terrorism, there are many cases requiring application of sophisticated scientific instruments. Not only at the stage of analysis of evidences, but also at the preliminary stages, in particular, at the stage of crime identification.

Description of criminal activity and identification of a crime is a challenge in some cases, for example, in the areas of crisis, conflict, and fighting. This is due to significant limitations of existing information and data available.

In such cases it is necessary to use many different sources of information, including social networks, with adapted statistical approaches to assessment of this data. Correct statistical methods of data collection, analysis, filtering and regularization in such cases are critical.

Obtained with the mathematical methods, robust spatial-temporal distributions of data could be used to define the event and to identify a crime.

Social networks reflect the motivations of the behaviour of different groups of society and varied social environments (Wasserman, Galaskiewicz, 1994). So, because of the large scope, it is a good base for sociometrics and behavior analysis (Krause, Croft, James, 2007). However, formal numerical methods of social network data analysis are still not sufficiently developed for a wide range of important cases, in particular for crisis management (Lerbinger, 2012) and conflict analysis (Scott, 2012).

The task of extraction of structured distributions of data regularized by determining parameters from the large sets of non-structured data is solving. The algorithm based on specifics of dig data distribution, and on data source characteristics (e.g. group behavior). Realization of this approach aimed to detection of stable indicators of criminal and/or terroristic activity. At the same time the analyzed data and cyber activity, producing it might be legal in most cases (if propaganda is beyond our scope).

The proposed algorithm allows to analyze the content of social networks on the base of the set of selected indicators. These indicators allow to control the social dynamics of different social groups represented in social networks and analyze their behavior, including identification of evidences of terrorism.

Assuredly, social media activity in itself is not a terrorist in the strict sense, since it is not a method of achieving of political goals using a direct violence (physical or psychological). However, an important evidence of terrorism is its demonstrative character. The attack requires a nationwide, or ideally a global audience. Therefore, an information component of terroristic activity is extremely important.

In particular, there are several main objectives of the information campaign of terrorist activity. This is propaganda of impotence of central power and calls for the creation of alternative authorities. Second, this is making the precedents of active disobedience and military confrontation with a central power. Third, this is dissemination of appeals to the people to join the active opposition to the authorities, glorification of terrorists, promoting their ideas and lifestyle. Also, positive information about terrorism activates any local power and social mood, the opposed government, including distanced from terrorist tactics before. Additionally, the attack is treated as an indisputable sign of the acute crisis in the society. All this is pushing the society and the power to make concessions to political forces that use terrorist tactics. Terrorism and its propaganda make strikes on the economy, reduces the investment attractiveness of the country, degrades the image of the country, pushing the country to the radicalization of the political course, to authoritarian forms of government (often this evolution is a purpose of terrorists).

Terrorism is the most dangerous and the most effective (by the criterion of the invested resources / result) way to the political destabilization of society. And also this is one of the most effective destabilizing tools of the enemy within the modern hybrid conflicts.

At the same time, the use of terrorist tactics presupposes a set of specific socio-cultural and political characteristics of the target society. Terrorism is a phenomenon inherent to the crisis stage of the modernization of society. Commonly, the completion of modernization transformations removes grounds for terrorism. Terrorism occurs at the boundaries of cultures and periods of history.

Terrorism is an indicator of the crisis processes. This is the "emergency channel" of feedback between society and government, between a separate part of the community and society as a whole. It indicates to the acute trouble in a certain area of social space. In this respect, terrorism have not a pure power, police solutions. Localization and suppression of terrorists is only a part of the job. Another part involves political, social and cultural changes, which should remove the reasons for the radicalization of society and base of the terrorism.

In this context, the analysis of social dynamics and the behavior of social groups is becoming an important factor in the control of terrorist activities. Not only and not so much as for the identification of the criminals, but the identification of the crime, the identification of its social base, to determine the causes and driving forces of terrorism is a task to be solved by approaches proposed in this study.

Case Study

In this paper, using the data on the conflict in some regions of Lugansk and Donetsk regions (Eastern Ukraine) shows the algorithm and method of data collection, filtering and regularization in tasks of child involvement in illegal military groups.

In other words, it shows how to use statistical algorithms in tasks of criminology and terrorism study for decision-making under deep uncertainty.

Antiterrorist operation - the armed conflict in the Donbas (Eastern Ukraine – see Figure 1) extends in spatially distributed urban agglomerations, e.g. with widely available and accessible information infrastructure and communications. The conflict zone is an area about 16,800 square kilometers with varying number of people: from 2.1 to 2.6 million up to 3.1 in maximum, depending on the intensity of fighting, economic situation and the season. In this area is available 990,620 telephone lines, 1,786,300 Internet lines (including optical fiber), there are 1,196,820 Internet users, 3,705,750 mobile phones serviced by 5 providers.

While traditional media functioning under external pressure and censorship, and often using as an "information weapon", social media, first of all, social networks have become an important source of information about the development of conflict. In this case, the question of the correct methods of collecting, processing and analyzing data from social networks is important in the context of analysis of the conflict.



Figure 1. Area of the conflict.

METHODOLOGY

In the conflicts and crises the accurate, timely and adequate information acquiring is a daunting challenge. Usually, parties of the conflict use different tools of censorship, all public information is distorted, and access to the crisis area for observers is usually limited. In such situation, social networks became a good source of operational information.

Psychological features and characteristics of group behavior of social network users have determine the distribution of data and information flows (Lerman, 2013).

Usually, when we consider the conflicts of values, and also in many cases of racial, religious and ethnic conflicts, there are stable sets of motivations of group behavior, presented in the form of symbols. It serves as a justification for violence and are widely used in propaganda. These symbols are the value-markers of certain social groups. In social networks, these symbols are distributed in the form of a hashtag, and are convenient search criteria for gathering information (Sultana, Paul, and Gavrilova, 2015).

In each case, the information contained in social networks is incomplete, with high noise, distributed by interconnected clusters. So, we should apply robust statistical algorithms for collection, filtering and classification of information using a set of interrelated criteria. Furthermore, the classified information should be regularized to obtain spatial-temporal distributions with controlled reliability. As result we obtain a coherent description of a situation in which all data will be statistically related to each other and mutually verified. Only in this case our data will be meaningful and has a sense as evidence, rather than fragmentary information from unknown and unconfirmed sources (Caverlee, Cheng, Sui, et al, 2013).

The first stage is the search and collection of data by identifying hashtags on the area of conflict, location, group membership, age/date of birth.

Data classification approach

The accumulated massive of data should be classified. In other words, a decision making procedure should be applied to each entry of general community - the multidimensional dataset to extract a class of “infant participant of illegal armed groups”. We propose to use for it a data classification approach based on the Bayes rule for minimum classification error in terms of maximum-a-posterior decision task in Markov random field model representation of multi-temporal, multi-source data (Duda, Hart, Stork, 2012).

Let $I_j \{x_1, x_2, \dots, x_N\}$ be a given find record describing a community of illegal armed groups combatants, modeled as a set of N identically distributed n -variate random vectors of identifiers. We assume M classes $\omega_1, \omega_2, \dots, \omega_M$ to be present among a set of combatants and we denote the resulting set of classes by $\Omega = \{\omega_1, \omega_2, \dots, \omega_M\}$ and the class label of the k -th community member ($k=1, 2, \dots, N$)

by $s_k \in \Omega$. By operating in the context of community classification, we assume that verified events, recognized as a training set to be available, and we denote the index set of the training community member by $T \subset \{1, 2, \dots, N\}$ and the corresponding true class label of the k -th training community member ($k \in T$) by s_k^* .

When collecting all the feature vectors of the N member of community of combatants in a single ($N \cdot n$)-dimensional column vector $X = \text{col}[x_1, x_2, \dots, x_N]$ and all the community member labels in a discrete random vector $S = (s_1, s_2, \dots, s_N) \in \Omega^N$, the Bayes rule approach (Duda, Hart, Stork, 2001) assigns to the community of combatants data X the label vector \tilde{S} , which maximizes the joint posterior probability $P(S | X)$:

$$\tilde{S} = \arg \max_{S \in \Omega^N} P(S | X) = \arg \max_{S \in \Omega^N} [p(X | S)P(S)] \quad (1)$$

where $p(X | S)$ and $P(S)$ are the joint probability density function (PDF) of the global feature vector X conditioned to the label vector S and the joint probability mass function (PMF) of the label vector itself, respectively. The Markov random field (MRF) approach offers a computationally tractable solution to this maximization problem by passing from a global model for the statistical dependence of the class labels to a model of the local general community properties, defined according to a given neighborhood system (Geman, S., Geman, D., 1993). Specifically, for each k -th member of general community, a neighborhood $N_k \subset \{1, 2, \dots, N\}$ is assumed to be defined, such that, for instance, N_k includes the four (first-order neighborhood) or the eight (second-order neighborhood) member surrounding the k -th member ($k=1, 2, \dots, N$). More formally, a neighborhood system is a collection $\{N_k\}_{k=1}^N$ of subsets of member such that each member is outside its neighborhood (i.e., $k \notin N_k \forall k = 1, 2, \dots, N$) and neighboring member are always mutually neighbors (i.e., $k \in N_h$ if and only if $h \in N_k \forall k, h = 1, 2, \dots, N, k \neq h$). This simple discrete topological structure attached to the general community data is exploited in the Markov random field framework to model the statistical relationships between the class labels of spatially distinct member and to provide a computationally affordable solution to the classification problem of (1). Specifically, we assume the feature vectors x_1, x_2, \dots, x_N to be conditionally independent and identically distributed with probability density function $p(x | s)$ ($x \in \mathfrak{R}^n, s \in \Omega$), that is (Geman, S., Geman, D., 1993):

$$p(X | S) = \prod_{k=1}^N p(x_k | s_k) \quad (2)$$

and the joint prior probability mass function $P(S)$ to be a Markov random field with respect to the abovementioned neighborhood system. The probability distribution of each k -th general community label, conditioned to all the other general community labels, is equivalent to the distribution of the k -th label conditioned only to the labels of the neighboring members ($k=1, 2, \dots, N$):

$$P\{s_k = \mathbf{a}_i | s_h : h \neq k\} = P\{s_k = \mathbf{a}_i | s_h : h \in N_k\}, i = 1, 2, \dots, N \quad (3)$$

The probability mass function of S is a strictly positive function on Ω^N , so, $P(S) > 0 \forall S \in \Omega^N$.

The Markov assumption expressed by (3) allows restricting the statistical relationships among the general community labels to the local relationships inside the predefined neighborhood, thus greatly simplifying the contextual model for the label distribution as compared to a generic global model for the joint probability mass function of all the general community labels.

Basing on described approach, the Ho–Kashyap method (Ho, Kashyap, 1965) has been applied to classification of the general community.

As the result of classification we obtain a dataset with all records that meet the specified condition. For example, distribution of members of community “infant participant of illegal armed groups” with

age, sex, social status, accessory, spatial and temporal marks inside the general community of members of illegal armed groups' combatants.

Data regularization algorithm

Further, we need regularized spatial-temporal distribution of classified records. It is necessary to avoid duplication and fix falsification, inter-verify data, trace robust trends in the distributions, so get a basis for interpretation of the data. We propose to use a two-stage procedure of data regularization.

The method proposed is based on non-linear kernel-based principal component algorithm (KPCA) modified according to specific of data. Using this method the set of selected records has been analyzed. The robust technique of data regularization for normalization of data reliability is proposed. The technique utilizes data from different sources, different nature, and with different metrics. This approach allows to calculate regularized distributions in units invariant toward data properties and quality. We can analyze simultaneously different types of data, regardless spatial and temporal scales and heterogeneities using this approach.

Correct statistical analysis requires the set of data x_i with controlled reliability (above mentioned "training set"), which reflects distribution of investigated parameters during whole observation period (taking into account variances of reliability of data x_i). Set of data x_i ($x_i \in R^m$) consists of multi-source data, including data with sufficient reliability x_j ($x_j \in R^m$), where $j = 1, \dots, N$. Problem of determination of controlled quality and reliability spatial-temporal distribution of investigated parameters x_i might be solved in framework of tasks of multivariate random processes analysis and multidimensional processes regularization (Kostyuchenko, Movchan, Kopachevsky, et al, 2015).

Required regularization may be provided by 2-stage procedure. If we able to formulate stable hypothesis on distribution of reliability of data in the framework of defined problem we may to propose relatively simple way to determine investigated parameters distributions $x_i^{(x,y)}$ towards distributions on measured sites x_i^m basing on (Kostyuchenko, Movchan, 2015):

$$x_i^{(x,y)} = \sum_{m=1}^n w_{x,y}(\tilde{x}_i^m) x_i^m \quad (4)$$

where weighting coefficients $w_{x,y}(\tilde{x}_i^m)$ determined as:

$$\min \left\{ \sum_{m=1}^n \sum_{x_i^m \in R^m} w_{x,y}(\tilde{x}_i^m) \left(1 - \frac{x_i^m}{\tilde{x}_i^m}\right)^2 \right\} \quad (5)$$

according to (Kostyuchenko, Movchan, 2015). Here m – number of records; n – number of sources/series; x_i^m – distribution of data; R^m – set (aggregate collection) of data; \tilde{x}_i^m - mean distribution of searching parameters.

This is the simple way to obtain a regular spatial distribution of analyzed parameters, on which we can apply further analysis, in particular temporal regularization. At the same time, this is the first stage of regularization. This algorithm may be interpreted as the general form of Kolmogorov regularization procedure (Ermoliev, Makowski, Marti, 2012).

Further, second stage of regularization should take into account both data distribution temporal non-linearity (caused by imperfection of available multi-source statistics) and features of temporal-spatial heterogeneity of data distribution caused by systemic complexity of studied community. According to (Kostyuchenko, Movchan, 2015) the kernel based non-linear approaches are quite effective for analysis of such types of distributions.

Proposed method is based on modified kernel principal component analysis (KPCA). In the framework of this approach the algorithm of non-linear regularization might be described as following rule (Kostyuchenko, Movchan, Kopachevsky, et al, 2015):

$$x_i = \sum_{i=1}^N \alpha_i \tilde{k}_t(x_i, x_i) \quad (6)$$

In equation (6) the coefficients α selected according to optimal balance of relative validation function and covariance matrix, for example as (Kostyuchenko, 2015):

$$C^F v = \frac{1}{N} \sum_{j=1}^N \Phi(x_j) \Phi(x_j)^T \cdot \sum_{i=1}^N \alpha_i \Phi(x_i) \quad (7)$$

Where non-linear mapping function of input data distribution Φ determined as:

$$x_i = \sum_{i=1}^N \alpha_i \tilde{k}_t(x_i, x_i) \quad (8)$$

$$\sum_{k=1}^N \Phi(x_k) = 0 \quad (9)$$

And \tilde{k}_t - is mean values of kernel-matrix $\mathbf{K} \in R^N$ ($[\mathbf{K}]_{ij} = [k(\mathbf{x}_i, \mathbf{x}_j)]$). Vector components of matrix determined as $\mathbf{k}_i \in R^N$; $[\mathbf{k}_i]_j = [k_t(\mathbf{x}_i, \mathbf{x}_j)]$. Matrix calculated according to modified rule: $\mathbf{k}_t(\mathbf{x}_i, \mathbf{x}_i) = \left\langle \prod_{j,t}^{\rho} (1 - \prod_{j,i}^{\rho})^{x_j} \right\rangle$, where ρ – empirical parameters, selected according to the classification model of study community.

Using described algorithm, it is possible to obtain regularized spatial-temporal distribution of investigating parameters over the whole observation period with rectified reliability and controlled uncertainty (Kostyuchenko, 2015; Kostyuchenko, 2016). As the result we obtain a regular spatial-temporal distribution of child soldiers registered, prepared for the interpretation.

DATA AND RESULTS

By the requested hashtag over 21,500 profiles in social networks (VK, Facebook, Instagram, Twitter) in five languages (Russian, Ukrainian, English, Serbian and German) has been analyzed. Following hashtags were used: #донбасс(Donbas), #новороссия(NovoRossia), #юныегероиновороссии(NovoRossiaYoungHeroes) #героиновороссии(NovorossiaHeroes) #память(Memory). More than 480 000 entries had satisfied to determined criteria. After filtering by age (12-18 years) / date of birth (after 1998) and of belonging to an illegal military group, has been collected about 5400 entries.

A classification procedure (1)-(3) has been applied to these entries. After it data has been regularized by a two-stage procedure (4)-(9).

The used method allowed to obtain an overview of distribution of key characteristics of illegal armed groups in the conflict zone. There were identified from 40 to 60 separated armed groups during the observation period. All of them are significantly varied by weapons, equipment, ideology (Soviet patriots, Russian imperialists, neo-Stalinists, anarchists, neo-Nazi, Christian orthodox extremists), and united only anti-Western ideas and the pro-Russian rhetoric.

They acted quite independently. The opinion of the nominal leaders of the region determined their actions no more that they are dependent on the supply of arms, ammunition, fuel and food from Russia

via the central leadership. However, to the end of the observation period, the number of autonomous groups has decreased, and their centralization is increased.

At the beginning of 2016 a number of armed groups was about 30. During the active phase of conflict in April 2014 – December 2015 the total number of illegal armed group members ranged from 19.5 to 52 thousand, and to the end of 2015 it amounted to 35.7 thousand. Noticeable is heterogeneity and uncertainty of this population: more than 60% of the militants are citizens of Russia and other countries, about 15% are Russian military personnel.

For a detailed study of the general population, it is necessary to use additional tools based on stochastic methods of assessment (Ermoliev, Makowski, Marti, 2012). This makes possible to analyze distributions of socioeconomic parameters of members of illegal armed groups.

In this case, we limited ourselves to a narrow concrete case of child soldiers. Calculated distributions of the study parameters are demonstrating following results.

Data were obtained based on the information resources of local communities, their supporters groups, homepages of militants and their families in social networks, and media publications. This allows us to estimate the number of children, their distribution and dynamics, age and gender characteristics, and in most cases, to identify their origin and identity, including determination of residency and schools, to restore the history of their participation in armed conflict.

First case of involvement of children into illegal armed group was detected in town Slovyansk in end of April 2014. Underage personnel were recruited into pro-Russian troop headed by Igor Stryelkov (a.k.a. Girkin), who captured the town. Then, during the fighting in Slovyansk in June 10-30, 2014, the first 2 child soldiers were killed and a few kids were injured. The estimated (by calculation using procedures (1)-(3), (6)) total number of people under the age of 18, who was involved in the conflict as military personnel on the side of illegal groups, during the events in Slovyansk is about 35 - 40.

Later, during 2014 the number of children in the illegal military groups gradually increased. Almost every group had underage participants, which in many cases were involved into the fighting.

The total number of underage soldiers in terrorist groups during the conflict in Donbas in the period April 2014 - January 2015 could be estimated at 210-250 persons. Their age ranges from 13-18, mostly 16-17 years, but registered even cases of 12-year-old kids (Fig.2).

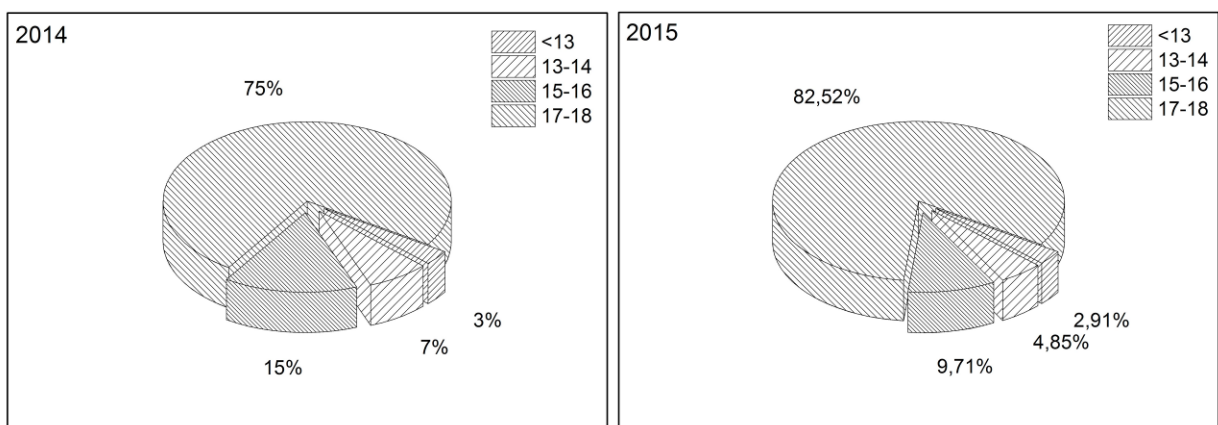


Figure 2. Distribution of child soldiers by the age.

Most (90%) underage members of illegal armed groups in 2014 and in 2015 are boys with an insignificant tendency to increase their part (Fig.3).

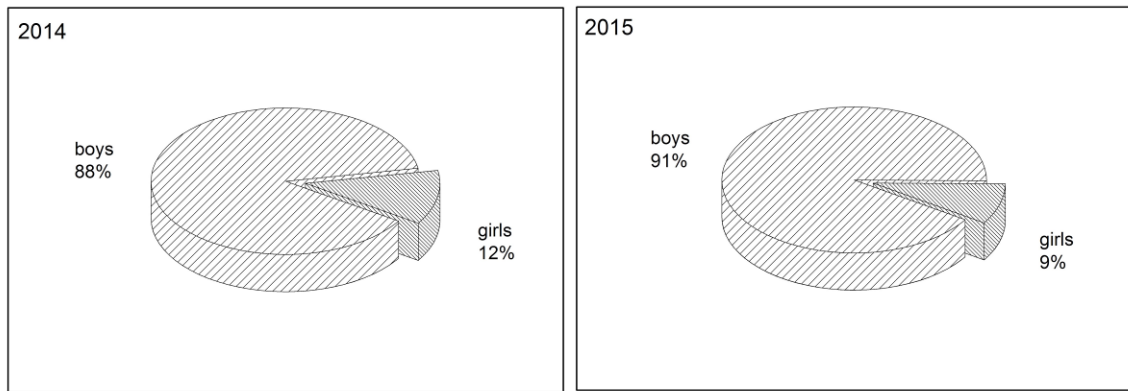


Figure 3. Distribution of child soldiers by the gender.

In 2015 the number of child soldiers began to decline. The reason for this is likely to organizing the chaotic terrorist groups and reducing its support from Russian Federation. However, local children continue to cooperate with militants using them as a source of income during the economic degradation of the territory. During 2014-2015 the direct subsidy from Russia of conflict region was about 88% of their total budget, while local economic activity (including illegal) accounted for only 12%. At the same time in 2015 losses of underage soldiers are increasing.

Question of involvement of children in the illegal armed groups is also important. In 2014, especially in the period April-October, the majority (80%) of underage soldiers were kids of members of these groups (Fig.4). Their parents (mainly fathers) were also detected among the members of armed groups. About 10% were children who have no parental care (registered by the regional social service or with permanent residency in other localities), and 10% - children from local families who accidentally joined armed groups (kids with local addresses and registered in local schools). Ways of recruitment of children are usually the military-patriotic and sport clubs where children were indoctrinated by specific pro-Russian and pro-Soviet ideas. Soviet symbols of WWII were widely used in propaganda among child.

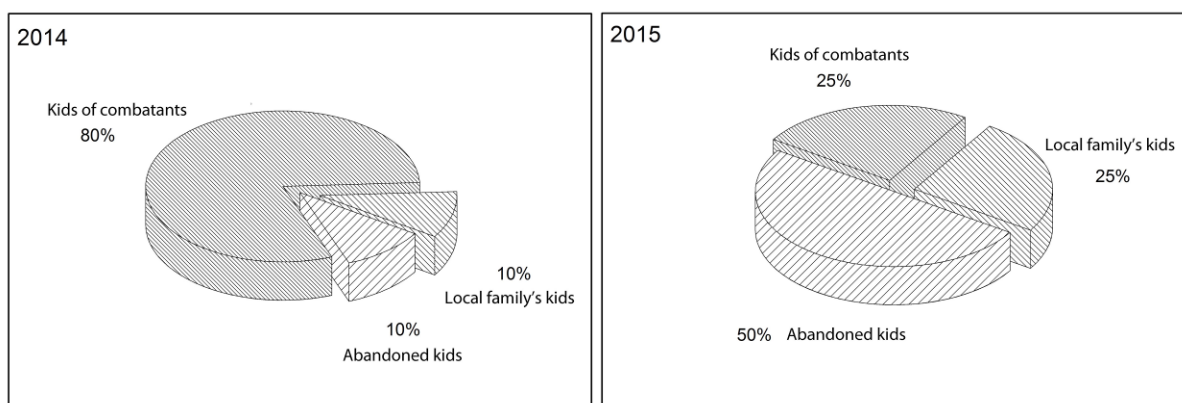


Figure 4. Distribution of child soldiers by the social status.

In 2015 the situation is changing: the proportion of kids of members of armed groups is reduced to 25%, and the proportion of children without parental care increased to 50%. The number of children from ordinary families also increased to 25%. This is due to changes both in the structure of local communities and in the economy of the territory.

In 2015 at the background of the observed influx of mercenaries from Russia, the number of local residents among illegal armed groups as well as local child soldiers is decreased (Fig.5).

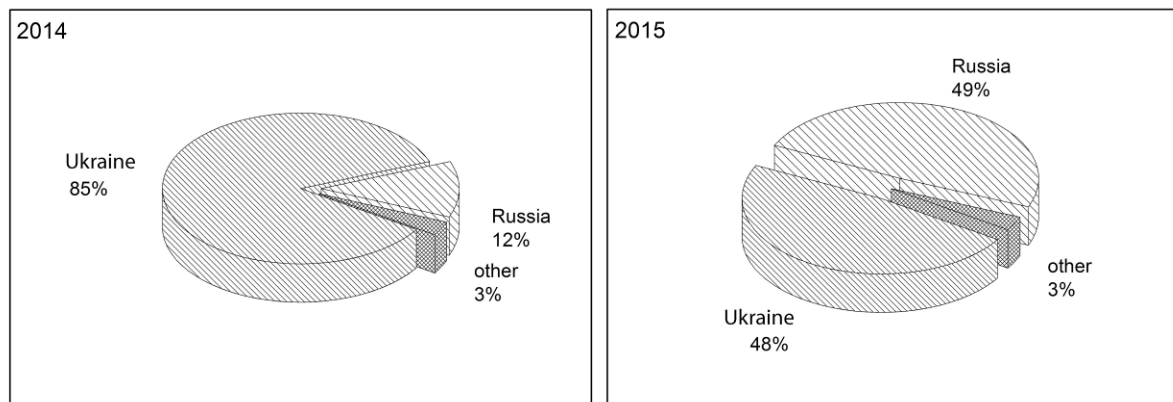


Figure 5. Distribution of child soldiers by the nationality.

With a significant reduction in intensity of fighting, over 2015 children are mainly used by militants for intelligence activity, housekeeping and security functions.

The total number of underage soldiers at the end of 2015 in the Russian-supported illegal armed groups could be estimated at 70-90 persons. This number is not constant, children who are local residents, are not a part of the group permanently. The most stable groups of children - up to 50% of them - are Russian citizens.

Thus, as the data demonstrate, for the period 2014-2015 during the armed conflict in the Donbas (Eastern Ukraine) on the side of the Russian-supported illegal armed groups about 300 underage soldiers (under 18 years) has been involved. About 150-180 child soldiers directly participated in the combats. Up to 50 minors were killed during the fighting.

Estimations, based on calculation procedures (1)-(3) and (4)-(9), are corresponding to the assessments of other sources, such as The 2016 Trafficking in Persons Report.

CONCLUSIONS AND DISCUSSION

Social networks are an important and valuable source of information in conflict situations. The social network data analysis could be an important instrument in criminology, conflict research and terrorism study.

However, the specific of big data requires application of correct techniques of data collection processing and analysis. In this context, it is important not only to correctly collect and filter the data, but also to apply the regularization procedure, which will provide regular spatially-temporally distributed data, mutually verified, with controlled reliability.

For data filtering and classification has been proposed an approach based on the Bayes rule for minimum classification error in terms of maximum-a-posterior decision task in Markov random field model representation of multi-temporal, multi-source data. As the result of classification we obtain a dataset with all records that meet the specified condition. For collected, filtered and classified data regularization it has been proposed non-linear kernel-based principal component algorithm (KPCA) based method, modified according to specific of data. Using this algorithm was obtained regularized spatial-temporal distributions of investigating parameters over the whole observation period with rectified reliability and controlled uncertainty, prepared for the interpretation.

Proposed in this paper methodological framework requires some a priori information about the studied distributions. Consequently, methods based on this methodology require regional adaptations.

Thus, application of the proposed method resulting to distribution, which cannot be obtained by another way, at least during the active conflict. We can recognize a calculated data as unique, reliable, and meaningful and has a sense as evidence, rather than fragmentary information from unknown and unconfirmed sources.

Using of children as soldiers and military personnel is a widespread practice of criminal and terrorist regimes worldwide. It is clearly condemned by the international community and recognized as a war crime. According to international law (in particular, Article 77.2 of the Additional Protocol I to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts, 1977, and the Optional Protocol to the Convention on the Rights of the Child on the

involvement of children in armed conflict, 2002), the participation of children under 15 years in army is prohibited, as well as the recruitment of persons under 18 years to paramilitary and guerrilla nongovernmental groups with any goals. Actually, child soldiers are separate, independent indicator of terrorism. According to international statistics, more than 300,000 children and teenagers were used in the last 10 years as a soldier, a partisan military personnel to support the armed forces in the world. The most famous historical cases of mass use of minors in armed conflict are a children's group "The Liberation Tigers of Tamil Eelam" in Sri Lanka, the young militants of Hamas and "Islamic Jihad" in the Palestinian Authority, "Lord's Resistance Army" in Uganda, left-wing guerrilla groups in Colombia and other.

Unfortunately, in the framework of the conflict in Donbas, the practice of using children as soldiers and military personnel also widely used by Russia supported illegal armed groups mixed from Russian invaders and local collaborators. Mass systematic use of persons under 18 as soldiers and military personnel during the conflict proves that we are dealing with a classic terrorist organization, controlled, managed and supported by Russia.

ACKNOWLEDGMENT

The authors are deeply grateful to the anonymous referees for constructive suggestions that resulted in important improvements to the paper, to colleagues from the American Statistical Association (ASA) for their critical and constructive comments and suggestions.

REFERENCES

- Caverlee, J., Cheng, Z., Sui, D. Z., & Kamath, K. Y. (2013). Towards Geo-Social Intelligence: Mining, Analyzing, and Leveraging Geospatial Footprints in Social Media. *IEEE Data Eng. Bull.*, 36(3), 33-41.
- Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons
- Ermoliev, Y., Makowski, M. and Marti, K. (2012) *Managing Safety of Heterogeneous Systems*. Springer, Heidelberg. ISBN 978-3-642-22883-4
- Geman, S., & Geman, D. (1993). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *Journal of Applied Statistics*, 20(5-6), 25-62.
- Ho, Y. C., & Kashyap, R. L. (1965). An algorithm for linear inequalities and its applications. *IEEE Transactions on Electronic Computers*, 5(EC-14), 683-688.
- Kostyuchenko Yu.V. (2015) "Geostatistics and remote sensing for extremes forecasting and disaster risk multiscale analysis" in: S. Kadry and A. El Hami (eds.), *Numerical Methods for Reliability and Safety Assessment: Multiscale and Multiphysics Systems*, Springer International Publishing Switzerland, 2015, XII, 805 p. 328 il., 404-423, DOI 10.1007/978-3-319-07167-1_16, ISBN 978-3-319-07166-4
- Kostyuchenko Yu.V. (2016) *Risk Perception Based Approach to Analysis of Social Vulnerability in: Risk Perception: Theories and Approaches*. eds by Theodore Spencer, NY, Nova Publ., pp. 17-51, ISBN: 978-1-63484-623-3
- Kostyuchenko Yu.V., Movchan D. (2015) Quantitative parameter of risk perception: can we measure a geoethic and socio-economic component in disaster vulnerability? in: Peppoloni, S. & Di Capua, G. (eds) *Geoethics: the Role and Responsibility of Geoscientists*. Geological Society, London, Special Publications, 419. – 2015, First published online Feb 16, 2015, <http://dx.doi.org/10.1144/SP419.10>
- Kostyuchenko Yu.V., Movchan D., Kopachevsky I., Bilous Yu. (2015) *Robust Algorithm of Multi-Source Data Analysis for Evaluation of Social Vulnerability in Risk Assessment Tasks // Proc. of SAI IntelSys 2015*, Nov 10-11, IEEEXplore, 2015, pp. 944- 949, London, UK, doi 978-1-4673-7606-8/15
- Krause, J., Croft, D. P., & James, R. (2007). Social network theory in the behavioural sciences: potential applications. *Behavioral Ecology and Sociobiology*, 62(1), 15-27.
- Lerbinger, O. (2012). *The crisis manager*. Routledge.
- Lerman, K. (2013). *Social Informatics: Using Big Data to Understand Social Behavior*. In *Handbook of Human Computation* (pp. 751-759). Springer New York.

Scott, J. (2012). *Social network analysis*. Sage.

Sultana, M., Paul, P. P., & Gavrilova, M. (2015). Social Behavioral Biometrics: An Emerging Trend. *International Journal of Pattern Recognition and Artificial Intelligence*, 29(08), 1556013.

The 2016 Trafficking in Persons Report (2016) U.S. Department of State Publications, Office of the Under Secretary of Civilian Security, Democracy and Human Rights, June 2016, 422p.

Wasserman, S., & Galaskiewicz, J. (Eds.). (1994). *Advances in social network analysis: Research in the social and behavioral sciences* (Vol. 171). Sage Publications.